Quantifying Bell non-locality with the trace distance

S. G. A. Brito, B. Amaral, and R. Chaves

1International Institute of Physics, Federal University of Rio Grande do Norte, 59078-970, P. O. Box 1613, Natal, Brazil
2Departamento de Física e Matemática, CAP - Universidade Federal de São João del-Rei, 36.420-000, Ouro Branco, MG, Brazil
(Dated: September 14, 2017)

Measurements performed on distant parts of an entangled quantum state can generate correlations incompatible with classical theories respecting the assumption of local causality. This is the phenomenon known as quantum non-locality that, apart from its fundamental role, can also be put to practical use in applications such as cryptography and distributed computing. Clearly, developing ways of quantifying non-locality is an important primitive in this scenario. Here, we propose to quantify the non-locality of a given probability distribution via its trace distance to the set of classical correlations. We show that this measure is a monotone under the free operations of a resource theory and that furthermore can be computed efficiently with a linear program. We put our framework to use in a variety of relevant Bell scenarios also comparing the trace distance to other standard measures in the literature.

I. INTRODUCTION

With the establishment of quantum information science, the often called counter-intuitive features of quantum mechanics such as entanglement and non-locality have been raised to the status of a physical resource that can be used to enhance our computational and information processing capabilities in a variety of applications. To that aim, it is of utmost importance to devise a resource theory to such quantities, not only allowing for operational interpretations but as well for the precise quantification of resources. Given its ubiquitous importance, the resource theory of entanglement is arguably the most well understood and vastly explored and for this reason has become the paradigmatic model for further developments.

As discovered by John Bell, one of the consequences of entanglement is the existence of quantum non-local correlations, that is, correlations obtained by local measurements on distant parts of a quantum system that are incompatible with local hidden variable (LHV) models. In spite of their close connection, it has been realized that entanglement and non-locality refer to truly different resources, the most striking demonstration given by the fact that there are entangled states that can only give rise to local correlations. In view of that and the wide applications of Bell’s theorem in quantum information processing, several ways of quantifying non-locality have been proposed. However, only recently a proper resource theory of non-locality has been developed thus allowing for a formal proof that previously introduced quantities indeed provide good measures of non-local behavior. Importantly, different measures will have different operational meanings and do not necessarily have to agree on the ordering for the amount of non-locality. For instance, a natural way to quantify non-locality is the maximum violation of a Bell inequality allowed by a given quantum state. However, we might also be interested in quantifying the non-locality of a state by the amount of noise (e.g., detection inefficiencies) it can stand before becoming local. Interestingly, these two measures can be inversely related as demonstrated by the fact that in the CHSH scenario the resistance against detection inefficiency increases as we decrease the entanglement of the state (also reducing the violation of CHSH inequality).

Bell’s theorem is a statement about the incompatibility of probabilities predicted by quantum mechanics with those allowed by classical theories. Thus, it seems natural to quantify non-locality using standard measures for the distinguishability between probability distributions, the paradigmatic example being the trace distance. Apart from being a valid distance in the space of probabilities, it also has a clear operational interpretation. However, and somehow surprisingly, apart from the exploratory results in, to our knowledge an in-depth analysis of the trace distance as a quantifier of non-locality has never been presented before.

That is precisely the gap we aim to fill in this paper. In Sec. we describe the scenario of interest and propose a novel quantifier for non-locality based on the trace distance. Further, in Sec. we show how our measure can be evaluated efficiently via a linear program and in Sec. we show that it is a valid quantifier by employing the resource theory presented in. We then apply our framework for a variety of Bell scenarios in Sec. including bipartite as well as multipartite ones. In Sec. we discuss the relation of the trace distance with other measures of non-locality. Finally, in Sec. we discuss our findings and point out possible venues for future research.
II. SCENARIO

We are interested in the usual Bell scenario setup where a number of distant parts perform different measurements on their shares of a joint physical system. Without loss of generality, here we will restrict our attention to a bipartite scenario (with straightforward generalizations to more parts, see Sec. [V]) where two parts, Alice and Bob, perform measurements labeled by the random variables \( X \) and \( Y \) obtaining measurement outcomes described by the variables \( A \) and \( B \), respectively (see Fig. 1).

A central goal in the study of Bell scenarios is the characterization of what are the distributions \( p(a, b|x, y) \) compatible with a classical description based on the assumption of local realism implying that

\[
p_C(a, b|x, y) = \sum_\lambda p(\lambda) p(a|x, \lambda) p(b|y, \lambda). \tag{1}
\]

All the correlations between Alice and Bob are assumed to be mediated by common hidden variable \( \lambda \) that thus suffices to compute the probabilities of each of the outcomes, that is, \( p(a|x, y, b, \lambda) = p(a|x, \lambda) \) (and similarly for \( b \)).

The central realization of Bell’s theorem [14] is the fact that there are quantum correlations obtained by local measurements \( (M_A^x \text{ and } M_B^y) \) on distant parts of a joint entangled state \( \rho \), that according to quantum theory are described as

\[
p_Q(a, b|x, y) = \text{Tr} \left[ \left( M_A^x \otimes M_B^y \right) \rho \right], \tag{2}
\]

and cannot be decomposed in the LHV form (11). Moreover, even more general set of correlations, beyond those achievable by quantum theory and called non-signalling (NS) correlations, can be defined. NS correlations are defined by the linear constraints

\[
\begin{align*}
p(a|x) &= \sum_b p(a, b|x, y) = \sum_b p(a, b|x, y') \quad (3) \\
p(b|y) &= \sum_a p(a, b|x, y) = \sum_a p(a, b|x', y),
\end{align*}
\]

that is, as expected from their spatial distance, the outcome of a given part is independent of the measurement choice of the other. The set of classical correlations \( C_C \) (those compatible with (11)) is a strict subset of the quantum correlations \( C_Q \) (compatible with (2)) that in turn is a strict subset of \( C_{\text{NS}} \) (compatible with (3)).

Suppose we are given a probability distribution and want to test if it is non-local or not, that is, whether it admits a LHV decomposition (11). The most general way of solving that is resorting to a linear program (LP) formulation. First notice that we can represent a probability distribution \( q(a, b|x, y) \) as a vector \( q \) with a number of components given by \( n = |x||y||a||b| \) (\(| \cdot | \) representing the cardinality of the random variable). Thus, (11) can be written succinctly as \( p_c = A \cdot \lambda \), with \( \lambda \) being a vector with components \( \lambda_i = p(\lambda = i) \) and \( A \) being a matrix indexed by \( i \) and \( j = (x, y, a, b) \) (a multi-index variable) with \( A_{ij} = \delta_{a_i a_j} \delta_{b_i b_j} \) (where \( f_a \) and \( f_b \) are deterministic functions). Thus, checking whether \( q \) is local amounts to a simple feasibility problem that can be written as the following LP:

\[
\begin{align*}
\min_{\lambda \in \mathbb{R}^n} & \quad v \cdot \lambda \\
\text{subject to} & \quad q = A \cdot \lambda \\
& \quad \lambda_i \geq 0 \\
& \quad \sum_i \lambda_i = 1,
\end{align*}
\]

where \( v \) represents an arbitrary vector with the same dimension \( m = |x||a||y||b| \) as the vector representing the hidden variable \( \lambda \).

The measure we propose to quantify the degree of non-locality is based on the trace distance between two probability distributions \( q = q(x) \) and \( p = p(x) \):

\[
D(q, p) = \frac{1}{2} \sum_x |q(x) - p(x)|. \tag{5}
\]

The trace distance is a metric on the space of probabilities since it is symmetric and respect the triangle inequality. Furthermore, it has a clear operational meaning since

\[
D(q, p) = \max_S |q(S) - p(S)|. \tag{6}
\]
where the maximization is performed over all subsets $S$ of the index set $\{x\}$. That is, $D(q, p)$ specifies how well the distributions can be distinguished if the optimal event $S$ is taken into account. Consider for instance distributions $p(x)$ and $q(x)$ for a variable $X$ assuming $d$ values as $x = 1, \ldots, d$ and such that $p(x) = \delta_{x,1}$ and $q(x) = \frac{1}{d-1}(1 - \delta_{x,1})$. In this case $D(q, p) = 1$, also meaning that $q$ and $p$ can be perfectly distinguished in a single shot since if we observe $x = 1$, we can be sure to have $p(x)$ (or $q(x)$ otherwise).

In our case, we are interested in quantifying the distance between the probability distribution generated out of a Bell experiment and the closest classical probability distribution under test and the set of local correlations. Geometrically, it can be understood (see Fig. 2) as how far we are from the local polytope defining the correlations. Therefore, the more we violate a Bell inequality the higher it will be its value. However, as will be further discussed in Sec. [VI], the violation of a given Bell inequality will in general only provide a lower bound to its value.

III. LINEAR PROGRAM FORMULATION

Given a distribution $q = q(a, b| x, y)$ of interest, in order to compute $\text{NL}(q)$ we have to solve the following optimization problem

$$\min_{\lambda \in \mathbb{R}^n} \| q - A \cdot \lambda \|_{\ell_1} \tag{8}$$

subject to $\lambda \geq 0$

$$\sum_i \lambda_i = 1.$$

First notice that the $\ell_1$ norm of a vector $p$ ($p \in \mathbb{R}^n$, with components $p_i$)

$$\|p\|_{\ell_1} = \sum_i |p_i| \tag{9}$$

Using that, we can rewrite our problem as a linear program

$$\min_{t \in \mathbb{R}, \lambda \in \mathbb{R}^n} \langle 1_n, t \rangle \tag{11}$$

subject to $-t \leq q - A \cdot \lambda \leq t$

$$\sum_i \lambda_i = 1 \quad \lambda \geq 0,$$

where $q$ is a known vector of probability distribution to which we want to quantify the non-locality. This way, given an arbitrary distribution of interest we can compute, in an efficient manner, $\text{NL}(q)$.

Alternatively, we might be interested not on the full distribution but simply on a linear function of it, for example, the violation of a given Bell inequality in the form $I_{\text{Bell}} \cdot q = c$. In this case, further linear constraints need to be added to the LP such as normalization and the fact that the distribution is non-signalling (see Sec. [V]).
for examples):  

\[
\begin{align*}
\min_{t \in \mathbb{R}^n, \lambda \in \mathbb{R}^n, q \in \mathbb{R}^n} & \langle I_n, t \rangle \\
\text{subject to} & \quad -t \leq q - A \cdot \lambda \leq t \\
& \quad \sum_{i} \lambda_i = 1 \\
& \quad I_{\text{Bell}} \cdot q = c \\
& \quad \sum_{a,b} q(ab|xy) = 1 \\
& \quad \sum_{a} q(ab|xy) - \sum_{a} q(ab|x'y) = 0 \quad \forall (b,y) \\
& \quad \sum_{b} q(ab|xy) - \sum_{b} q(ab|xy') = 0 \quad \forall (a,x) \\
& \quad \lambda \geq 0 \\
& \quad q \geq 0.
\end{align*}
\]  

Notice that, instead of adding only NS constraints, one could also be interested in imposing quantum constraints \([\alpha, \beta] \leq 11\). However, in this case we would need to resort to a semi-definite program (that asymptotically converges to the quantum) instead of a linear program.

Finally, often we might be interested in having an analytical rather than numerical tool. To that aim we can rely on the dual of the LP \([11]\) and \([12]\). We refer the reader to \([22]\) for a very detailed account of the dualization procedure but, in short, the optimum solution of \([11]\) and \([12]\) is achieved in one of the extremal points of the convex set defined by the dual constraints. This way, being able to compute such extremal points \(v_i\), we have an analytical solution, valid for arbitrary test distributions \(q\), given by

\[
\text{NL}(q) = \max_{v_i} q \cdot v_i. \tag{13}
\]

\[\text{IV. PROVING THAT TRACE DISTANCE IS A NON-LOCALITY MEASURE} \]

A resource theory provides a powerful framework for the formal treatment of a physical property as a resource, enabling its characterization, quantification and manipulation \([3, 4, 5]\). Such a resource theory consists in three main ingredients: a set of objects, specifying the physical property that may serve as a resource, and a characterization of the set of free objects, which are the ones that do not contain the resource; a set of free operations, that map every free object into a free object; resource quantifiers that provide a quantitative characterization of the amount of resource a given object contain.

One of the essential requirements for a resource quantifier is that it must be monotonous under free operations, that is, the quantifier must not increase when a free operation is applied. Hence, to define proper quantifiers, one needs first to establish the set of free operations that will be considered, which can vary depending on the applications or the physical constraints under consideration.

In a resource theory of non-locality, the set of objects is the set of non-signaling correlations \(\mathcal{NC}\), and the set of free objects is the set \(\mathcal{C}\) of local correlations. A detailed discussion of several physically relevant free operations for non-locality can be found in Refs. \([2, 20]\).

In what follows we sketch the proof that our measure \(\text{NL}(q)\) is a monotone for a resource theory of non-locality defined by a wide class of free operations. The detailed proof the results below can be found in the Appendix \([A]\).

The first free operation we consider is relabeling \(\mathcal{R}\) of inputs and outputs, defined by a permutation of the set of inputs \(x\) and \(y\), and outputs \(a\) and \(b\). This operation corresponds to a permutation of the entries of the correlation vectors \(q\), and hence we have that

\[
\text{NL}(\mathcal{R}(q)) = \text{NL}(q). \tag{14}
\]

Another natural free operation is to take convex sums between a distribution \(q\) and a local distribution \(p \in \mathcal{C}\). In this case, the triangular inequality for the \(\ell_1\) norm implies that, if \(\pi \in [0, 1]\),

\[
\text{NL}(\pi q + (1 - \pi) p) \leq \pi \text{NL}(q). \tag{15}
\]

Combining monotonicity under relabellings and convexity of the \(\ell_1\) norm one can show that NL is monotonous under convex combinations of relabeling operations.

Next, we sketch the proof of the monotonicity of NL under more sophisticated free operations, namely post-processing and pre-processing operations. Given \(q\), we define a post-processing operation as one that transforms \(q\) into \(\mathcal{O}(q)\), where

\[
\mathcal{O}(q)(a, \beta | x, y) = \sum_{a,b} O^L(a, \beta | a, b, x, y) \times q(a, b | x, y), \tag{16}
\]

and \(O^L\) is a \(x, y\)-dependent correlation with inputs \(a, b\) and outputs \(a, \beta\) that satisfies

\[
O^L(a, \beta | a, b, x, y) = \sum_{\lambda} p(\lambda) O_A^L(a | a, x) \times O_B^L(\beta | b, y). \tag{17}
\]

As shown in Refs. \([3, 33]\), output operations preserve the set of local distributions: if \(q \in \mathcal{C}\), then \(\mathcal{O}(q) \in \mathcal{C}\). As a consequence of convexity of the \(\ell_1\) norm, one can prove that if \(\mathcal{O}\) is an post-processing operation, then

\[
\text{NL}(\mathcal{O}(q)) \leq \text{NL}(q). \tag{18}
\]

Regarding pre-processing operations, it is possible to show that NL is monotonous under uncorrelated input enlarging operations defined in \([3]\). More generally, one
can define a pre-processing operation that transforms \( q \) into \( \mathbf{I}(q) \), where
\[
\mathbf{I}(q)(a,b|x,\psi) = \sum_{x,y} q(a,b|x,y) I^\ell(x,y|x,\psi),
\]
and \( I^\ell \) is a local correlation with inputs \( x,\psi \) and outputs \( x,y \). It was also shown in Refs. [2,1] that pre-processing operations preserve the set of local distributions: if \( q \in \mathcal{C}_C \), then \( \mathbf{I}(q) \in \mathcal{C}_C \). In what follows, we will consider the restricted class of pre-processing operations that satisfy \( |x| = |\chi|, |y| = |\psi| \) and
\[
\sum_{\chi,\psi} I^\ell(x,y|x,\psi) \leq 1.
\]
Intuitively, these restrictions forbid one to increase artificially the number of inputs of the scenario and correlated input enlarging operations, defined in [4]. As a consequence of this restriction and convexity of the \( \ell_1 \) norm, one can prove that if \( \mathbf{I} \) is an output operation with \( |x| = |\chi|, |y| = |\psi| \) satisfying Eq. (20), then
\[
\text{NL}(\mathbf{I}(q)) \leq \text{NL}(q).
\]

V. APPLICATIONS TO VARIOUS BELL SCENARIOS

A. Bipartite

We start considering the paradigmatic CHSH scenario [23], where each of the two parts have two measurement settings with two outcomes each, that is, \( x,y,a,b = 0,1 \). The only Bell inequality (up to symmetries) characterizing this scenario is the CHSH one, that using the notation in [12] can be written as
\[
\text{CHSH} = q_{a=0,b=0}^0 + q_{a=0,b=0}^1 + q_{a=1,b=0}^0 - q_{a=0,b=1}^1 - q_{a=1,b=1}^0 \leq 0,
\]
where in the inequality above we have used the shorthand notation \( q_{a,b}^k = q(a=0,b=0|x,y) \) and similarly to the other terms. Using the dualization procedure described in Sec. [8] and assuming \( q \) to be a non-signaling distribution (respecting [8]), one can prove that
\[
\text{NL}(q) = \frac{1}{2} \max \{0, \Pi(\text{CHSH})\},
\]
where \( \Pi(\text{CHSH}) \) stand for all the 8 symmetries (under permutation of parts, inputs and outputs) of the CHSH inequality. As expected, in the CHSH scenario the CHSH inequality completely characterizes the trace distance of a given test distribution to the set of local correlations [43].

Moving beyond the CHSH scenario, we have also considered the CGLMP scenario [44] where two parts perform two possible measurement with a number \( d \) of outcomes. The CGLMP inequality can be succinctly written for any \( d \) as:
\[
\Pi_{\text{CGLMP}}^d = \frac{1}{4} \sum_{k=0}^{d/2-1} \left( 1 - \frac{2k}{d-1} \right)
\]
and that we conjecture to hold true to any \( d \). Given that the quantum violation of the CGLMP inequality increases with \( d \) [34] it follows that the maximum quantum value of \( \text{NL}(q) \) will follow a similar trend as shown in Fig. [3].

FIG. 3. Value of \( \text{NL}(q) \) as a function of the value of the CGLMP inequality \( \Pi_{\text{CGLMP}}^d \) for \( d = 2,3,4,5 \). The inset shows the maximum known quantum violation for each of these values of \( d \) that in turn implies that \( \text{NL}(q) = 0.1035 \) (d=2), \( \text{NL}(q) = 0.1143 \) (d=3), \( \text{NL}(q) = 0.1215 \) (d=4) and \( \text{NL}(q) = 0.1269 \) (d=5).

Finally, another scenario we have considered is the one introduced in [32], where each of the two parts measure a number \( n \) of observables with 2 outcomes.
each. We analyze the $I_{nn22}$ inequality that has the form

$$I_{nn22} \leq 0,$$

(26)

where, for example, for $n = 3$ we have

$$I_{3322} = q_{AB}^{0,0} + q_{AB}^{1,0} + q_{AB}^{2,0} + q_{AB}^{1,1} - q_{AB}^{0,1} - q_{AB}^{2,1} + q_{AB}^{0,0} - q_{AB}^{1,0} - q_{AB}^{2,0} - q_{AB}^{1,1} - q_{AB}^{0,1} - q_{AB}^{2,1}.$$

(27)

The maximum non-signaling violation of these inequalities grow linearly with the number of settings $n$ as $I_{nn22} = (n - 1)/2$.

Again, by fixing the value of the inequality and imposing the non-signalling constraints we have obtained the corresponding value of $\text{NL}(q)$, with the results shown in Fig. 4. Interestingly, even achieving the maximal non-signaling violation of the $I_{nn22}$ we obtain a value for $\text{NL}(q)$ that decreases with the number of settings $n$. Therefore, when restricted to this class of inequalities, the best situation is already achieved with $n = 2$ (the CHSH scenario). Furthermore this illustrates well the fact that different measures of non-locality do not coincide in general: even though the violation of a Bell inequality can grow with number of settings, the trace distance of the corresponding distribution might decrease.

FIG. 4. Value of $\text{NL}(q)$ as a function of the value of the inequality $I_{nn22}$ for $n = 2, 3, 4, 5, 6$. The inset shows the maximum known quantum violation (achieved with qubits) for $n = 2, 3$. Interestingly, even if the maximum NS violation of the inequality increases with the number of settings $n$, its trace distance to the set of local correlations decreases.

B. Tripartite

In the tripartite scenario, considering that each of the parts measure two dichotomic observables, all the different classes of Bell inequalities have been classified into two groups: the CHSH class and the Sliwa class. The results are shown in Fig. 5.

FIG. 5. a) Values of $\text{NL}(q)$ as a function of the violation of the 46 Sliwa classes of inequalities. b) The values of $\text{NL}(q)$ for the maximal quantum $I_{22}$ and nonsignalling violation of each one of these inequalities. The maximum value of $\text{NL}(q) = 1/8$ is achieved by the maximum violation of the Mermin inequality $I_{22}$ (corresponding to the 2nd Sliwa inequality).

There are 46 of them, under the name of Sliwa inequalities.

Following a similar approach to the CGLMP and $I_{nn22}$ discussed above, we have computed the value of $\text{NL}(q)$ as a function of the various Sliwa inequalities. The result in shown in Fig. 5 together with the values associated to the maximum violation of these inequalities. Regarding quantum violations we have used the probability distributions presented in [17] where the maximum quantum violation of the Sliwa inequalities has been considered. The results are shown in Fig. 5, where we can see that the optimum quantum value of $\text{NL}(q) = 1/8$, higher than the one obtained for the maximum quantum violation of CHSH but smaller than the one for CGLPM already for $d = 5$. The maximum non-signalling violation of these inequalities leads to $\text{NL}(q) = 0.25$, the same value obtained for the CGLMP inequality (any $d$) in the bipartite case.

C. Mermin inequality for more parts

The results presented above show that considering the paradigmatic and Sliwa [16] inequalities, the best values we could achieve for our quantifier were...
\( \text{NL}(q) = 1/4 \) in the case of non-signalling correlations and \( \text{NL}(q) = 1/8 \) for quantum correlations. As we show next, these values can be improved analyzing multipartite scenarios beyond 3 parts, more precisely considering the generalization of the Mermin inequality \([12][43]\) in its form given by \([4]\)

\[ \langle M_N \rangle \leq 1, \quad (28) \]

that is defined recursively starting with \( M_1 = A_1 \) by

\[ M_i = \frac{M_{i-1}}{2}(A_i + \bar{A}_i) + \frac{M_{i-1}}{2}(A_i - \bar{A}_i), \quad (29) \]

where \( M_{i-1} \) is obtained from \( M_i \) by exchanging all the observables \( A \leftrightarrow \bar{A} \). By choosing suitable projective observables in the \( X-Y \) plane of the Bloch sphere and GHZ states \([43]\), the maximum quantum violation is given by \( \langle |M_N| \rangle = 2^{\frac{N + 1}{2}} \). For \( N \) odd this is also the algebraic/non-signalling maximum of the inequality. For \( N \) even the algebraic/non-signalling maximum is given by \( \langle |M_N| \rangle = 2^\frac{N}{2} \). Succinctly, the maximum NS is \( \langle |M_N| \rangle = 2^{\frac{N+1}{2}} \).

Following the same approach as before, we have computed the value of \( \text{NL}(q) \) as a function of \( M_N \). The results are shown in Fig. 6. Interestingly, we see a clearly increase of \( \text{NL}(q) \) as we consider the maximum violation of \( M_N \) with increasing \( N \). Notice, however, that going from \( N \) even to \( N + 1 \) seems to decrease the value of \( \text{NL}(q) \). The reason for that comes from the fact that from the all possible \( 2^N \) measurement settings allowed by the scenario, only \( 2^{N-1} \) enter in the evaluation of the \( M_N \). Since we are fixing the probability of the inputs to be identical distributed, that amounts to reduce \( \text{NL}(q) \) by a factor of half. If instead, we now choose the probability of inputs to be 1/2\(^{N-1}\) for all those appearing in \( M_N \) and zero otherwise, we then recover a monotonically increasing value of \( \text{NL}(q) \) with \( N \). In particular, notice that by doing that we achieve a value of \( \text{NL}(q) = 1/4 \) for the maximum quantum violation of the Mermin inequality in the tripartite scenario.

In this case, we can also provide an analytical construction, providing an upper bound for \( \text{NL}(q) \), that perfectly coincides with the LP results with \( N = 2, \ldots, 6 \) and that for this reason we conjecture provides the actual value of \( \text{NL}(q) \) for any \( N \). For simplicity in what follows we will restrict our attention to the case of \( N \) even.

We choose a fixed distribution \( q = vq_{\text{max}} + (1-v)\frac{1}{2^N} \), where \( q_{\text{max}} \) is the distribution given the maximum NS of the inequality. Notice that \( M_N \) only contain full correlators and can be succinctly written as

\[ M_N = \frac{1}{2^N} \sum_{x_1, \ldots, x_N = 0, 1} c_{x_1, \ldots, x_N} \langle A_1^{x_1} \ldots A_N^{x_N} \rangle \quad (30) \]

\[ \text{NL}(q) = v(1/2N)\alpha_N = M_N(1/2^{3N/2})\alpha_N \quad (31) \]

that tends to \( \text{NL}(q) = 1/2 \) with \( N \to \infty \) and \( v = 1 \) (\( M_N = 2^{N/2} \)).

**VI. RELATION TO OTHER NON-LOCALITY MEASURES**

In the following we will compare the trace distance measure with 3 other standard measures of non-locality: the amount of violation of Bell inequalities, the EPR-2 decomposition \([12]\) and the relative entropy \([12]\).
A. The violation of Bell inequalities

The violation of Bell inequalities are a standard way of quantifying the degree of non-locality and in some cases can as well find operational interpretations, for instance, as the probability of success in some distributed computation protocols \[14,15\]. Clearly, one can expect that the more a given distribution violates a Bell inequality the higher is the trace distance from the local set.

However, in general, the violation of a given Bell inequality only provides a lower bound to such trace distance as can be clearly seen in Fig. 7. In this figure we consider the I_{1322} inequality (see \[28\]). Using the LP formulation we have computed NL(q) for a distribution of the form q = v q_{\max} + (1 – v) 1/4, that is, a mixture of the distribution maximally violating the I_{1322} with white noise rendering I_{1322} = 2v – 1. Clearly, simply imposing the value of the violation of the inequality only gives a non-tight lower bound to NL(q).

Moreover, we see that optimizing over p(x, y) such that \( p(x, y) = 1/4 \) iff \( x, y \neq 2 \) (in such a way that we recover the CHSH inequality) gives us a higher value for NL(q).

![Fig. 7. The black line shows NL(q) for the distribution q = v q_{max} + (1 – v) 1/4 as a function of I_{1322} = 2v – 1 considering that p(x, y) = 1/4 iff x, y \neq 2 (in which case we basically recover I_{2222}). The red solid line shows NL(q) as a function of I_{2222} = 2v – 1 considering the full distribution q. Finally, the purple dashed line shows NL(q) as a function of I_{2222} = 2v – 1 considering only the value of the inequality (thus optimizing over all q compatible with it). Clearly, imposing the value of a given Bell inequality only provides a (in general non-tight) lower bound for NL(q).](image)

B. The EPR-2 decomposition

Any probability distribution \( q = q(a, b|x, y) \) in a Bell experiment can be decomposed into convex mixture of a local part \( q_L \) and a non-local and non-signalling part \( q_{NL} \) as \( q = (1 - w_{NL}) q_L + w_{NL} q_{NL}, \) with \( 0 \leq w_{NL} \leq 1. \) The minimum non-local weight over all such decompositions,

\[
\bar{w}_{NL}(q) = \min_{q_L, w_{NL}} w_{NL}.
\]  

defines the nonlocal content of \( q \), a natural quantifier of the non-locality in \( q \). Nicely, the violation of any Bell inequality \( I \leq I^2 \) (with \( I^2 \) the local bound) yields a non-trivial lower bound to \( \bar{w}_{NL} \) given by

\[
\bar{w}_{NL}(q) \geq \frac{I(q) - I^2}{I^2 - I^2},
\]  

with \( I^L \) being the maximum value of I obtainable with non-signaling correlations. Similarly to the trace distance the non-local content can also be computed via a linear program. However, differently from the trace distance, we see that any extremal non-local point of the NS-polytope will achieve the maximum according to this measure, independently of its actual distance to the set of local correlations.

Considering the CHSH scenario and following an identical approach as the one use to get \[23\] one can show that \( \bar{w}_{NL} = 2 \max \{0, \Pi(CHSH)\} \). That is, in this particular case we have a simple relation \( \bar{w}_{NL}(q) = 4NL(q) \). This picture, however, changes drastically already at the tripartite scenario. For each of 45 classes of extremal NS points (those violating maximally the Sliwa inequalities) it follows from \[13\] that they achieve the maximum \( \bar{w}_{NL}(q) = 1 \). So, from the perspective of the non-local content all non-local extremal points of the NS polytope display the same amount of non-locality. That is not the case for NL(q), as can be clearly seen in Fig. 5 as the different extremal points have different values for it, illustrating the fact that they are at different distances from the local polytope.

C. The relative entropy

Another important non-locality quantifier is defined in terms of the relative entropy (also known as the Kullback-Leibler divergence) given by \[16,17\]

\[
NL_{KL}(q) = \frac{1}{|x||y|} \min_{p \in C} KL(q, p),
\]  

with \( KL(q(x), p(x)) = \sum_x q(x) \log q(x) / p(x) \). Similarly to \[7\] we have also chosen an identically distributed distribution for the outcomes. This quantity is also a monotone for the resource theory of non-locality.
and differently for the results

\[ q \text{ (left) and } NL \text{ (right) inequality as function of } \gamma \text{ for the distribution } q \text{ arising from measurements on the state } |\phi\rangle = \gamma |00\rangle + \sqrt{1 - 2\gamma^2} |11\rangle + \gamma |22\rangle. \]

From the critical value of \( \gamma_c < 0.369 \) we have that \( \gammaNL_{CGLMP} > 0 \) and \( NL(q) > 0 \). In both figures, the circle, square and diamond corresponds to the maximal entangled state (ME), the state that maximal violate (MV) the \( \gammaNL_{CGLMP} \) and the optimal state (OS) that provides the maximal KL(\( q, p \)) respectively. As we can see, the maximal \( \gammaNL_{CGLMP} \) violation is achieved by \( \gamma = 0.617 \) and the maximal of \( NL(q) \) is achieved for the same \( \gamma \). This is opposed to the result found in [17], where the quantum distribution with maximum \( KL(q, p) \) do not violated \( \gammaNL_{CGLMP} \) maximally.

In [17] special attention has been given to the relation between \( NL_{KL}(q) \) and the violation of the CGLMP inequality. It has been shown that quantum correlations maximally violating the CGLMP inequality do not necessarily imply the optimum relative entropy. For that, the set of considered quantum distributions comes from a fixed set of measurements on a two-qutrit state of the form \( |\phi\rangle = \gamma |00\rangle + \sqrt{1 - 2\gamma^2} |11\rangle + \gamma |22\rangle \). Inspired by that result we have analyzed what is the value of \( NL(q) \) in the same setup. The results are shown in Fig. 8 with 3 distributions being specially relevant [17]: i) the distribution obtained from the maximum violation of CGLPM with maximally entangled states (\( \gamma = 1\sqrt{3} \)), ii) the maximum violation of the CGLMP (obtained with \( \gamma = 0.617 \)) and iii) the maximum value of \( NL_{KL}(q) \) (achieved with \( \gamma = 0.642 \)). As expected from the results in Sec. V] and differently for the results obtained for the relative entropy, the more we violate the CGLMP inequality the higher is the trace distance measure \( NL(q) \).

Interestingly, via the Pinsker inequality [18] we can relate the trace and relative entropy measure. That is, the trace distance provides a non-trivial bound to the relative entropy. Furthermore, the LP solutions to the minimization of the trace distance naturally give an ansatz solution providing an upper bound for \( NL_{KL}(q) \). These results are shown in Fig. 9 where we plot \( NL_{KL}(q) \) as a function of the CGLMP violation. In this plot we also show the curve obtained using the optimal non-signalling distribution \( q^* \) obtained from the LP used to minimize \( NL(q) \) subjected to a specific value of the CGLMP inequality (see the caption in Fig. 9 for more details).

![FIG. 8. In this figure we show the behaviour of the \( \gammaNL_{CGLMP} \) (left) and \( NL(q) \) (right) inequality as function of \( \gamma \) for the distribution \( q \) arising from measurements on the state \( |\phi\rangle = \gamma |00\rangle + \sqrt{1 - 2\gamma^2} |11\rangle + \gamma |22\rangle \). From the critical value of \( \gamma_c > 0.369 \) we have that \( \gammaNL_{CGLMP} > 0 \) and \( NL(q) > 0 \). In both figures, the circle, square and diamond corresponds to the maximal entangled state (ME), the state that maximal violate (MV) the \( \gammaNL_{CGLMP} \) and the optimal state (OS) that provides the maximal KL(\( q, p \)), respectively. As we can see, the maximal \( \gammaNL_{CGLMP} \) violation is achieved by \( \gamma = 0.617 \) and the maximal of \( NL(q) \) is achieved for the same \( \gamma \). This is opposed to the result found in [17], where the quantum distribution with maximum KL(\( q, p \)) do not violated \( \gammaNL_{CGLMP} \) maximally.

![FIG. 9. Plot of \( \min_{p\in NS} KL(q, p) \) as a function of the CGLMP inequality for \( d = 3 \). The black dashed curve represents the lower bound provided the Pinsker inequality, defined by

\[ KL(q, p) \geq 4CLNL(q)^2, \]

where \( C = \frac{1}{2} \log_2 e \) and \( NL(q) = \frac{1}{2} \| q - p \|_{\text{F}} \). The red curve provides an upper bound for \( \min_{p\in NS} KL(q, p) \) since it computes \( KL(q^*, p^*) \), where \( q^* \) and \( p^* \) are the solutions provided by the LP minimization of \( NL(q) \) subjected to NS constraints and a given value of the CGLMP inequality. The blue curve is obtained by a brute force minimization of \( \min_{p\in NS} KL(q, p) \). The color points represent the 3 points found in [17] and discussed in the main text. Interestingly, for the green (circle) and pink (diamond) points there are NS correlations (possibly post-quantum) with the same value of the CGLMP inequality but providing a higher value of \( \min_{p\in NS} KL(q, p) \).

VII. DISCUSSION

Apart from its primal importance in the foundations of quantum physics, non-locality has also found several applications as a resource in quantum cryptography [47], randomness generation/amplification [48–49], self-testing [50] and distributed computing [51]. Within these both fundamental and applied contexts, quantifying non-locality is undoubtedly an important primitive.

Here we have introduced a natural quantifier for non-locality, a geometrical measure based on the trace dis-
tance between the probability distribution under test and the set of classical distributions compatible with LHV models. Nicely, our quantifier can be efficiently computed (as well as closed form analytical expressions can be derived) via a linear programming formulation. We have shown that it provides a proper quantifier since it is monotonous under a wide class of free operations defined in the resource theory of non-locality [32]. Finally we have applied our framework to a few scenarios of interest as compared our approach to other standard measures.

It would be interesting to use the trace distance measure to analyze previous results in the literature. For instance, the non-locality distillation and activation protocols proposed in [53] or the role of the trace distance in cryptographic protocols [47]. From the statistical perspective one can investigate the relation of the trace distance to p-values [25], and their interconnections to non-locality [50–52] and the statistical significance of Bell tests with limited data [53,54]. Finally, we point out that even though here we have focused on Bell non-locality, the same measure can also be applied to quantify non-classical behavior in more general notions of non-locality [55–57] as well as in quantum contextuality [8,9]. We hope our results might motivate further research in these directions.

ACKNOWLEDGMENTS

We thank Daniel Cavalcanti for valuable comments on the manuscript and Thiago O. Maciel for interesting discussions. SGAB, BA and RC acknowledges financial support from the Brazilian ministries MEC and MCTIC. BA also acknowledges financial support from CNPq.

Appendix A: Detailed proof of the results in Sec. IV

1. Relabeling operations

**Theorem 1.** If \( \mathcal{R} \) is a relabeling of inputs or outputs of \( \mathcal{q} \), then \( \text{NL}(\mathcal{R}(\mathcal{q})) = \text{NL}(\mathcal{q}) \).

**Proof.** Let \( \mathcal{p}^* \in C_C \) be such that

\[
\text{NL}(\mathcal{p}) = \frac{1}{2|x| |y|} \min_{\mathcal{p} \in C_C} \| \mathcal{R}(\mathcal{q}) - \mathcal{p} \|_{\ell_1}.
\]

The distributions \( \mathcal{R}(\mathcal{q}) \) and \( \mathcal{R}(\mathcal{p}^*) \) are obtained respectively from \( \mathcal{q} \) and \( \mathcal{p}^* \) by a permutation of entries. Then, if follows that

\[
\| \mathcal{R}(\mathcal{q}) - \mathcal{R}(\mathcal{p}^*) \|_{\ell_1} = \| \mathcal{q} - \mathcal{p}^* \|_{\ell_1}.
\]

Since relabeling operations preserve the set of local distributions, \( \mathcal{R}(\mathcal{p}^*) \in C_C \) and hence

\[
\text{NL}(\mathcal{R}(\mathcal{q})) = \frac{1}{2|x| |y|} \min_{\mathcal{p} \in C_C} \| \mathcal{R}(\mathcal{q}) - \mathcal{p} \|_{\ell_1} \quad (A3)
\]

\[
\leq \frac{1}{2|x| |y|} \| \mathcal{R}(\mathcal{q}) - \mathcal{R}(\mathcal{p}^*) \|_{\ell_1} \quad (A4)
\]

\[
= \frac{1}{2|x| |y|} \| \mathcal{q} - \mathcal{p}^* \|_{\ell_1} \quad (A5)
\]

\[
= \text{NL}(\mathcal{q}) \quad (A6)
\]

Relabeling operations are invertible, and hence there is a relabeling operation \( \mathcal{R}^{-1} \) such that \( q = \mathcal{R}^{-1}(\mathcal{R}(q)) \). A similar argument shows that

\[
\text{NL}(\mathcal{q}) = \text{NL}(\mathcal{R}^{-1}(\mathcal{R}(\mathcal{q}))) \geq \text{NL}(\mathcal{R}(\mathcal{q})) \quad (A7)
\]

which proves the desired result.

2. Convex combinations

**Theorem 2.** If \( \mathcal{q} = \sum_k \tau_k \mathcal{q}_k \), where \( \tau_k \leq 0 \) and \( \sum_k \tau_k = 1 \), then

\[
\text{NL}(\mathcal{q}) \leq \sum_k \tau_k \text{NL}(\mathcal{q}_k) . \quad (A8)
\]

**Proof.** Let \( \mathcal{p}_k^* \in C_C \) be such that

\[
\text{NL}(\mathcal{q}_k) = \frac{1}{2|x| |y|} \| \mathcal{q}_k - \mathcal{p}_k^* \|_{\ell_1} \quad (A9)
\]

and let \( \mathcal{p}^* = \sum_k \tau_k \mathcal{p}_k^* \in C_C \). Then, we have

\[
\text{NL}(\mathcal{q}) = \frac{1}{2|x| |y|} \min_{\mathcal{p} \in C_C} \| \mathcal{q} - \mathcal{p} \|_{\ell_1} \quad (A10)
\]

\[
\leq \frac{1}{2|x| |y|} \| \mathcal{q} - \mathcal{p}^* \|_{\ell_1} \quad (A11)
\]

\[
= \frac{1}{2|x| |y|} \left\| \sum_k \tau_k \mathcal{q}_k - \sum_k \tau_k \mathcal{p}_k^* \right\|_{\ell_1} \quad (A12)
\]

\[
\leq \frac{1}{2|x| |y|} \sum_k \tau_k \| \mathcal{q}_k - \mathcal{p}_k^* \|_{\ell_1} \quad (A13)
\]

\[
= \sum_k \tau_k \text{NL}(\mathcal{q}_k) \quad (A14)
\]

3. Post-processing operations

**Theorem 4.** If \( \mathcal{O} \) is an post-processing operation, defined as in Eq. (10), then \( \text{NL}(\mathcal{O}(\mathcal{q})) \leq \text{NL}(\mathcal{q}) \).

**Proof.** Let \( \mathcal{p}^* \in C_C \) be the distribution satisfying Equation [A1]. Then,
\[
\text{NL}(O(q)) = \frac{1}{2|x||y|} \min_{p \in C} \|O(q) - p\|_{\ell_1}
\]  
(A16)

\[
\leq \frac{1}{2|x||y|} \|O(q) - O(p^*)\|_{\ell_1}
\]  
(A17)

\[
= \frac{1}{2|x||y|} \sum_{a,b} \left| \sum_{a,b} O^L(\alpha, \beta | a, b, x, y) (q(a,b|x,y) - p^*(a,b|x,y)) \right|
\]  
(A18)

\[
\leq \frac{1}{2|x||y|} \sum_{a,b} O^L(\alpha, \beta | a, b, x, y) |q(a,b|x,y) - p^*(a,b|x,y)|
\]  
(A19)

\[
= \frac{1}{2|x||y|} \sum_{a,b, x, y} |q(a,b|x,y) - p^*(a,b|x,y)|
\]  
(A20)

\[
= \frac{1}{2|x||y|} \|q - p^*\|_{\ell_1} = \text{NL}(q).
\]  
(A21)

4. Pre-processing operations

In Ref. [3] the author defines the uncorrelated input enlarging operation, which consists in one or more parts adding an uncorrelated measurement locally. Without loss of generality we can assume that this measurement is deterministic, since convexity of the \(\ell_1\) norm implies that NL is also a monotone under the addition of a non-deterministic uncorrelated measurement.

Given a correlation \(q\), suppose part \(A\) adds one uncorrelated measurement at her side. Denoting \(|x| = \ldots\)

\[
\sum_{a,b} |q_f(a,b|m_A + 1, y) - p_f^*(a,b|m_A + 1, y)| =
\]  
(A24)

\[
\sum_b |q(b|y) - p^*(b|y)|
\]  
(A25)

\[
= \sum_b \left| \sum_a q(a,b|x,y) - p^*(a,b|x,y) \right| \forall x \leq m_A
\]  
(A26)

\[
\leq \sum_{a,b} q(a,b|x,y) - p^*(a,b|x,y) \forall x \leq m_A
\]  
(A27)

Hence,

\[
\sum_{a,b} |q_f(a,b|m_A + 1, y) - p_f^*(a,b|m_A + 1, y)| \leq \min_{x \leq m_A} \sum_{a,b} \left| \sum_a q(a,b|x,y) - p^*(a,b|x,y) \right|.
\]  
(A28)

This implies that \(q_f\) is obtained from \(q\) by adding pairs of inputs for which the distance of the probability distributions \(q_f\) and \(p_f^*\) decreases. Since NL is defined by taking the average over the pairs of inputs, this implies the desired result.
We now proceed to prove monotonicity under the pre-processing operations defined in Eq. (10).

**Theorem 6.** If $I$ is an pre-processing operation such that $|x| = |y|$, $|ψ⟩ = |y⟩$ and $∑_x χ^I (x, y | x, ψ) ≤ 1$, then

$$\text{NL}(I(q)) ≤ \text{NL}(q). \quad \text{(A30)}$$

**Proof.** Let $p^* \in C_C$ be the distribution satisfying Equation $\text{NL}(q)$. Then,

$$\text{NL}(I(q)) = \frac{1}{2|x|} \min_{p \in L} \|I(q) - p\|_{I_1} \leq \frac{1}{2|x|} \|I(q) - I(p^*)\|_{I_1} \leq \frac{1}{2|x|} \sum_{a,b,x,ψ} \left| ∑_x (q(a,b|x,y) - p^*(a,b|x,y)) I^L(x,y|x,ψ) \right| \leq \frac{1}{2|x|} \sum_{a,b,x,ψ} \sum_{x,y} I^L(x,y|x,ψ) \|q(a,b|x,y) - p^*(a,b|x,y)| \leq \frac{1}{2|x|} \sum_{a,b,x,ψ} \sum_{x,y} |q(a,b|x,y) - p^*(a,b|x,y)| \leq \frac{1}{2|x|} \|q - p^*\|_{I_1} = \text{NL}(q). \quad \text{(A36)}$$

\[\square\]

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, “Quantum entanglement,” Rev. Mod. Phys. 81, 865–942 (2009).
[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, “Bell nonlocality,” Rev. Mod. Phys. 86, 419–478 (2014).
[3] G. Gour and R. W. Spekkens, “The resource theory of quantum reference frames: manipulations, and nonmonotones,” New J. Phys. 10, 033023 (2008).
[4] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, “Resource theory of quantum states out of thermal equilibrium,” Phys. Rev. Lett. 111, 250404 (2013).
[5] J. I. de Vicente, “On nonlocality as a resource theory and nonlocality measures,” J. Phys. A: Math. Theor. 47, 424017 (2014).
[6] A. Winter and D. Yang, “Operational resource theory of coherence,” Phys. Rev. Lett. 116, 120404 (2016).
[7] B. Coecke, T. Fritz, and R. W. Spekkens, “A mathematical theory of resources,” Information and Computation 250, 59 – 86 (2016). quantum Physics and Logic.
[8] S. Abramsky, R. S. Barbosa, and S. Mansfield, “Contextual fraction as a measure of contextuality,” Phys. Rev. Lett. 119, 050504 (2017).
[9] Barbara Amaral, Adán Cabello, Marcelo Terra Cunha, and Leandro Aolita, “Noncontextual wirings,” arXiv preprint arXiv:1705.07911 (2017).
[10] J. S. Bell, “On the Einstein–Podolsky–Rosen paradox,” Physics 1, 195 (1964).
[11] R. F. Werner, “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model,” Phys. Rev. A 40, 4277–4281 (1989).
[12] S. Popescu and D. Rohrlich, “Quantum nonlocality as an axiom,” Foundations of Physics 24, 379–385 (1994).
[13] P. H. Eberhard, “Background level and counter efficiencies required for a loophole-free Einstein-Podolsky-Rosen experiment,” Phys. Rev. A 47, R747–R750 (1993).
[14] B. F. Toner and D. Bacon, “Communication cost of simulating bell correlations,” Phys. Rev. Lett. 91, 187904 (2003).
[15] S. Pironio, “Violation of Bell inequalities as lower bounds on the communication cost of nonlocal correlations,” Phys. Rev. A 68, 062102 (2003).
[16] W. Van Dam, R. D. Gill, and P. D. Grunwald, “The statistical strength of nonlocality proofs,” IEEE Transactions on Information Theory 51, 2812–2837 (2005).
[17] A. Acín, R. Gill, and N. Gisin, “Optimal Bell tests do not require maximally entangled states,” Phys. Rev. Lett. 95, 230402 (2005).
[18] M. Junge, C. Palazuelos, D. Pérez-García, I. Villanueva, and M. M. Wolf, “Operator space the-
ory: A natural framework for Bell inequalities,” Phys. Rev. Lett. 104, 170405 (2010)
[19] M. J. W. Hall, “Relaxed Bell inequalities and kochen-specker theorems,” Phys. Rev. A 84, 022102 (2011)
[20] R. Chaves, D. Cavalcanti, L. Aolita, and A. Acín, “Multiparticle quantum nonlocality under local decoherence,” Phys. Rev. A 86, 012108 (2012)
[21] E. A. Forsha and Fernando Parisio, “Measure of nonlocality which is maximal for maximally entangled qutrits,” Phys. Rev. A 92, 050101 (2015)
[22] R. Chaves, R. Kueng, J. B. Brask, and D. Gross, “Unifying framework for relaxations of the causal assumptions in Bell’s theorem,” Phys. Rev. Lett. 114, 140403 (2015)
[23] M. Ringbauer, C. Giarmatzi, R. Chaves, F. Costa, A. G. White, and A. Fedrizzi, “Experimental test of nonlocal causality,” Science Advances 2 (2016), 10.1126/sciadv.1600162
http://advances.sciencemag.org/content/2/8/e1600162.full.pdf
[24] A. Montina and S. Wolf, “Information-based measure of nonlocality,” New J. Phys. 18, 013035 (2016)
[25] J. B. Brask and R. Chaves, “Bell scenarios with communication,” J. Phys. A: Math. Theor. 50, 094001 (2017)
[26] R. Gallego and L. Aolita, “Nonlocality free wirings and the distinguishability between bell boxes,” Physical Review A 95, 032118 (2017).
[27] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt, “Proposed experiment to test local hidden-variable theories,” Phys. Rev. Lett. 23, 880–884 (1969)
[28] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, 2010).
[29] C Bernhard, B Bessire, A Montina, M Pfaffhauser, A Stefanov, and S Wolf, “Non-locality of experimental qutrit pairs,” J. Phys. A: Math. Theor. 47, 424013 (2014).
[30] M. Navascues, S. Pironio, and A. Antonio, “Bounding the set of quantum correlations,” Phys. Rev. Lett. 98, 010401 (2007)
[31] B. Lang, T. Vértesi, and M. Navascués, “Closed sets of correlations: answers from the zoo,” J. Phys. A: Math. Theor. 47, 424029 (2014).
[32] D. Collins and N. Gisin, “A relevant two qubit bell inequality inequivalent to the CHSH inequality,” J. Phys. A: Math. Theor. 37, 1775 (2004)
[33] Ll. Masanes, A. Acín, and N. Gisin, “General properties of nonsignaling theories,” Phys. Rev. A 73, 012112 (2006)
[34] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, “Bell-type inequalities to detect true n-body nonseparability,” Phys. Rev. Lett. 88, 170405 (2002)
[35] We have normalized the inequality in order to obtain the local bound equal 0.
[36] Cezary Śliwa, “Symmetries of the Bell correlation inequalities,” Physics Letters A 317, 165 – 168 (2003)
[37] S. López-Rosa, Z. Xu, and A. Cabello, “Maximum nonlocality in the (3,2,2) scenario,” Phys. Rev. A 94, 062121 (2016)
[38] N. David Mermin, “Extreme quantum entanglement in a superposition of macroscopically distinct states,” Phys. Rev. Lett. 65, 1838–1840 (1990)
[39] M. Ardehali, “Bell inequalities with a magnitude of violation that grows exponentially with the number of particles,” Phys. Rev. A 46, 3775–3778 (1992)
[40] A. V. Belinskii and D. N. Klyshko, “Interference of light and Bell’s theorem,” Physics-Uspekhi 36, 653–693 (1993)
[41] A. V. Belinskii, “Bell’s theorem for trichotomic observables,” Physics-Uspekhi 40, 305 (1997).
[42] N Gisin and H Bechmann-Pasquinucci, “Bell inequality, Bell states and maximally entangled states for n qubits,” Physics Letters A 246, 1 – 6 (1998)
[43] D. M. Greenberger, M. A. Horne, and A. Zeilinger, “Going beyond Bell’s theorem,” in Bell’s theorem, quantum theory and conceptions of the universe (Springer, 1989) pp. 60–72.
[44] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, “Nonlocality and communication complexity,” Rev. Mod. Phys. 82, 665–698 (2010)
[45] C. Brukner, M. Zukowski, J. Pan, and A. Zeilinger, “Bell’s inequalities and quantum communication complexity,” Phys. Rev. Lett. 92, 127901 (2004)
[46] A. A. Fedotov, Pr Harremoës, and F. Topsoe, “Refinements of pinsker’s inequality,” IEEE Transactions on Information Theory 49, 1491–1498 (2003)
[47] A. Acin, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, “Device-independent security of quantum cryptography against collective attacks,” Phys. Rev. Lett. 98, 230501 (2007)
[48] S. Pironio, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, “Random numbers certified by Bell’s theorem,” Nature 464, 1021–1024 (2010).
[49] R. Colbeck and R. Renner, “Free randomness can be amplified,” Nature Physics 8, 350 (2012).
[50] D. Mayers and A. Yao, “Self testing quantum apparatus,” Quantum Information & Computation 4, 273–286 (2004)
[51] N. Brunner, D. Cavalcanti, A. Salles, and P. Skrzypczyk, “Bound nonlocality and activation,” Phys. Rev. Lett. 106, 020402 (2011).
[52] D. Elkouss and S. Wehner, “(nearly) optimal P-values for all Bell inequalities,” arXiv preprint arXiv:1510.07233 (2015)
[53] Y. Zhang, S. Glancy, and E. Knill, “Asymptotically optimal data analysis for rejecting local realism,” Phys. Rev. A 84, 062118 (2011)
[54] P. Bierhorst, “A robust mathematical model for a loophole-free Clauser–Horne experiment,” J. Phys. A: Math. Theor. 48, 195302 (2015)
[55] G. Svetlichny, “Distinguishing three-body from two-body nonseparability by a bell-type inequality,” Phys. Rev. D 35, 3066–3069 (1987)
[56] R. Gallego, L. E. Wurflinger, A. Acín, and M. Navascués, “Operational framework for nonlocality,” Phys. Rev. Lett. 109, 070401 (2012)
[57] R. Chaves, D. Cavalcanti, and L. Aolita, “Causal hierarchy of multipartite Bell nonlocality,” Quantum 11, 23 (2017)