White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/161975/

Version: Published Version

Article:
Vilasini, V. and Colbeck, Roger (2020) Limitations of entropic inequalities for detecting nonclassicality in the postselected Bell causal structure. Physical Review Research. 033096. ISSN: 2643-1564

https://doi.org/10.1103/PhysRevResearch.2.033096

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Limitations of entropic inequalities for detecting nonclassicality in the postselected Bell causal structure

V. Vilasini* and Roger Colbeck†
Department of Mathematics, University of York, Heslington, York YO10 5DD, United Kingdom

(Received 6 December 2019; accepted 16 June 2020; published 17 July 2020)

Classical and quantum physics impose different constraints on the joint probability distributions of observed variables in a causal structure. These differences mean that certain correlations can be certified as nonclassical, which has both foundational and practical importance. Rather than working with the probability distribution itself, it can instead be convenient to work with the entropies of the observed variables. In the Bell causal structure with two inputs and outputs per party, a technique that uses entropic inequalities is known that can always identify nonclassical correlations. Here we consider the analog of this technique in the generalization of this scenario to more outcomes. We identify a family of nonclassical correlations in the Bell scenario with two inputs and three outputs per party whose nonclassicality cannot be detected through the direct analog of the previous technique. We also show that use of Tsallis entropy instead of Shannon entropy does not help in this case. Furthermore, we give evidence that natural extensions of the technique also do not help. More precisely, our evidence suggests that even if we allow the observed correlations to be postprocessed according to a natural class of nonclassicality nongenerating operations, entropic inequalities for either the Shannon or Tsallis entropies cannot detect the nonclassicality, and hence that entropic inequalities are generally not sufficient to detect nonclassicality in the Bell causal structure.

In addition, for the bipartite Bell scenario with two inputs and three outputs we find the vertex description of the polytope of nonsignalling distributions that satisfy all of the CHSH-type inequalities, which is one of the main regions of investigation in this work.

DOI: 10.1103/PhysRevResearch.2.033096

I. INTRODUCTION

Causal structures are a useful tool for understanding correlations between observed events. Such correlations may be mediated by an influence traveling from one to the other, or come about due to common causes, which may not be observed. The nature of any unobserved causes depends on the theory being considered. For instance, they may be classical, quantum or from a generalized probabilistic theory (GPT) [1], and the kinds of observed correlations that are possible in general depends on this. At the foundational level, studying the differences gives us insight into how the notion of causality differs between theories, while, on a practical level, these differences are crucial for applications in device-independent cryptography [2–7].

One way to establish a difference is to violate a *Bell inequality* [8], where we use the term to mean a necessary condition on the observed correlations when any unobserved systems are classical. Bell inequalities are often introduced using the (bipartite) Bell causal structure [see Fig. 1(a)]. Here there are four observed variables: A and B corresponding to the inputs of each party, and X and Y corresponding to the outputs. In the case that the numbers of possible inputs are i_A and i_B and likewise the number of possible outputs are o_A and o_B, we call the scenario the (i_A, i_B, o_A, o_B) Bell scenario. For the (2,2,2,2) case, the CHSH inequalities [9] are known to be the only class of Bell inequalities required to completely characterize the scenario (i.e., all extremal 2-setting, 2-outcome Bell inequalities are equivalent to the CHSH inequalities up to symmetry). In the (2,2,3,3) scenario, which will be the main focus of this paper, there is only one new class of Bell inequalities inequivalent to CHSH, the I_{2233} class [10,11]. In other words, given a no-signalling distribution for the (2,2,3,3) scenario, the distribution is local if and only if all the CHSH and I_{2233} inequalities hold.

As i_A, i_B, o_A, and o_B increase, many new classes of extremal Bell inequalities are found and these scenarios quickly become difficult to fully characterize [12–15]. One attempt at avoiding this difficulty is to move away from probability space to instead consider inequalities expressed in terms of the entropies of the variables involved. There are two ways that these can be used: either directly using the causal structure under consideration, or by using the postselection technique in which the original causal structure is first modified (more details can be found later in this paper). Braunstein and Caves [16] were the first to derive an entropic Bell inequality. They considered the postselected version of the Bell causal structure shown in Fig. 1(b) and found entropic inequalities that hold for all classical distributions. These can be violated when one or more of the unobserved nodes are quantum, and hence behave like entropic versions of Bell inequalities.

*vv577@york.ac.uk
†roger.colbeck@york.ac.uk

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

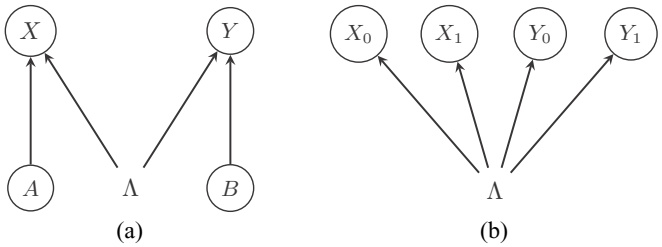

FIG. 1. (a) The bipartite Bell causal structure. The nodes A and B represent the random variables corresponding to independently chosen inputs, while X and Y represent the random variables corresponding to the outputs. Λ is an unobserved node representing the common cause of X and Y. (b) The postselected Bell causal structure for two parties. The observed nodes X_a represent the outputs when the input is $a \in \{0, 1\}$ and likewise for Y_b. Note that X_0 and X_1 are never simultaneously observed and likewise Y_0 and Y_1.

Although entropic inequalities can be useful as a way to detect nonclassicality, the entropic approach has several disadvantages. For instance, in the bipartite Bell causal structure without postselection the set of achievable Shannon entropies over the observed variables for classical and quantum causes coincide [17], so without postselecting, nonclassicality cannot be detected through entropic Bell inequalities in this case. The use of other entropic measures such as Tsallis entropies to analyze this problem in the absence of postselection has also been shown to have limitations [18] and no quantum violations are known for the entropic Bell inequalities derived in Ref. [18]. Because of this, we focus on postselected causal structures in this paper.

It is natural to ask whether the nonclassicality of a distribution can always be detected through postselected entropic inequalities. For the $(d, d, 2, 2)$ Bell scenarios with $d \geqslant 2$, this is known to be the case [19] in the following sense. For every nonclassical distribution in the $(d, d, 2, 2)$ Bell scenario, there is a transformation that does not make any classical distribution nonclassical, and such that the resulting distribution violates one of the BC entropic inequalities. The main purpose of this work is to investigate whether a similar result holds for nonbinary outcomes. To do so, we need to specify a class of postprocessing operations. The most general operations that we could consider are the *nonclassicality nongenerating (NCNG)* operations, i.e., those that do not map any classical distribution to a nonclassical one. An interesting subset of these is the class of postprocessings achievable through *local operations and shared randomness (LOSR)*, which are physical in the sense that two separated parties with shared randomness could perform them.[1] Because of the difficulty of dealing with arbitrary NCNG operations, for the majority of our analysis we consider LOSR supplemented with the additional (NCNG) operation where the parties are exchanged (and convex combinations). We use LOSR + E to refer to this supplemented set.

We study the (2,2,3,3) Bell scenario with LOSR + E postprocessing operations, to see whether when applied to any nonclassical distribution the result violates an entropic Bell inequality. We investigate this using both Shannon and Tsallis entropies. Our motivation for considering Tsallis entropies is that they are known to provide an advantage over the Shannon entropy in detecting nonclassicality in the absence of postprocessing [20] in the sense that there are nonclassical distributions that violate Tsallis entropic inequalities but not the analogous Shannon-entropic inequality. In the (2,2,2,2) case, however, due to the result of Ref. [19], this advantage is less apparent when postprocessings are considered. It is unclear whether or not this is also the case for the (2,2,3,3) scenario, and hence we consider Tsallis entropies in this work.

The result of Ref. [19] that Shannon inequalities can always be used in the case that CHSH is violated readily extends to the (2,2,3,3) scenario (cf. corollary 3). Thus of most interest to us is the region containing nonclassical distributions that satisfy all the CHSH-type inequalities. After finding the vertex description of this region, we show that some distributions in this region violate entropic inequalities and conjecture that there are others that cannot violate either Shannon or Tsallis entropic inequalities after processing with LOSR + E. For our conjecture, we consider a class of isotropic nonclassical distributions in the (2,2,3,3) scenario and give numerical evidence that arbitrary Shannon entropic inequalities and a class of Tsallis entropic inequalities cannot detect the nonclassicality of these distributions under any processing with LOSR + E. If correct, our conjecture implies that the method of Ref. [19] for the (2,2,2,2) scenario does not work for all nonclassical distributions in the (2,2,3,3) scenario. That said, it remains a useful technique in many cases.

The structure of the remainder of the paper is as follows. After introducing our notation and reviewing some existing work in Sec. II, we proceed to investigate the (2,2,3,3) scenario. We present new results for this scenario in probability space (Sec. III) as well as in entropy space (Sec. IV). Finally, in Sec. V, we conclude and discuss some open questions. These results, along with those of Refs. [17,18], highlight some of the limitations of the entropic approach to analyzing causal structures.

II. PRELIMINARIES

A. Probability distributions and entropy

We begin with some notation. Given a conditional probability distribution $p_{XY|AB}$, where A, B, X, and Y have cardinalities i_A, i_B, o_A, and o_B, respectively, we can express the distribution using a matrix. For instance, in the case where all the variables take values in $\{0, 1\}$ and using $p(xy|ab)$ as an abbreviation for $p_{XY|AB}(xy|ab)$, this is done as

$$p_{XY|AB} = \left[\begin{array}{cc|cc} p(00|00) & p(01|00) & p(00|01) & p(01|01) \\ p(10|00) & p(11|00) & p(10|01) & p(11|01) \\ \hline p(00|10) & p(01|10) & p(00|11) & p(01|11) \\ p(10|10) & p(11|10) & p(10|11) & p(11|11) \end{array}\right] \quad (1)$$

and the generalization to larger alphabets is analogous (see, e.g., Ref. [21]). This format is convenient because it makes it easy to check whether a distribution is no-signalling, i.e., to check that $p_{X|AB}$ is independent of B and that $p_{Y|AB}$ is independent of A.

[1] In general, NCNG operations used on the correlations prior to evaluating an entropic inequality need not be physical in this sense.

In a given scenario, we will be interested in the set of distributions that are nonsignalling, and the set of classical (local) distributions. These sets form convex polytopes that are highly symmetric. In particular, such polytopes are invariant under local relabellings and/or relabelling parties. By *local relabellings* we mean combinations of relabelling the inputs (e.g., $A \mapsto A \oplus 1$) and outputs conditioned on the local input (e.g., $X \mapsto X \oplus \alpha A \oplus \beta$ where $\alpha, \beta \in \{0, 1\}$ and $\oplus$ denotes modulo-2 addition). One might also think about more general *global relabellings* that depend on both inputs (for instance maps of the form $X \mapsto X \oplus \alpha A \oplus \beta B \oplus \gamma$ with $\alpha, \beta, \gamma \in \{0, 1\}$), but these do not preserve the no-signalling set in general so will not be considered here. The only global relabelling we consider is exchange of the two parties, which corresponds to transposing the distribution in the matrix notation of Eq. (1) and always preserves no-signalling.

Given a random variable X, distributed according to the discrete probability distribution $p_x := p_X(x)$ the Shannon entropy is defined by[2] $H(X) = -\sum_{\{x:p_x>0\}} p_x \ln p_x$. Given two random variables, X and Y, the conditional Shannon entropy is defined by $H(X|Y) = -\sum_{\{xy:p_{xy}>0\}} p_{xy} \ln p_{x|y}$. The following properties hold: (P1) monotonicity: $H(X) \leqslant H(XY)$; (P2) strong-subadditivity: $H(XY) + H(YZ) \geqslant H(XYZ) + H(Y)$; and (P3) chain rule: $H(X|Y) = H(XY) - H(Y)$.

The order q Tsallis entropy of X for a real parameter q is defined as [23]

$$S_q(X) = \begin{cases} -\sum_{\{x:p_x>0\}} p_x^q \ln_q p_x & q \neq 1 \\ H(X) & q = 1 \end{cases}. \tag{2}$$

In this expression, we have used the q-logarithm function $\ln_q p_x = \frac{p_x^{1-q}-1}{1-q}$, which converges to the natural logarithm function as $q \to 1$. This means that $\lim_{q\to 1} S_q(X) = H(X)$ and hence that $S_q(X)$ is continuous in q. We henceforth write $\sum_x$ instead of $\sum_{\{x:p_x>0\}}$ with the implicit understanding that probability zero events are excluded from the sum.

The Tsallis entropies for $q \geqslant 1$ satisfy many of the same properties as the Shannon entropy. In particular, monotonicity, strong subadditivity and chain rule all hold for Tsallis entropies for all $q \geqslant 1$ [24,25], making these polymatroids like the Shannon entropy.

B. Causal structures

The relationships between different variables of interest can be conveniently expressed as a causal structure. This is a *directed acyclic graph* (DAG) where the observed variables are nodes, and there may be additional nodes representing unobserved systems. Given such a causal structure, we distinguish the cases where the hidden systems are classical, quantum or are from some generalized probabilistic theory. For every classical causal structure that has at least one parentless observed node, a postselected causal structure can be defined. The general technique for doing this can be found in Ref. [26] (for example).

[2]Note that we choose to use the natural logarithm in this work, in contrast to the usual choice of using base 2, in order to make the relation to the Tsallis entropy more straightforward.

In this work we will only consider the Bell causal structure with two inputs per party and the postselected version thereof (see Fig. 1). The postselected causal structure is obtained by removing the parentless observed nodes A and B in the original causal structure 1(a) and replacing the descendants X and Y with two copies of each i.e., $X_{A=0}$, $X_{A=1}$, $Y_{B=0}$, $Y_{B=0}$ such that the original causal relations are preserved and there is no mixing between the copies [this is shown in Fig. 1(b)]. It makes sense to do this in the classical case because classical information can be copied, so we can simultaneously consider the outcome X given $A = 0$ and that given $A = 1$. By contrast, in the quantum case the values of A correspond to different measurements that are used to generate X, and the associated variables $X_{A=0}$ and $X_{A=1}$ may not co-exist. It hence does not make sense to consider a joint distribution over $X_{A=0}$ and $X_{A=1}$ in this case. We therefore only consider the subsets of the observed variables that co-exist

$$\mathcal{S} := \{X_0,\ X_1,\ Y_0,\ Y_1,\ X_0Y_0,\ X_0Y_1,\ X_1Y_0,\ X_1Y_1\}, \tag{3}$$

where we use the short form X_0Y_0 for the set $\{X_0, Y_0\}$ etc. Any nontrivial inequalities derived for the co-existing sets in the classical case can admit quantum or GPT violations.

C. The (2,2,2,2) Bell scenario in probability space

For the bipartite Bell causal structure of Fig. 1(a), the set of all observed distributions $p_{XY|AB}$ that can arise when Λ is classical corresponds to the set of correlations that admit a local hidden variable model, i.e., the set of distributions that have the form

$$p_{XY|AB} = \int_\Lambda d\Lambda\ p_\Lambda\ p_A\ p_B\ p_{X|A\Lambda}\ p_{Y|B\Lambda}. \tag{4}$$

In this work, we will refer to such correlations either as *local* or as *classical* and denote the set of all such distributions $\mathcal{L}$. We also use $\mathcal{L}^{(2,2,2,2)}$ to denote the local distributions in the (2,2,2,2) case (and analogously for other cases).

The set of local correlations form a convex polytope, which can be specified in terms of a finite set of Bell inequalities, each a necessary condition for classicality. In the (2,2,2,2) case, there are eight extremal Bell inequalities (facets of the local polytope). One of these has the form

$$\begin{aligned} I_{\text{CHSH}} := {} & p(X = Y|A = 0, B = 0) + p(X = Y|A = 0, B = 1) \\ & + p(X = Y|A = 1, B = 0) \\ & + p(X \neq Y|A = 1, B = 1) \leqslant 3 \end{aligned} \tag{5}$$

and the other seven are equivalent under local relabellings [21]. We denote these by I^k_{CHSH} for $k \in [8]$, where $I^1_{\text{CHSH}} := I_{\text{CHSH}}$ and $[n]$ stands for the set $\{1, 2, \ldots, n\}$, where n is a positive integer. This provides the facet description of the (2,2,2,2) local polytope. One can also express I_{CHSH} in matrix form using

$$M_{\text{CHSH}} = \left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right], \tag{6}$$

so that the Bell inequality can be written $\text{tr}(M^T_{\text{CHSH}} P) \leqslant 3$, where P is the matrix form of the distribution and T denotes the transpose.

In the vertex picture, the (2,2,2,2) local polytope has 16 local deterministic vertices and the (2,2,2,2) nonsignalling polytope shares the vertices of the local polytope and has eight more: the Popescu-Rohrlich (PR) box and seven distinct local relabellings [21,22]. The PR box distribution satisfies $X \oplus Y = A.B$ and has the form

$$p_{\mathrm{PR}} = \left[\begin{array}{cc|cc} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \hline \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right]. \tag{7}$$

We denote the eight extremal nonsignalling vertices equivalent under local relabellings to p_{PR} by p^k_{PR}, $k \in [8]$ where $p^1_{\mathrm{PR}} := p_{\mathrm{PR}}$. Note that the eight CHSH inequalities $\{I^k_{\mathrm{CHSH}}\}$ are in one-to-one correspondence with these eight extremal nonsignalling points i.e., each p^k_{PR} violates exactly one CHSH inequality and each CHSH inequality is violated by exactly one p^k_{PR}.

D. Entropic inequalities and postselection

In Ref. [16], Braunstein and Caves derived a set of constraints on the postselected causal structure of Fig. 1(b) and showed that these constraints can be violated by quantum correlations. To discuss these we introduce the notion of *entropic classicality*. For every distribution $p_{XY|AB}$ in the Bell causal structure [Fig. 1(a)], we can associate an entropy vector $v \in \mathbb{R}^8$ in the postselected causal structure [Fig. 1(b)] whose components are the Shannon entropies of each element of the set $\mathcal{S}$ [Eq. (3)] distributed according to $p_{X_aY_b} := p_{XY|A=a,B=b}$. Let $\mathbf{H}$ be the map that takes the observed distribution to its corresponding entropy vector in the postselected causal structure.

Definition 1. (Entropic classicality) An entropy vector $v \in \mathbb{R}^8$ is *classical* with respect to the bipartite Bell causal structure [Fig. 1(a)] if there exists a classical distribution $p_{XY|AB} \in \mathcal{L}$ such that $\mathbf{H}(p_{XY|AB}) = v$. Further, a distribution $p_{XY|AB}$ is *entropically classical* if there exists a classical distribution with the same entropy vector, i.e., if there exists a classical entropy vector v such that $\mathbf{H}(p_{XY|AB}) = v$.

The set of all classical entropy vectors forms a convex cone. The distribution p_{PR} [Eq. (7)] is an example of a nonclassical distribution that is entropically classical (see Sec. II E).

We now review how the Braunstein-Caves (BC) inequalities are derived for the case when the observed parentless nodes A and B are binary. In this case, the postselected causal structure 1(b) imposes no additional constraints on the distribution (or entropies) of the observed nodes X_0, X_1, Y_0, and Y_1 because they share a common parent and thus any joint distribution over X_0, X_1, Y_0, and Y_1 can be realized in the causal structure 1(b). By contrast, any correlations in the original causal structure 1(a) must obey the no-signalling constraints over the observed nodes A, B, X, and Y since A does not influence Y and B does not influence X in this causal structure. The inequalities derived by Braunstein and Caves follow by applying properties P1–P3 to the variables $\{X_0, X_1, Y_0, Y_1\}$. The derived relations hold for the classical causal structure (and not necessarily for the quantum and GPT cases) because only in the classical case does it make sense to consider a joint distribution over these four variables that in the quantum and GPT cases do not co-exist (cf. Sec. II B). It is worth remarking that without postselection, no quantum-violatable entropic constraints exist for this causal structure [17]. The BC inequalities are entropic Bell inequalities i.e., they hold for every classical entropy vector in the postselected causal structure 1(b). There are four BC inequalities

$$\begin{aligned} I^1_{\mathrm{BC}} &:= H(X_0Y_0) + H(X_1) + H(Y_1) - H(X_0Y_1) - H(X_1Y_0) - H(X_1Y_1) \leqslant 0, \\ I^2_{\mathrm{BC}} &:= H(X_0Y_1) + H(X_1) + H(Y_0) - H(X_0Y_0) - H(X_1Y_0) - H(X_1Y_1) \leqslant 0, \\ I^3_{\mathrm{BC}} &:= H(X_1Y_0) + H(X_0) + H(Y_1) - H(X_0Y_0) - H(X_0Y_1) - H(X_1Y_1) \leqslant 0, \\ I^4_{\mathrm{BC}} &:= H(X_1Y_1) + H(X_0) + H(Y_0) - H(X_0Y_0) - H(X_0Y_1) - H(X_1Y_0) \leqslant 0. \end{aligned} \tag{8}$$

It has been shown in Ref. [27] that these four inequalities are complete in the following sense (the lemma below is implied by corollary V.3 in Ref. [27]).

Lemma 1. A distribution in the postselected Bell scenario with binary A and B is entropically classical if and only if it satisfies the four BC inequalities (8).

It turns out that in the (2,2,2,2) Bell scenario, nonclassical distributions that do not violate the BC inequalities can be made to do so with some additional postprocessing, as shown in Ref. [19]. We review this result below before analyzing the same question in the (2,2,3,3) scenario.

E. Detecting nonclassicality in the (2,2,2,2) Bell scenario in entropy space

The current section summarizes the relevant results of Ref. [19] regarding the sufficiency of entropic inequalities in the (2,2,2,2) scenario. As previously mentioned, it is possible for a nonclassical distribution to have the same entropy vector as a classical one and hence to be entropically classical. For example, the maximally nonclassical distribution in probability space, p_{PR} [Eq. (7)] is entropically classical since it has the same entropy vector as the classical distribution

$$p_{\mathrm{C}} = \left[\begin{array}{cc|cc} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \hline \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}\right] \tag{9}$$

and hence cannot violate any of the BC inequalities.[3] However, the distribution $\frac{1}{2}p_{\mathrm{PR}} + \frac{1}{2}p_{\mathrm{C}}$ maximally violates $I^4_{\mathrm{BC}} \leqslant 0$

[3] p_{PR} and p_{C} are related by a permutation of the entries in the bottom right 2×2 block and entropies are invariant under such permutations.

attaining a value of ln 2. That convex mixtures of nonviolating distributions can lead to a violation is due to the fact that entropic inequalities are nonlinear in the underlying probabilities (in contrast to the facet Bell inequalities in probability space).

In Ref. [19], it was shown that such a procedure is possible for every nonclassical distribution in the $(d, d, 2, 2)$ Bell scenario with $d \geqslant 2$ i.e., for every such distribution, there exists an LOSR transformation such that the resultant distribution violates a Shannon entropic BC inequality (8). Thus nonclassicality can be detected in this scenario by processing the observed correlations (in a way that cannot generate nonclassicality) before using a BC entropic inequality on the result. In this sense the BC entropic inequalities provide a necessary and sufficient test for nonclassicality in these scenarios.

In more detail, for the (2,2,2,2) case, this works as follows. First, one defines a special class of distributions, *isotropic distributions*, as follows for some $k \in [8]$ and $\epsilon \in [0, 1]$.

$$p^k_{\text{iso}} = \epsilon p^k_{\text{PR}} + (1 - \epsilon) p_{\text{noise}}, \tag{10}$$

where p_{noise} is white noise i.e., the distribution with all entries equal to $1/4$. In the (2,2,2,2) Bell scenario the isotropic distribution p^k_{iso} is nonclassical if and only if $\epsilon > 1/2$. The LOSR transformation used in [19] involves first transforming the observed distribution into an isotropic distribution through a local depolarization procedure that cannot generate nonclassicality. Second, it is shown that for any nonclassical isotropic distribution i.e., a p^k_{iso} with $\epsilon > 1/2$, there exists a classical distribution p^k_{C} such that the distribution $p^k_{\text{E},v} = v p^k_{\text{iso}} + (1 - v) p^k_{\text{C}}$ violates one of the BC entropic inequalities for sufficiently small $v > 0$. In particular, the value of I^k_{BC} for $p^k_{\text{E},v}$ can be expanded for small v as

$$I^k_{\text{BC}} \approx \frac{v}{\ln 4}[f(\epsilon) - (4\epsilon - 2)\ln v], \tag{11}$$

where $f(\epsilon)$ is a function of ϵ, independent of v (see [19] for details). Thus, for any $\epsilon > 1/2$, the corresponding isotropic distributions are nonclassical and taking v arbitrarily small can make I^k_{BC} positive which is a violation of the entropic inequality. We summarize the main result of Ref. [19] for (2,2,2,2) Bell scenarios in the following theorem (which is implicit in Ref. [19]).

Theorem 2. For every nonclassical distribution, $p_{XY|AB}$ in the (2,2,2,2) Bell scenario, there exists an LOSR transformation $\mathcal{T}$, such that $\mathcal{T}(p_{XY|AB})$ violates one of the BC entropic inequalities (8).

One of the aims of the present paper is to study whether this result extends to the case where the number of outcomes per party is more than two. In general, the $(2, 2, d, d)$ Bell polytope for $d > 2$ has new, distinct classes of Bell inequalities and extremal nonsignalling vertices other than the CHSH inequalities and the PR boxes. In the following, we analyze this problem for the $d = 3$ case, for which it is helpful to first describe the (2,2,3,3) scenario in probability space.

F. The (2,2,3,3) Bell scenario in probability space

In the (2,2,3,3) Bell scenario, there are two classes of Bell inequalities that completely characterize the local polytope: the CHSH inequalities and the I_{2233} inequalities [10,11] of which a representative example is

$$\begin{aligned} I_{2233} := [&p(X = Y|A = 0, B = 1) \\ &+ p(X = Y - 1|A = 1, B = 1) \\ &+ p(X = Y|A = 1, B = 0) \\ &+ p(X = Y|A = 0, B = 0)] \\ &- [p(X = Y - 1|X = 0, B = 1) \\ &+ p(X = Y|A = 1, B = 1) \\ &+ p(X = Y - 1|A = 1, B = 0) \\ &+ p(X = Y + 1|A = 0, B = 0)] \leqslant 2 \end{aligned} \tag{12}$$

where all the random variables take values in $\{0, 1, 2\}$ and all additions and subtractions of the random variables are modulo 3. In matrix form a representative CHSH-type inequality and I_{2233} are

$$M^{(2,2,3,3)}_{\text{CHSH}} = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \end{array}\right) \quad \text{and}$$

$$M_{I_{2233}} = \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & -1 & 0 \\ -1 & 1 & 0 & 0 & 1 & -1 \\ 0 & -1 & 1 & -1 & 0 & 1 \\ \hline 1 & -1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 \\ -1 & 0 & 1 & 1 & 0 & -1 \end{array}\right). \tag{13}$$

The (2,2,3,3) local polytope has a total of 1116 facets, 36 of which correspond to positivity constraints, 648 to CHSH facets (these are equivalent to first coarse-graining two of the outputs into one (for each party and each input) and then applying one of the eight (2,2,2,2) CHSH inequalities[4]), and the remaining 432 are I_{2233}-type [28] (we label these I^i_{2233} for $i \in \{1, 2, \ldots, 432\}$ with $I^1_{2233} = I_{2233}$).

The facets of the no-signalling polytope correspond to positivity constraints. Converting this facet description to the vertex description (e.g., using the PORTA software [29]) one can obtain all the vertices of the (2,2,3,3) nonsignalling polytope. This comprises 81 local deterministic vertices, 648 PR-box type vertices, and 432 extremal nonsignalling vertices (for each of the I_{2233} inequalities there is one of the latter that gives maximal violation). We call these new vertices the I_{2233} vertices. The specific vertex that maximally violates (12) is

$$p_{\text{NL}} := \left(\begin{array}{ccc|ccc} \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ \hline \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 \end{array}\right) \tag{14}$$

[4]For instance, evaluating the CHSH-type inequality represented by $M^{(2,2,3,3)}_{\text{CHSH}}$ is equivalent to coarse graining the distribution by always mapping outcomes 1 and 2 to 1 and then evaluating M_{CHSH} (6).

The 432 I_{2233} vertices of the (2,2,3,3) nonsignalling polytope are related to each other through local relabellings.

III. RESULTS 1: THE (2,2,3,3) SCENARIO IN PROBABILITY SPACE

In this section, we compute the vertex description of the *CHSH-classical polytope*, $\Pi_{\text{CHSH}}^{(2,2,3,3)}$, i.e., the polytope whose facets are the 648 CHSH inequalities and the positivity constraints. As previously mentioned, this will be the main region of interest in the remainder of this work since the nonclassicality of distributions not belonging to this region can always be certified using Shannon-entropic inequalities (corollary 3). The following result allows us to significantly speed up the vertex enumeration problem.

Proposition 1. Every nonclassical distribution in $\Pi_{\text{CHSH}}^{(2,2,3,3)}$ violates only one I_{2233} inequality.

Proof. Let $i \in \{2, 3, \ldots, 432\}$ and consider the linear program that maximizes the value of $\epsilon \geqslant 0$ subject to there existing a no-signalling distribution that (1) violates $I_{2233}^1 \leqslant 2$ and $I_{2233}^i \leqslant 2$ by at least ϵ, i.e., $I_{2233}^1 - \epsilon \geqslant 2$ and $I_{2233}^i - \epsilon \geqslant 2$; and (2) satisfies all CHSH-type inequalities.

We run over $i \in \{2, 3, \ldots, 432\}$ and check that in all cases either the output of this linear program is $\epsilon = 0$ (meaning that the two I_{2233} inequalities can be jointly saturated but not violated) or that the program is infeasible (the two I_{2233} inequalities cannot even be jointly saturated). By symmetry it follows that no pair of I_{2233} inequalities can be simultaneously violated when all the CHSH-type inequalities are satisfied. ■

Note that in the (2,2,3,3) scenario there exist extremal nonsignalling distributions that violate multiple Bell inequalities. For example, the distributions p_{NL} [Eq. (14)] and p_{NL}^* [Eq. (18)] violate $I_{2233} \leqslant 2$ [cf. (12)] although only p_{NL} violates it maximally. By symmetry, p_{NL} also violates another I_{2233} inequality. In addition, p_{NL} violates the CHSH-type inequality whose evaluation is equivalent to applying the output coarse-graining $0 \mapsto 0$, $1 \mapsto 1$ and $2 \mapsto 1$ for each party and then evaluating (5). This is in contrast to the (2,2,2,2) scenario where there is a one-to-one correspondence between the extremal nonsignalling vertices and the CHSH inequalities in the sense that each such vertex violates exactly one CHSH inequality.[5]

Due to the symmetries, all the vertices of $\Pi_{\text{CHSH}}^{(2,2,3,3)}$ can be enumerated by first finding all the vertices for which the I_{2233} inequality of Eq. (12) is saturated or violated i.e., $I_{2233} \geqslant 2$, and taking the orbit of these vertices under local relabellings and exchange of the parties. The vertex enumeration for this case yields 47 extremal points of which 30 are the local deterministic points that saturate $I_{2233} \geqslant 2$ and 17 are nonclassical points that violate only this inequality. These are listed in Table I. By applying all symmetries and removing duplicate vertices we find that $\Pi_{\text{CHSH}}^{(2,2,3,3)}$ has 7425 vertices (including the 81 local deterministic vertices).

[5]Note that this correspondence breaks down in the (2,2,3,3) scenario where it is possible for a CHSH-type vertex to violate multiple CHSH-type inequalities (these correspond to the same 2-outcome CHSH inequality after coarse-graining).

IV. RESULTS 2: THE (2,2,3,3) SCENARIO IN ENTROPY SPACE

We now investigate whether entropic inequalities are necessary and sufficient for nonclassicality in the (2,2,3,3) Bell scenario. In $(d, d, 2, 2)$ scenarios with $d \geqslant 2$, only $2d$ Shannon entropic inequalities are required for the Shannon entropic characterization of the scenario [27,30]; in the (2,2,2,2), these are the four inequalities of (8). It may at first seem surprising that these can always be used to decide whether a distribution is classical because the number of extremal Bell inequalities grows very rapidly in d in the $(d, d, 2, 2)$ scenario [15], and deciding whether a distribution is classical is NP-complete [31]. The reduction in the number of inequalities in entropy space is compensated by the need to identify a suitable postprocessing operation (of which there are uncountably many possibilities) in order to detect violations.

The first observation is a corollary of theorem 2.

Corollary 3. Let $p_{XY|AB}$ be a distribution in the (2,2,3,3) Bell scenario that violates at least one CHSH-type inequality. Then there exists an LOSR transformation $\mathcal{T}$, such that $\mathcal{T}(p_{XY|AB})$ violates one of the BC entropic inequalities (8).

Proof. For each CHSH-type inequality in the (2,2,3,3) scenario, there exists a coarse-graining in which two of the outcomes are mapped to one (for each party and each input) such that for any initial distribution in the (2,2,3,3) scenario that violates the CHSH-type inequality the coarse-grained distribution violates one of the CHSH-inequalities in the (2,2,2,2) scenario. Hence, for the given $p_{XY|AB}$, after applying the corresponding coarse-graining for the violated CHSH-type inequality, followed by the LOSR operation from theorem 2, we violate one of the BC entropic inequalities. ■

This corollary means that we can limit our analysis to $\Pi_{\text{CHSH}}^{(2,2,3,3)}$, the polytope in which all the CHSH inequalities are satisfied, and, in particular, the nonclassical region of this. This is the region in which one of the I_{2233} inequalities is violated.

In going from the (2,2,2,2) to (2,2,3,3) scenario, a new class of inequalities (the I_{2233} inequalities) become relevant in probability space but the entropic characterization remains unchanged, since entropic inequalities do not depend on the number of measurement outcomes. It is natural to ask whether all nonclassical distributions in the (2,2,3,3) scenario that satisfy all the CHSH inequalities cannot be certified entropically. However, this is not the case as shown by the following proposition.

Proposition 2. The polytope $\Pi_{\text{CHSH}}^{(2,2,3,3)}$ is not entropically classical.

Proof. Consider the distribution

$$p_e := \frac{1}{50}\left[\begin{array}{ccc|ccc} 21 & 0 & 0 & 21 & 0 & 0 \\ 0 & 2 & 0 & 1 & 1 & 0 \\ 11 & 0 & 16 & 0 & 1 & 26 \\ \hline 31 & 0 & 0 & 20 & 1 & 10 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 16 & 1 & 1 & 15 \end{array}\right]. \tag{15}$$

This is formed by mixing the nonlocal extremal point number 8 of $\Pi_{\text{CHSH}}^{(2,2,3,3)}$ (see Table I) with the three local deterministic points 18, 26, and 47 with respective weights $1/10$, $3/10$, $1/5$, and $2/5$, and hence is in $\Pi_{\text{CHSH}}^{(2,2,3,3)}$. It achieves a I_{BC}^4 value

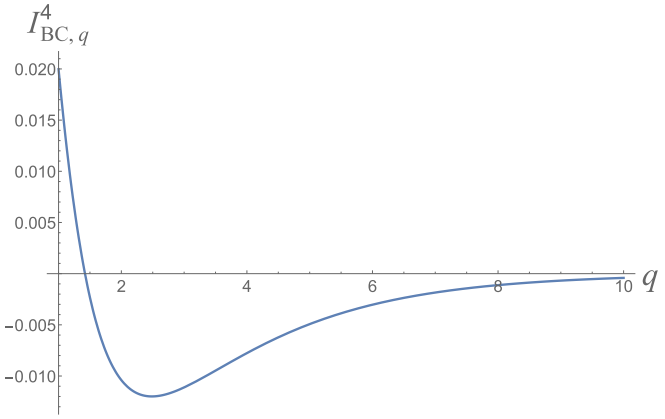

FIG. 2. Plot of the $I^4_{\mathrm{BC},q}$ value as a function of the Tsallis parameter q for the distribution p_e [Eq. (15)]. As seen from the plot, the distribution violates the BC inequality $I^4_{\mathrm{BC},q} \leqslant 0$ for q values between 1 (Shannon case) and just under 1.5 and the violation is maximum in the Shannon case, indicating that it is preferable to use Shannon rather than Tsallis entropies with $q > 1$ in this case.

of 0.0199733, in violation of $I^4_{\mathrm{BC}} \leqslant 0$, so is not entropically classical. ■

We remark that by mixing with more local deterministic distributions and varying the weights, larger violations of $I^4_{\mathrm{BC}} \leqslant 0$ can be found; the distribution p_e used in the previous proposition was chosen for its relative simplicity.

Interestingly, we find that the Shannon entropic BC inequalities appear to give the largest violation among the Tsallis entropic inequalities for $q \geqslant 1$ when applied to p_e. This can be seen in Fig. 2.

In light of proposition 2, it is natural to ask whether the nonclassicality of all distributions in $\Pi^{(2,2,3,3)}_{\mathrm{CHSH}}$ can be detected through entropic inequalities. We find numerical evidence that suggests the contrary, i.e., that there are nonclassical distributions in $\Pi^{(2,2,3,3)}_{\mathrm{CHSH}}$ whose nonclassicality cannot be detected through entropic inequalities using a general class of postprocessing operations, and hence that these entropic inequalities are not sufficient for detecting nonclassicality in the (2,2,3,3) scenario. Before presenting these results, we briefly overview the postprocessing operations considered in this work.

A. Post-processing operations

In this paper, we study whether entropic inequalities can always detect nonclassicality in the (2,2,3,3) Bell scenario. In order to do so we could in principle consider applying any NCNG operation to the distribution prior to evaluating the entropic inequality. However, due to the difficulty in dealing with arbitrary NCNG operations, we consider the subset of these corresponding to LOSR + E instead. In Ref. [32], it was shown that all LOSR operations can be generated by convex combinations of local deterministic operations. These can be thought of in the following way. Each party first does a deterministic function on their input, uses the result as the input to their device, then does a deterministic function on their input and the output of their device to form the final output. All such operations correspond to local relabellings and local coarse-grainings. Note that deterministic classical distributions can be formed as a special case of coarse-graining (a local deterministic distribution is formed when each party coarse-grains all of their outputs to one output for each of their inputs). For the distributions we consider for our main conjectures, it turns out that all the coarse-grainings give rise to local distributions (cf. proposition 6), so, by considering mixing with deterministic classical distributions, local relabelling, and exchange of parties we can cover all LOSR + E operations. We hence start by separately considering mixing with classical distributions, and then consider relabelling and exchange of parties.

B. Mixing with classical distributions

Analogously to the (2,2,2,2) case, we can define a family of distributions $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} = \epsilon p_{\mathrm{NL}} + (1-\epsilon)p^{(2,2,3,3)}_{\mathrm{noise}}$, where $p^{(2,2,3,3)}_{\mathrm{noise}}$ is the uniform distribution with all entries equal to $1/9$ and $\epsilon \in [0,1]$. This class of distributions is isotropic in the sense that the marginal distributions are uniform for each input of each party. In order to show the insufficiency of entropic inequalities, one needs to identify at least one nonclassical distribution whose nonclassicality cannot be detected through entropic inequalities. We will discuss this for the class $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ and only consider distributions of this form in the rest of the paper. Further, without loss of generality, we consider only the BC inequality $I^4_{\mathrm{BC}} \leqslant 0$ in what follows (by symmetry all the arguments will also hold for isotropic distributions corresponding to relabelled versions of p_{NL} and the corresponding BC inequalities).

1. Using Shannon entropy

In the entropic picture of the (2,2,3,3) scenario, the four BC inequalities (8) still hold (these are valid independently of the cardinality of the random variables). Again, analogously to the (2,2,2,2) case, the maximally nonlocal distribution, p_{NL} has the same entropy vector as the classical distribution

$$p^{(2,2,3,3)}_{\mathrm{C}} = \left[\begin{array}{ccc|ccc} \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ \hline \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \end{array}\right] \tag{16}$$

(amongst others). The distribution p_{NL} is hence entropically classical. However, in contrast to the (2,2,2,2) case, we have evidence suggesting that there are values of ϵ for which $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ is nonclassical, but such that the mixture $v p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + (1-v)p_{\mathrm{L}}$ is entropically classical for all classical distributions p_{L} and all $v \in [0,1]$, i.e., there exist nonclassical distributions in the (2,2,3,3) scenario for which mixing with classical distributions never gives rise to a nonclassical entropy vector.

We begin by considering mixing $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ with $p^{(2,2,3,3)}_{\mathrm{C}}$ in analogy with the treatment of the (2,2,2,2) case. Although we have not fully proven this, from our numerics, this mixing appears to be optimal in the sense that when it does not allow for entropic violations, no other mixing can either. This allows us to identify a range of ϵ for which the mixture $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ is nonclassical, yet appears to remain entropically classical

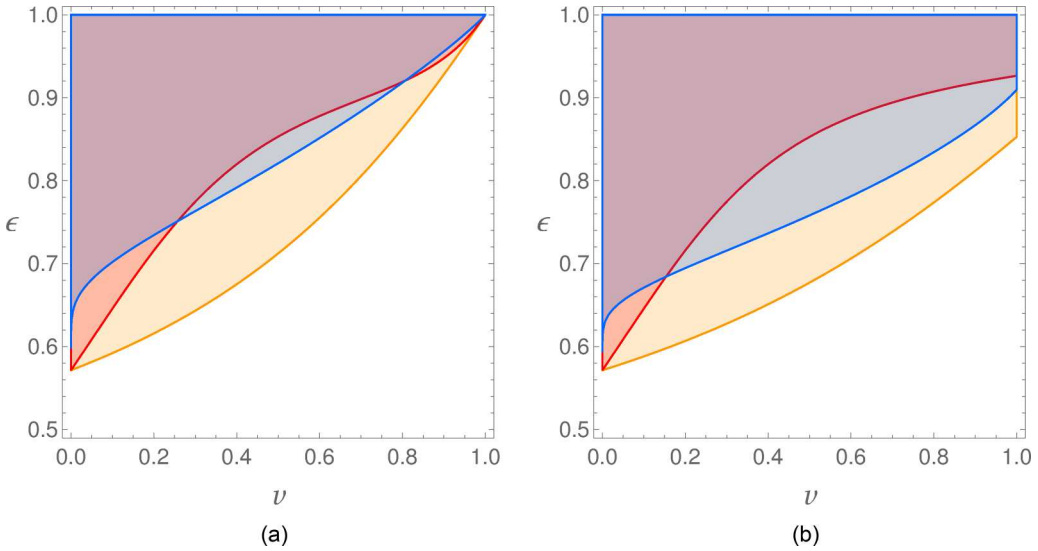

FIG. 3. The regions in the v-ϵ plane where the Shannon entropic inequality $I^4_{\mathrm{BC},1} := I^4_{\mathrm{BC}} \leqslant 0$ (blue), the Tsallis entropic inequality, $I^4_{\mathrm{BC},q} \leqslant 0$ for $q = 2$ (orange) and $q = 8$ (red) are violated by the distributions (a) $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v} = v p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + (1-v) p^{(2,2,3,3)}_{\mathrm{C}}$ (b) $\tilde{p}^{(2,2,3,3)}_{\mathrm{E},\epsilon,v} = v \tilde{p}^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + (1-v) p^{(2,2,3,3)}_{\mathrm{C}}$. Here, $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} = \epsilon p_{\mathrm{NL}} + (1-\epsilon) p^{(2,2,3,3)}_{\mathrm{noise}}$ and $\tilde{p}^{(2,2,3,3)}_{\mathrm{iso},\epsilon} = \epsilon(1/2 p_{\mathrm{NL}} + 1/2 p^{*}_{\mathrm{NL}}) + (1-\epsilon) p^{(2,2,3,3)}_{\mathrm{noise}}$. For both (a) and (b), $I^4_{\mathrm{BC},q} \leqslant 0$ is not violated when $\epsilon \leqslant 4/7 \approx 0.5714$ but for $\epsilon > 4/7$, there is a violation of this inequality for a larger range of v values in the latter case, and also for a larger range in the $q = 2$ case as compared to the other two cases.

even when mixed with arbitrary classical distributions. We begin with two propositions whose proofs can be found in Appendix C.

Proposition 3. $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ is nonclassical if and only if $\epsilon > 1/2$. Further, for $\epsilon \leqslant 4/7$, $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ satisfies all the CHSH-type inequalities, while for $\epsilon > 4/7$ it violates at least one CHSH-type inequality.

By analogy with the (2,2,2,2) case, we consider the violation of I^4_{BC} attainable by mixing $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ with $p^{(2,2,3,3)}_{\mathrm{C}}$. We find that for $\epsilon \in (1/2, 4/7]$, $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ is nonclassical but does not violate any of the BC inequalities. As shown in the above proposition, these distributions are in the CHSH-classical polytope $\Pi^{(2,2,3,3)}_{\mathrm{CHSH}}$ and hence lie in our region of interest.

Proposition 4. For $\epsilon \leqslant 4/7$, $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v} = v p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + (1-v) p^{(2,2,3,3)}_{\mathrm{C}}$ does not violate any of the BC entropic inequalities (8) for any $v \in [0, 1]$. However, for all $\epsilon > 4/7$, there exists a $v \in [0, 1]$ such that the entropic inequality $I^4_{\mathrm{BC}} \leqslant 0$ is violated by $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v}$.

The second part of this proposition already follows from corollary 3 and proposition 3.

Corollary 4. For $\epsilon \leqslant 4/7$, $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v} = v p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + (1-v) p^{(2,2,3,3)}_{\mathrm{C}}$ is entropically classical for all $v \in [0, 1]$.

Proof. This follows from proposition 4 and lemma 1. ■

While proposition 4 shows that the proof strategy of Ref. [19] does not directly generalize to all nonclassical distributions in the (2,2,3,3) case, it does not rule out the possibility that there may exist other mixings with classical distributions that could transform $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ for $\epsilon \in (1/2, 4/7]$ into a distribution that violates one of the BC inequalities. To investigate this, we can consider the polytope formed by mixing $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ with classical distributions for some $\epsilon \leqslant 4/7$, i.e., the polytope $\mathrm{Conv}(\{p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}\} \bigcup \{p^{\mathrm{L},k}\}_k)$, where $\{p^{\mathrm{L},k}\}_k$ denotes the set of all (81) local deterministic vertices of the (2,2,3,3) Bell-local polytope, $\{p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}\}$ is a set with a single element and Conv() denotes the convex hull. We considered several values of $\epsilon \leqslant 4/7$ and numerically optimized the entropic expression I^4_{BC} over the nonclassical region of this polytope[6] for each[7] and were unable to find violations. The optimization involves a nonlinear objective function with linear constraints. Hence, it is possible that the numerical approach missed the global optimum. Nevertheless, this is evidence for the following conjecture and is presented in more detail in Appendix B. Proposition 4 along with Fig. 3 and the evidence in Appendix B also suggest this conjecture.

Conjecture 1. Let $\epsilon \leqslant 4/7$. For all mixtures of the distribution $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ with classical distributions in the (2,2,3,3) Bell scenario, the resulting distribution is entropically classical, i.e., all distributions in $\mathrm{Conv}(\{p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}\} \bigcup \{p^{\mathrm{L},k}\}_k)$ are entropically classical.

The interesting cases of conjecture 1 are for nonclassical distributions (i.e., for $\epsilon > 1/2$), and the most relevant of these are those that can be achieved in quantum theory. The next remark addresses this case.

Remark 1. There exist nonclassical quantum distributions that lie in the polytope $\mathrm{Conv}(\{p^{(2,2,3,3)}_{\mathrm{iso},\epsilon=4/7}\} \bigcup \{p^{\mathrm{L},k}\}_k)$ and in this case, our results suggest that the nonclassicality of the corresponding distributions cannot be detected through entropic inequalities as we now explain. Let p_{QM} be the quantum distribution from [11, Eq. (14) with $d = 3$] with Bob's inputs relabelled. This violates $I^4_{\mathrm{BC}} \leqslant 0$ through mixing with $p^{(2,2,3,3)}_{\mathrm{C}}$. Consider then mixing p_{QM} with uniform noise $p^{(2,2,3,3)}_{\mathrm{noise}}$ to obtain $p_{\mathrm{mix}}(u) := u p_{\mathrm{QM}} + (1-u) p^{(2,2,3,3)}_{\mathrm{noise}}$

[6]To restrict to the nonclassical region, it is sufficient to mix with a subset of these 81 locals—see Appendix B for more detail.

[7]Note however that it is enough that these results hold for one value of $\epsilon \in (1/2, 4/7]$ in order to conclude that entropic inequalities are not sufficient for detecting nonclassicality in this scenario.

($u \in [0, 1]$). We found that for some values of u (e.g., $u = 7/10$), $p_{\text{mix}}(u)$ is nonclassical. Further, $p_{\text{mix}}(u)$ is quantum achievable since it can be obtained from the density operator $u|\psi'\rangle\langle\psi'| + (1-u)\frac{\mathbb{I}}{9}$ (where $|\psi'\rangle$ is the two qutrit state producing p_{QM}) and the same quantum measurements that produce p_{QM} from $|\psi'\rangle$.

2. Using Tsallis entropies

Given the results (proposition 4 and conjecture 1) of the previous section for Shannon entropic inequalities, a natural question is whether other entropic measures can provide an advantage over the Shannon entropy in detecting nonclassicality. Here, we look at Tsallis entropies and find that similar results hold in this case as well, suggesting that Tsallis entropies also do not allow us to completely solve the problem.

The properties of monotonicity, strong subadditivity, and the chain rule are sufficient to derive the BC inequalities, which hence also hold for Tsallis entropy when $q \geqslant 1$. [Other generalized entropies such as Rényi or min/max entropies do not satisfy one or more of these properties in general and it is not clear whether the analogues of (8) hold for these.] In other words, for all $q \geqslant 1$, we have

$$\begin{aligned}
I^1_{\text{BC},q} &= S_q(X_0Y_0) + S_q(X_1) + S_q(Y_1) - S_q(X_0Y_1) - S_q(X_1Y_0) - S_q(X_1Y_1) \leqslant 0,\\
I^2_{\text{BC},q} &= S_q(X_0Y_1) + S_q(X_1) + S_q(Y_0) - S_q(X_0Y_0) - S_q(X_1Y_0) - S_q(X_1Y_1) \leqslant 0,\\
I^3_{\text{BC},q} &= S_q(X_1Y_0) + S_q(X_0) + S_q(Y_1) - S_q(X_0Y_0) - S_q(X_0Y_1) - S_q(X_1Y_1) \leqslant 0,\\
I^4_{\text{BC},q} &= S_q(X_1Y_1) + S_q(X_0) + S_q(Y_0) - S_q(X_0Y_0) - S_q(X_0Y_1) - S_q(X_1Y_0) \leqslant 0,
\end{aligned} \tag{17}$$

and we refer to these as the Tsallis entropic BC inequalities. Entropic classicality in Tsallis entropy space can be defined analogously to definition 1, in terms of Tsallis entropy vectors over the set of variables $\mathcal{S}$ [Eq. (3)]. We say that a distribution is *q-entropically classical* if its entropy vector written in terms of the Tsallis entropy of order q is achievable using a classical distribution. In the case of the Shannon entropy, we used the fact (lemma 1) that the BC inequalities (8) are known to be necessary and sufficient for entropic classicality for 2-input Bell scenarios [27]. However, it is not clear if the result of Ref. [27] generalizes to Tsallis entropies for $q > 1$. Thus our results in the Tsallis case are weaker than those for Shannon, being stated only for the BC inequalities. We leave the generalization to arbitrary Tsallis entropic inequalities as an open problem.

Proposition 5. For $\epsilon \leqslant 4/7$, $p^{(2,2,3,3)}_{\text{E},\epsilon,v} = v p^{(2,2,3,3)}_{\text{iso},\epsilon} + (1-v)p^{(2,2,3,3)}_{\text{C}}$ does not violate any of the Tsallis BC inequalities (17) for any $v \in [0, 1]$ and $q > 1$. However, for $\epsilon > 4/7$ and every $q > 1$, there always exists a $v \in [0, 1]$ such that the entropic inequality $I^4_{\text{BC},q} \leqslant 0$ is violated by $p^{(2,2,3,3)}_{\text{E},\epsilon,v}$.

We refer the reader to Appendix C for a proof of this proposition. To investigate the extension to other mixings, we tried the same computational procedure (see Appendix B) as in the Shannon case. We found no violation of the Tsallis entropic BC inequalities for any mixings of $p^{(2,2,3,3)}_{\text{iso},\epsilon}$ with classical distributions, for several values of $q > 1$ and $\epsilon \in (1/2, 4/7]$, leading to the following conjecture, which is similar to conjecture 1.

Conjecture 2. Let $\epsilon \leqslant 4/7$. For all mixtures of the distribution $p^{(2,2,3,3)}_{\text{iso},\epsilon}$ with classical distributions in the (2,2,3,3) Bell scenario, the resulting distribution does not violate any of the Tsallis entropic BC inequalities for any $q > 1$, i.e., all distributions in $\text{Conv}(\{p^{(2,2,3,3)}_{\text{iso},\epsilon}\} \bigcup \{p^{\text{L},k}\}_k)$ satisfy the Tsallis entropic BC Inequalities (17) for all $q > 1$.

Figure 3(a) shows the values of ϵ and v for which $I^4_{\text{BC},q}$ (for $q = 1,\ 2,$ and 8) evaluated with $p^{(2,2,3,3)}_{\text{E},\epsilon,v}$ is positive, which is also suggestive of this conjecture.

Remark 2. Any impossibility result for the (2,2,3,3) scenario also holds in the $(2, 2, d, d)$ case for $d > 3$ because the former is always embedded in the latter, i.e., every distribution in the (2,2,3,3) scenario has a corresponding distribution in all the $(2, 2, d, d)$ scenarios with $d > 3$ which can be obtained by assigning a zero probability to the additional outcomes. Further, the entropic inequalities (8) remain the same for all these scenarios as they do not depend on the cardinality of the random variables involved. Thus the existence nonclassical distributions for the $d = 3$ case whose nonclassicality cannot be detected by entropic inequalities implies the same result for all $d > 3$.

C. Beyond classical mixings

So far, we only considered mixing with classical distributions to obtain entropic violations and gave evidence that this does not work for some nonclassical distributions in the (2,2,3,3) scenario. This motivates us to study whether using arbitrary LOSR + E operations allows us to detect this nonclassicality through entropic violations. We show in this section that if conjectures 1 and 2 hold then they also hold for all LOSR + E operations. First consider the following example.

The maximum possible violation of the BC inequalities in the (2,2,2,2) case is $I^4_{\text{BC}} = \ln 2$ [19]. This is derived by considering only Shannon inequalities within the coexisting sets, and the bound that the maximum entropy of a binary variable is ln 2. An analogous proof holds in the (2,2,3,3) case, except that the bound is then ln 3. In the former case, we have $p_{\text{E},\epsilon=1,v=1/2} = \frac{1}{2}p_{\text{PR}} + \frac{1}{2}p_{\text{C}}$, which maximally violates $I^4_{\text{BC}} \leqslant 0$, while in the latter case, one such distribution is

formed by $(p_{\mathrm{NL}} + p^{*}_{\mathrm{NL}} + p_{\mathrm{C}}^{(2,2,3,3)})/3$, where p^{*}_{NL} is another extremal nonlocal distribution:

$$p^{*}_{\mathrm{NL}} = \left[\begin{array}{ccc|ccc} \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ \hline \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \end{array}\right]. \tag{18}$$

Since the equal mixture $(p_{\mathrm{NL}} + p_{\mathrm{C}}^{(2,2,3,3)})/2$ violates $I^{4}_{\mathrm{BC}} \leqslant 0$ nonmaximally, one may be motivated to use the nonlocal distribution $\tilde{p}_{\mathrm{NL}} = (p_{\mathrm{NL}} + p^{*}_{\mathrm{NL}})/2$ in place of p_{NL} in the definition of $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$, i.e., to take

$$\tilde{p}_{\mathrm{iso},\epsilon}^{(2,2,3,3)} = \epsilon \tilde{p}_{\mathrm{NL}} + (1-\epsilon) p_{\mathrm{C}}^{(2,2,3,3)}.$$

One could then consider whether for $\epsilon \in (1/2, 4/7]$, $\tilde{p}_{\mathrm{E},\epsilon,v}^{(2,2,3,3)} = v\tilde{p}_{\mathrm{iso},\epsilon}^{(2,2,3,3)} + (1-v)p_{\mathrm{C}}^{(2,2,3,3)}$ violates $I^{4}_{\mathrm{BC}} \leqslant 0$. Interestingly, while $\tilde{p}_{\mathrm{E},\epsilon,v}^{(2,2,3,3)}$ violates $I^{4}_{\mathrm{BC}} \leqslant 0$ for a larger range of v values whenever $\epsilon > 4/7$, it does not give any violation (for any value of v) when $\epsilon \leqslant 4/7$, and Propositions 3 and 4 also hold if $\tilde{p}_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$ replaces $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$ [see Fig. 3(b) for an illustration]. The corresponding results also hold for the Tsallis case with $q > 1$, i.e., proposition 5 also holds with $\tilde{p}_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$ replacing $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$ (see also Fig. 3). These suggest that mixing with relabellings in addition to mixing with classical distributions may also not help to violate entropic inequalities when $\epsilon \leqslant 4/7$.

In the remainder of this section, we consider the full set of LOSR + E operations. We first note that all input coarse-grainings of $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$ result in local distributions (there are no Bell inequalities if one party has only one input). Similarly, considering output coarse-grainings, whenever three outcomes are mapped to one the resulting distribution is always classical because there are no Bell inequalities if one party always makes a fixed outcome for one of their inputs. We henceforth only consider coarse-grainings that take two outcomes to one. We can choose two of the three outcomes to combine into one for each party and each local input. For the four input choices $\{A=0, A=1, B=0, B=1\}$, there are 81, 108, 54, and 12 distinct coarse-grainings of this type when the outcomes of either 4, 3, 2, or 1 input choices are coarse-grained. Thus there are a total of 255 coarse-grainings that remain.

If we apply all such coarse-grainings to $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$, this generates 255 possible distributions that we denote $\{p_{\epsilon}^{\mathrm{CG},i}\}_i$, $i \in [255]$. There are also 432 distinct local relabellings of $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$, which we denote by $\{p_{\epsilon}^{\mathrm{R},j}\}_j$, $j \in [432]$ (this set includes $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)} = p_{\epsilon}^{\mathrm{R},1}$). Due to symmetries of $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$, it turns out that exchanging parties can be achieved through local relabellings for these distributions, so we do not need to separately consider the exchange in our results pertaining to $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$. The set of all distributions that can be achieved through a convex mixture of $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$ with its coarse-grainings, relabellings and classical distributions is a convex polytope Π_ϵ for each ϵ and is the convex hull of these $255 + 432 + 81 = 768$ points, i.e.,

$$\Pi_\epsilon := \mathrm{Conv}\left(\{p_{\epsilon}^{\mathrm{CG},i}\}_i \bigcup \{p_{\epsilon}^{\mathrm{R},j}\}_j \bigcup \{p^{\mathrm{L},k}\}_k\right).$$

We present the results for the remaining coarse-grainings and relabellings separately below. Firstly, we show that the coarse-grainings $\{p_{\epsilon}^{\mathrm{CG},i}\}_i$ of $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$ are classical if and only if $\epsilon \leqslant 4/7$.

Proposition 6. The distribution $p_{\epsilon}^{\mathrm{CG},i}$ is classical for all i if and only if $\epsilon \leqslant 4/7$.

This is intuitive because $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$ satisfies all the CHSH-type inequalities if and only if $\epsilon \leqslant 4/7$. Since coarse-grainings cannot generate nonclassicality and correspond to reducing the number of outcomes, and since I_{2233} requires three outcomes, after coarse-graining the only relevant thing is whether there is a CHSH-violation. A full proof is given in Appendix C.

Proposition 6 implies that $\Pi_\epsilon = \mathrm{Conv}(\{p_{\epsilon}^{\mathrm{R},j}\}_j \bigcup \{p^{\mathrm{L},k}\}_k)$ $\forall \epsilon \leqslant 4/7$, and that it is not necessary to consider coarse-grainings for such values of ϵ. Our next results are that if conjectures 1 and 2 hold for $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$ for $\epsilon \leqslant 4/7$, then they continue to hold even when we consider arbitrary convex combinations with classical distributions and local relabellings of $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$.

Proposition 7. Let $\epsilon \leqslant 4/7$. If conjecture 1 holds, then every distribution in Π_ϵ is Shannon entropically classical.

Proposition 8. Let $\epsilon \leqslant 4/7$. If conjecture 2 holds, then every distribution in Π_ϵ satisfies the Tsallis entropic BC Inequalities (17) $\forall q > 1$.

These are proven in Appendix C and give the following corollary.

Corollary 5. Let $\epsilon \leqslant 4/7$. If conjectures 1 and 2 hold, then for any operation $\mathcal{O}$ in LOSR + E, $\mathcal{O}(p_{\mathrm{iso},\epsilon}^{(2,2,3,3)})$ does not violate a Shannon or Tsallis ($q > 1$) entropic BC inequality.

V. DISCUSSION

We have provided evidence that there are distributions in the (2,2,3,3) scenario for which arbitrary LOSR + E operations do not enable detection of nonclassicality with any Shannon entropic inequalities or Tsallis entropic BC inequalities. This is in contrast to the (2,2,2,2) scenario [19], where for any nonclassical distribution, there is always a simple LOSR operation that results in a distribution violating one of the Shannon BC inequalities. In order that BC inequalities do not detect nonclassicality we need that the distributions are nonclassical while at the same time satisfying all CHSH-type Bell inequalities. Having found all the vertices that characterize this region, we identify distributions in this region that violate BC entropic inequalities. Thus, the set of all nonclassical distributions in the (2,2,3,3) scenario that cannot be certified through entropic inequalities under LOSR + E postprocessings is not characterized by the CHSH-type inequalities.

Although we considered LOSR + E operations, a natural next question is to what extent the results can be generalized to more general NCNG operations. In particular, there could be a nonlinear NCNG map that allows the entropic BC inequalities to detect a wider range of nonclassical distributions. It would be interesting to see whether for any nonclassical distribution of the form $p_{\mathrm{iso},\epsilon}^{(2,2,3,3)}$ with $1/2 < \epsilon \leqslant 4/7$ (conjectured to be

entropically classical with respect to LOSR + E), one of these more general operations would allow its nonclassicality to be detected entropically. We leave this as an open question.

In Ref. [20], it was shown that (without postprocessing) Tsallis entropic inequalities can detect nonclassicality undetectable by Shannon entropic inequalities in the (2,2,2,2) and (2,2,3,3) Bell scenarios. In the presence of LOSR + E operations, we did not find any advantage of Tsallis entropies over the Shannon entropy in the (2,2,3,3) Bell scenario. In fact, for some nonclassical distributions such as that of Eq. (15), the Shannon entropic inequalities appear to give the largest violations among Tsallis entropies with $q \geqslant 1$ (which corresponds to those for which the BC inequalities can be derived in the classical case). On the other hand, for the family of distributions $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$, our results suggest that the range of ϵ for which postprocessing via mixing with classical distributions enables nonclassicality detection is the same for the Shannon as well as Tsallis entropic BC inequalities for $q > 1$. However, when entropic detection of nonclassicality is possible, using Tsallis entropy can make it easier to do this detection in the sense that there is a wider range of mixings that achieve this (see Fig. 3).[8]

In conclusion, while the entropic approach for detecting nonclassicality is useful in a number of scenarios, it is known to have disadvantages in others. In particular, in the absence of postselection we are not aware of any cases where entropic inequalities can be violated [17,18,33]. Here, we find that the entropic approach also suffers drawbacks in the presence of postselection as it may fail to detect nonclassicality under a natural class of postprocessing operations, both in the case of Shannon and Tsallis entropies. However, this method remains of use since in many cases nonclassicality can be detected using it.

ACKNOWLEDGMENTS

We thank Mirjam Weilenmann for useful discussions. V.V. acknowledges financial support from the Department of Mathematics, University of York. R.C. is supported by EPSRC's Quantum Communications Hub (Grant Nos. EP/M013472/1 and EP/T001011/1) and by an EPSRC First Grant (Grant No. EP/P016588/1).

APPENDIX A: CHARACTERIZATION OF $\Pi^{(2,2,3,3)}_{\mathrm{CHSH}}$ THE CHSH-CLASSICAL REGION

Table I enumerates the 47 extremal points of the polytope $\Pi^{(2,2,3,3)}_{\mathrm{CHSH}}$ that saturate or violate the inequality $I_{2233} \leqslant 2$ (12) while satisfying all the CHSH-type inequalities in the (2,2,3,3) Bell scenario. The first 17 of these are nonclassical while the remaining 30 are local deterministic vertices. Due to proposition 1 and the symmetries of the scenario, the remaining vertices of $\Pi^{(2,2,3,3)}_{\mathrm{CHSH}}$ can be generated by taking the orbit of these vertices under local relabellings and exchange of parties. In Table I, each extremal point is given by a single 36 dimensional vector which corresponds to writing the point in the notation explained in Sec. II A (a 6×6 matrix) and "flattening" it by writing one row after another in order.

APPENDIX B: EVIDENCE FOR CONJECTURES 1 AND 2

In order to check for violations of the Shannon and Tsallis entropic inequalities $I^4_{\mathrm{BC}} \leqslant 0$ and $I^4_{\mathrm{BC},q} \leqslant 0$ that could be obtained by mixing $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ [Eq. (14)] with arbitrary classical distributions, we maximized the left hand sides of these inequalities over the polytope $\mathrm{Conv}(\{p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}\} \bigcup \{p^{\mathrm{L},k}\}_k)$ for some ϵ values in $(1/2, 4/7]$ such as $\epsilon = 4/7, 5/9$ numerically using MATHEMATICA. Note that this polytope contains the local polytope where by definition, entropic inequalities cannot be violated. Thus we can simplify the optimization and increase its reliability by only optimizing over the nonclassical part of the polytope $\mathrm{Conv}(\{p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}\} \bigcup \{p^{\mathrm{L},k}\}_k)$. We find this region as follows. For $1/2 < \epsilon \leqslant 4/7$, we know from proposition 3 that $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ is nonclassical but does not violate eny of the CHSH inequalities, while it violates $I_{2233} \leqslant 2$ (12). By proposition 1, this is the only Bell inequality that $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ violates for this range of ϵ. Thus the nonclassical part of the polytope $\mathrm{Conv}(\{p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}\} \bigcup \{p^{\mathrm{L},k}\}_k)$ is the convex hull of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ and all the local deterministic points that satisfy $I_{2233} = 2$. These are the 30 local deterministic points of Table I. Hence we only need to optimize over convex combinations of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ with these 30 points and not all 81 local deterministic points, which reduces the size of the optimization (number of variables) and increases the chances of it being effective in detecting entropic violations if there are any.

Performing the optimization as outlined above, we found the maximum value to always be nonpositive for both Shannon case and the Tsallis case with $q = 1.1, 2, 3, 10, 50$. We obtained similar results when taking other values of $\epsilon \leqslant 4/7$ in the distribution $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ and also when considering the inequalities $I^i_{\mathrm{BC},q}$ for $i \in \{1, 2, 3\}$. This suggests that no point in the polytope $\mathrm{Conv}(\{p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}\} \bigcup \{p^{\mathrm{L},k}\}_k)$ violates any of the (Shannon or Tsallis entropic) BC inequalities for $\epsilon \leqslant 4/7$. For $\epsilon > 4/7$, some mixing of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ with $p^{(2,2,3,3)}_{\mathrm{C}}$ gives a distribution that violates $I^4_{\mathrm{BC},q} \leqslant 0 \; \forall q \geqslant 1$ (cf. proposition 4). Note that in the Shannon ($q = 1$) case, the range of values of the mixing parameter v for which a violation can be found becomes arbitrarily small as ϵ approaches $4/7$ from above (see Fig. 3). This limits the effectiveness of numerical tests for q close to 1. For instance, in the Shannon case, our program was not able to detect violations of $I^4_{\mathrm{BC}} \leqslant 0$ for $\epsilon < 4.2/7$ (even though our analytic argument shows that these are present), while it was for $\epsilon \geqslant 4.2/7$. Similarly, for the $q = 2$ Tsallis case, violations of $I^4_{\mathrm{BC},2} \leqslant 0$ could be found for $\epsilon \geqslant 4.00001/7$, but not below. The reason for this difference

[8] As a specific example, consider the distribution $p^{(2,2,3,3)}_{\mathrm{E},\epsilon=0.7,v=0.4}$. Figure 3(a) indicates that this distribution violates the Tsallis entropic inequality $I^4_{\mathrm{BC},q=2} \leqslant 0$ but does not violate the Shannon entropic inequality $I^4_{\mathrm{BC}} \leqslant 0$. However, we can always further mix $p^{(2,2,3,3)}_{\mathrm{E},\epsilon=0.7,v=0.4}$ with the classical distribution $p^{(2,2,3,3)}_{\mathrm{C}}$ to obtain $0.05p^{(2,2,3,3)}_{\mathrm{E},\epsilon=0.7,v=0.4} + 0.95p^{(2,2,3,3)}_{\mathrm{C}} = 0.02p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + 0.98p^{(2,2,3,3)}_{\mathrm{C}}$ which violates the Shannon entropic inequality $I^4_{\mathrm{BC}} \leqslant 0$. This is also in agreement with the results of [20], since when mixing is not considered, we also find examples where it is advantageous to use Tsallis entropy in the (2,2,3,3) scenario.

TABLE I. The vertices of $\Pi_{\text{CHSH}}^{(2,2,3,3)}$ that saturate or violate the I_{2233} inequality (12). All the vertices of the polytope can be obtained from the vertices listed here through local relabellings or exchange of parties.

Number	Vertex
1	$\frac{1}{6}$(1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 2, 0, 1, 1, 0, 0, 0, 2, 0, 1, 1, 2, 0, 0)
2	$\frac{1}{6}$(1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 2, 0, 0, 0, 1, 1, 0, 2, 0, 1, 0, 1, 0, 0, 2, 1, 1, 0)
3	$\frac{1}{6}$(1, 1, 0, 2, 0, 0, 0, 1, 1, 0, 2, 0, 1, 0, 1, 0, 0, 2, 1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0)
4	$\frac{1}{6}$(2, 0, 0, 1, 0, 1, 0, 2, 0, 1, 1, 0, 0, 0, 2, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0)
5	$\frac{1}{5}$(1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 0, 2, 1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0)
6	$\frac{1}{5}$(1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0, 2, 1, 1, 0)
7	$\frac{1}{5}$(1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 2, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0)
8	$\frac{1}{5}$(1, 0, 0, 1, 0, 0, 0, 2, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0)
9	$\frac{1}{5}$(1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 2, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 0)
10	$\frac{1}{5}$(1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 2, 0, 1, 1, 1, 1, 0)
11	$\frac{1}{5}$(1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 2, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 0)
12	$\frac{1}{5}$(1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 2, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 0)
13	$\frac{1}{5}$(1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0,1, 0, 1, 1, 2, 0, 0)
14	$\frac{1}{5}$(1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 2, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0)
15	$\frac{1}{5}$(1, 1, 0, 2, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0)
16	$\frac{1}{5}$(2, 0, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0)
17	$\frac{1}{9}$(2, 1, 0, 2, 0, 1, 0, 2, 1, 1, 2, 0, 1, 0, 2, 0, 1, 2, 2, 0, 1, 0, 2, 1, 1, 2, 0, 1, 0, 2, 0, 1, 2, 2, 1, 0)
18	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1)
19	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0)
20	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
21	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0)
22	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
23	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0)
24	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0)
25	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0)
26	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
27	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
28	(0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0)
29	(0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
30	(0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0)
31	(0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0)
32	(0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0)
33	(0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0)
34	(0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
35	(0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0)
36	(0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0)
37	(0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
38	(0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0)
39	(0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0)
40	(0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0)
41	(0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0)
42	(1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0)
43	(1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
44	(1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
45	(1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0)
46	(1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0)
47	(1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)

is in line with what one might expect by comparing the plots in Fig. 3, where for $\epsilon > 4/7$ the range of values of v for which a violation is possible is larger in the $q = 2$ case. This highlights an advantage of using Tsallis entropy and gives us further confidence that for $1/2 \leqslant \epsilon < 4/7$ there is no violation.

However, because of the form of our objective function, the optimization methods available do not guarantee to find the global maximum. Thus our findings only constitute evidence for the conjectures and are not conclusive. In general, finding global optima for nonlinear, nonconvex/concave

functions is an open question. A potential avenue for proving these conjectures is using DC (difference of convex) programming [34] since our objective function being a linear combination of entropies is a difference of convex functions.

APPENDIX C: PROOFS

For the proofs we need the concept of the local weight of a nonsignalling distribution [15,35].

Definition 2. The *local weight* of a no-signalling distribution $p_{XY|AB}$ is the largest $\alpha \in [0, 1]$ such that we can write

$$p_{XY|AB} = \alpha q^{\mathrm{L}}_{XY|AB} + (1-\alpha) q^{\mathrm{NL}}_{XY|AB},$$

where $q^{\mathrm{L}}_{XY|AB}$ is an arbitrary local distribution and $q^{\mathrm{NL}}_{XY|AB}$ is an arbitrary nonsignalling distribution. We denote the local weight by $l(p_{XY|AB})$.

The local weight of a distribution can be found by linear programming.

Proposition 3. $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ is nonclassical if and only if $\epsilon > 1/2$. Further, for $\epsilon \leqslant 4/7$, $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ satisfies all the CHSH-type inequalities, while for $\epsilon > 4/7$ it violates at least one CHSH-type inequality.

Proof. The distribution $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} = \epsilon p_{\mathrm{NL}} + (1-\epsilon) p^{(2,2,3,3)}_{\mathrm{noise}}$ can be written as follows.

$$p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} = \left[\begin{array}{ccc|ccc} A & B & B & A & B & B \\ B & A & B & B & A & B \\ B & B & A & B & B & A \\ \hline A & B & B & B & A & B \\ B & A & B & B & B & A \\ B & B & A & A & B & B \end{array}\right] \tag{C1}$$

where $A = (2\epsilon+1)/9$ and $B = (1-\epsilon)/9$. We used the LPASSUMPTIONS linear program solver [36] to find the local weight of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$, as a function of ϵ to be

$$l\left(p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}\right) = \begin{cases} 1 & 0 \leqslant \epsilon \leqslant \frac{1}{2} \\ 2(1-\epsilon) & \frac{1}{2} < \epsilon \leqslant 1 \end{cases},$$

which establishes the first part of the claim.

The second part can be confirmed by computing the value of each CHSH-type quantity for the distribution $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ and determining that each has a saturating ϵ of at most $4/7$. ∎

Proposition 4. For $\epsilon \leqslant 4/7$, $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v} = v p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + (1-v) p^{(2,2,3,3)}_{\mathrm{C}}$ does not violate any of the BC entropic inequalities (8) for any $v \in [0, 1]$. However, for all $\epsilon > 4/7$, there exists a $v \in [0, 1]$ such that the entropic inequality $I^4_{\mathrm{BC}} \leqslant 0$ is violated by $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v}$.

Proof. Consider the function $f : (0, 1) \times (0, 1) \to \mathbb{R}$ given by

$$\begin{aligned} f(\epsilon, v) := {} & 3(3-2(1-\epsilon)v)\ln[3-2(1-\epsilon)v] \\ & + 5(1-\epsilon)v\ln[(1-\epsilon)v] - (1+2\epsilon)v\ln[(1+2\epsilon)v] \\ & - (3-(2+\epsilon)v)\ln[3-(2+\epsilon)v] - 3\ln 9, \end{aligned}$$

where we implicitly extend the domain to $[0, 1] \times [0, 1]$ by taking the relevant limit. The Shannon entropic expression $I^4_{\mathrm{BC}}(\epsilon, v)$ evaluated for the distribution $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v} = v p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + (1-v) p^{(2,2,3,3)}_{\mathrm{C}}$ (seen as a function of ϵ and v) is then given as

$$I^4_{\mathrm{BC}} = \frac{1}{3\ln[2]} f(\epsilon, v)$$

Thus all the following arguments for $f(\epsilon, v)$ also hold for I^4_{BC}.

We first use that for $c > 0$ and $a \in \mathbb{R}$ for sufficiently small v, we have

$$\ln[c + av] = \ln[c] + \frac{av}{c} + O(v^2).$$

Using this we can expand $f(\epsilon, v)$ for small v as

$$\begin{aligned} f(\epsilon, v) = {} & (4-7\epsilon)v\ln[v] - (4-7\epsilon)v(1+\ln[3]) \\ & + v(5(1-\epsilon)\ln[1-\epsilon] \\ & - (1+2\epsilon)\ln[1+2\epsilon]) + O(v^2). \end{aligned} \tag{C2}$$

Thus, since $\lim_{v\to 0} v\ln[v] = 0$, we have $\lim_{v\to 0} f[\epsilon, v] = 0$.

We also have

$$\begin{aligned} \frac{\partial}{\partial v} f(\epsilon, v) = {} & (4-7\epsilon)\ln[v] + 5(1-\epsilon)\ln[1-\epsilon] \\ & - (1+2\epsilon)\ln[1+2\epsilon] - 6(1-\epsilon)\ln[3-2(1-\epsilon)v] \\ & + (2+\epsilon)\ln[3-(2+\epsilon)v]. \end{aligned} \tag{C3}$$

Note that $5(1-\epsilon)\ln[1-\epsilon] \leqslant 0$, $-(1+2\epsilon)\ln[1+2\epsilon] \leqslant 0$ $\forall \epsilon \in [0, 1]$. Further, since $3-2(1-\epsilon)v \geqslant 3-(2+\epsilon)v$ and both terms are positive, $6(1-\epsilon) \geqslant (2+\epsilon)$ $\forall \epsilon < 4/7$, and using the fact that ln[] is an increasing function, we have $-6(1-\epsilon)\ln[3-2(1-\epsilon)v] + (2+\epsilon)\ln[3-(2+\epsilon)v] \leqslant 0$ $\forall \epsilon \in [0, 4/7]$, $v \in [0, 1]$. This in turn implies that

$$\frac{\partial}{\partial v} f(\epsilon, v) \leqslant (4-7\epsilon)\ln[v] \qquad \forall\, 0 \leqslant \epsilon \leqslant 4/7, 0 \leqslant v \leqslant 1.$$

Hence we can conclude that for $\epsilon \leqslant 4/7$, $\frac{\partial}{\partial v} f(\epsilon, v) < 0$ for all $v \in [0, 1]$. Thus $f(\epsilon, v)$ is zero at $v = 0$ and, for $\epsilon \leqslant 4/7$, decreases with v, implying that $f(\epsilon, v) \leqslant 0$ $\forall \epsilon \in [0, 4/7]$, $v \in [0, 1]$. Note that $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v} = v p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + (1-v) p^{(2,2,3,3)}_{C}$ does not violate any of the analogous inequalities $I^i_{\mathrm{BC}} \leqslant 0$ for any $i \in \{1, 2, 3\}$, $\epsilon, v \in [0, 1]$. This is because for this distribution, we always have $H(X_0Y_0) = H(X_0Y_1) = H(X_1Y_0)$, $H(X_0) = H(X_1) = H(Y_0) = H(Y_1)$. Thus all three inequalities $I^1_{\mathrm{BC}} \leqslant 0$, $I^2_{\mathrm{BC}} \leqslant 0$ and $I^3_{\mathrm{BC}} \leqslant 0$ reduce to $H(X_1) + H(Y_1) - H(X_1Y_0) - H(X_1Y_1) \leqslant 0$, which is always satisfied since $H(X_1) \leqslant H(X_1Y_0)$ and $H(Y_1) \leqslant H(X_1Y_1)$ by the monotonicity of Shannon entropy.

Further, using the expression for the derivative of $f(\epsilon, v)$ with respect to v in Eq. (C3), we find that for $\epsilon > 4/7$, $\lim_{v\to 0} \frac{\partial}{\partial v} f(\epsilon, v) = \infty$. Thus, since $f(\epsilon, v) = 0$ for $v = 0$, sufficiently close to $v = 0$ there exists a v such that $f(\epsilon, v) > 0$. This proves the claim. ∎

Proposition 5. For $\epsilon \leqslant 4/7$, $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v} = v p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + (1-v) p^{(2,2,3,3)}_{\mathrm{C}}$ does not violate any of the Tsallis BC inequalities (17) for any $v \in [0, 1]$ and $q > 1$. However, for $\epsilon > 4/7$ and every $q > 1$, there always exists a $v \in [0, 1]$ such that the entropic inequality $I^4_{\mathrm{BC},q} \leqslant 0$ is violated by $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v}$.

Proof. The Tsallis entropic expression $I^4_{\mathrm{BC},q}(\epsilon, v)$ evaluated for the distribution $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v} = v p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} +$

$(1-v)p_{\mathrm{C}}^{(2,2,3,3)}$ (seen as a function of q, ϵ and v) is given as

$$\begin{aligned} I^4_{\mathrm{BC},q} &= \frac{1}{q-1}\Bigg[9\left(\frac{3-2(1-\epsilon)v}{9}\right)^q + 15\left(\frac{(1-\epsilon)v}{9}\right)^q \\ &\quad - \frac{6}{3^q} - 3\left(\frac{3-(2+\epsilon)v}{9}\right)^q - 3\left(\frac{(1+2\epsilon)v}{9}\right)^q\Bigg] \\ &=: \frac{g(q,\epsilon,v)}{q-1} \end{aligned}$$

For $q > 1$, the following arguments for $g(q,\epsilon,v)$ also hold for $I^4_{\mathrm{BC},q}$. Note that

$$\begin{aligned} \frac{\partial}{\partial v}g(q,\epsilon,v) = \frac{q}{3^{2q-1}}[&-6(1-\epsilon)(3-2(1-\epsilon)v)^{q-1} \\ &+5(1-\epsilon)((1-\epsilon)v)^{q-1} \\ &+(2+\epsilon)(3-(2+\epsilon)v)^{q-1} \\ &-(1+2\epsilon)((1+2\epsilon)v)^{q-1}]. \end{aligned} \tag{C4}$$

Then, since

$$6(1-\epsilon)(3-2(1-\epsilon)v\ )^{q-1} \geqslant (2+\epsilon)(3-(2+\epsilon)v)^{q-1}$$

and

$$(1+2\epsilon)((1+2\epsilon)v)^{q-1} \geqslant 5(1-\epsilon)((1-\epsilon)v)^{q-1}$$

hold for all $\epsilon \leqslant 4/7$, $v \in [0,1]$ and $q > 1$, we have

$$\frac{\partial}{\partial v}g(q,\epsilon,v) \leqslant 0 \qquad \forall \epsilon \leqslant 4/7, \quad v \in [0,1], \quad q > 1.$$

Since $g(q,\epsilon,v=0) = 0$, this implies that $g(q,\epsilon,v) \leqslant 0$ $\forall \epsilon \leqslant 4/7$, $v \in [0,1]$, $q > 1$. Hence, for $\epsilon \leqslant 4/7$, we cannot violate $I^4_{\mathrm{BC},q} \leqslant 0$ for any $v \in [0,1]$, $q > 1$.

Analogously to the Shannon case, $p^{(2,2,3,3)}_{\mathrm{E},\epsilon,v} = v p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + (1-v)p_{\mathrm{C}}^{(2,2,3,3)}$ also does not violate any of the inequalities $I^1_{\mathrm{BC},q} \leqslant 0$, $I^2_{\mathrm{BC},q} \leqslant 0$ or $I^3_{\mathrm{BC},q} \leqslant 0$ for any $\epsilon, v \in [0,1]$ and $q > 1$ by the same argument as in proposition 4.

Further, Eq. (C4) implies $\lim_{v\to 0}\frac{\partial}{\partial v}g(q,\epsilon,v) = \frac{q}{3^q}(7\epsilon - 4)$ which is always positive for $\epsilon > 4/7$. Since $g(q,\epsilon,v) = 0$ for $v = 0$, this allows us to conclude that for $\epsilon > 4/7$, there exists a v sufficiently close to $v = 0$ such that $g(q,\epsilon,v) > 0$. This establishes the claim. ■

Proposition 6. The distribution $p^{\mathrm{CG},i}_{\epsilon}$ is classical for all i if and only if $\epsilon \leqslant 4/7$.

Proof. For the "if" part of the proof, we calculated all the 255 coarse-grainings of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon=4/7}$ and used a linear programming algorithm to find that all of these are local (their local weight equals 1). Since decreasing ϵ in $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ cannot increase the violation of any (probability space) Bell inequality and neither can coarse-grainings, it follows that if $\epsilon \leqslant 4/7$, all coarse-grainings of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ are classical.

For the "only if" part we need to show that if all coarse-grainings of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ are classical then $\epsilon \leqslant 4/7$ or equivalently, if $\epsilon > 4/7$, there exists at least one coarse-graining of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ that is nonclassical. Consider the coarse-graining that involves combining the second output with the first for all 4 input choices. For $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ as in Eq. (C1), this coarse-graining gives

$$p^{(2,2,3,3)}_{\mathrm{CG},\epsilon} = \left[\begin{array}{ccc|ccc} 2(A+B) & 0 & 2B & 2(A+B) & 0 & 2B \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 2B & 0 & A & 2B & 0 & A \\ \hline 2(A+B) & 0 & 2B & 3B+A & 0 & A+B \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 2B & 0 & A & A+B & 0 & B \end{array}\right], \tag{C5}$$

where $A = (2\epsilon+1)/9$ and $B = (1-\epsilon)/9$. The I_{2233} value or left hand side of Eq. (12) for this distribution is $9A - 3B$. For this to be classical, we require that $9A - 3B \leqslant 2$ which gives $\epsilon \leqslant 4/7$. Again using the LPASSUMPTIONS linear program solver [36] we found the local weight of this distribution as a function of ϵ, which gives the following.

$$l\left(p^{(2,2,3,3)}_{\mathrm{CG},\epsilon}\right) = \begin{cases} 1, & 0 \leqslant \epsilon \leqslant \frac{4}{7}. \\ \frac{1}{9}(17-14\epsilon), & \frac{4}{7} < \epsilon \leqslant 1. \end{cases}$$

In other words, if $\epsilon > 4/7$, then the coarse-graining $p^{(2,2,3,3)}_{\mathrm{CG},\epsilon}$ of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ violates the I_{2233} Inequality (12) and is hence nonclassical. This concludes the proof. ■

We prove the following two propositions together as they only differ in one step.

Proposition 7. Let $\epsilon \leqslant 4/7$. If conjecture 1 holds, then every distribution in Π_ϵ is Shannon entropically classical.

Proposition 8. Let $\epsilon \leqslant 4/7$. If conjecture 2 holds, then every distribution in Π_ϵ satisfies the Tsallis entropic BC inequalities (17) $\forall q > 1$.

Proof. If conjectures 1 and 2 hold, then for any $\epsilon \leqslant 4/7$, $I^i_{\mathrm{BC},q} \leqslant 0$ holds $\forall q \geqslant 1$, $\forall i \in \{1,2,3,4\}$ and for all distributions in the polytope $\mathrm{Conv}(\{p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}\}\bigcup\{p^{\mathrm{L},k}\}_k)$.[9] We want to show that this implies the same for the larger polytope that comprises the convex hull of not just $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ and local deterministic distributions $\{p^{\mathrm{L},k}\}_k$, but also all the local relabellings of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$, i.e., the polytope $\Pi_\epsilon = \mathrm{Conv}(\{p^{\mathrm{R},j}_\epsilon\}_j\bigcup\{p^{\mathrm{L},k}\}_k)$. While we considered $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ in propositions 4 and 5 and conjectures 1 and 2, due to symmetries, these also apply to every relabelling of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$, i.e., $I^i_{\mathrm{BC},q} \leqslant 0$ ($\forall i \in \{1,2,3,4\}$) throughout every polytope in the set $\{\mathrm{Conv}(\{p^{\mathrm{R},j}_\epsilon\}\bigcup\{p^{\mathrm{L},k}\}_k)\}_j$ (where j runs over all the local relabellings of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$).[10] This is because for every input-output relabelling of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$, we can correspondingly relabel the inequality expression $I^4_{\mathrm{BC},q}$ (for $q \geqslant 1$) and the same arguments as propositions 4 and 5 again hold, and similarly conjectures 1 and 2 also follow for this case.[11] From this argument, it follows that if conjec-

[9] Note that $q = 1$ covers the Shannon case.

[10] A note on notation: $\{p^{\mathrm{R},j}_\epsilon\}$ is a set comprising a single element, while $\{p^{\mathrm{R},j}_\epsilon\}_j$ is a set whose elements are the distributions for every j.

[11] Note that output relabellings don't change the entropic expression but input relabellings (4 in number) can give either one of $I^i_{\mathrm{BC},q} \leqslant 0$ for $i \in \{1,2,3,4\}$. Thus for an output relabelling of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$, one can continue using $I^4_{\mathrm{BC},q}$ in propositions 4 and 5 and the following conjectures while for input relabellings, one simply needs to relabel the inequalities accordingly and run the same arguments.

tures 1 and 2 hold, then $I^i_{\mathrm{BC},q} \leqslant 0\ \forall q \geqslant 1,\ \forall i \in \{1,2,3,4\}$ everywhere in the union of the polytopes, i.e., everywhere in $\bigcup_j \mathrm{Conv}(\{p^{\mathrm{R},j}_\epsilon\} \bigcup \{p^{\mathrm{L},k}\}_k)$.

To conclude the proof it remains to show that $\bigcup_j \mathrm{Conv}(\{p^{\mathrm{R},j}_\epsilon\} \bigcup \{p^{\mathrm{L},k}\}_k) = \Pi_\epsilon \quad \forall \epsilon \leqslant 4/7$. This is established below. Then, proposition 8 automatically follows while proposition 7 follows from this and Lemma 1. ■

Proposition 9. $p^j_{\mathrm{mix},\epsilon} := \frac{1}{2} p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + \frac{1}{2} p^{\mathrm{R},j}_\epsilon$ is local $\forall j \neq 1$ if and only if $\epsilon \leqslant 4/7$.

Proof. For the "if" part of the proof, we used a linear program to confirm that $p^j_{\mathrm{mix},\epsilon=4/7}$ is local[12] $\forall j \neq 1$. Since reducing ϵ in $p^j_{\mathrm{mix},\epsilon}$ cannot decrease the local weight, this also holds for $\epsilon < 4/7$. The "only if" part of the proof is equivalent to showing that $\forall \epsilon > 4/7,\ \exists j$ such that $p^j_{\mathrm{mix},\epsilon}$ is nonclassical. Consider the particular local relabelling that corresponds to Alice swapping the outputs "1" and "2" only when her input is $A = 1$. Let the distribution obtained by applying this relabelling to $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ be p^{R}_ϵ. Then $p_{\mathrm{mix},\epsilon} = \frac{1}{2} p^{(2,2,3,3)}_{\mathrm{iso},\epsilon} + \frac{1}{2} p^{\mathrm{R}}_\epsilon$. More explicitly,

$$p^{\mathrm{R}}_\epsilon = \left[\begin{array}{ccc|ccc} A & B & B & A & B & B \\ B & A & B & B & A & B \\ B & B & A & B & B & A \\ \hline A & B & B & B & A & B \\ B & B & A & A & B & B \\ B & A & B & B & B & A \end{array}\right] \quad \text{and}$$

$$p_{\mathrm{mix},\epsilon} = \left[\begin{array}{ccc|ccc} A & B & B & A & B & B \\ B & A & B & B & A & B \\ B & B & A & B & B & A \\ \hline A & B & B & B & A & B \\ B & * & * & * & B & * \\ B & * & * & * & B & * \end{array}\right], \tag{C6}$$

where $* = \frac{A+B}{2}$, $A = (2\epsilon + 1)/9$, and $B = (1 - \epsilon)/9$. Now consider the Bell inequality $\mathrm{Tr}(M^T P) \geqslant 1$ where P is an arbitrary distribution in the (2,2,3,3) scenario and M is the following matrix:

$$M := \left[\begin{array}{ccc|ccc} 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{array}\right] \tag{C7}$$

Then, the condition for $p_{\mathrm{mix},\epsilon}$ to be nonclassical with respect to this Bell inequality i.e., $\mathrm{Tr}(M^T p_{\mathrm{mix},\epsilon}) < 1$ gives us $\epsilon > 4/7$. Since p^{R}_ϵ is a local relabelling of $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$, there exists a j such that $p^{\mathrm{R}}_\epsilon = p^{\mathrm{R},j}_\epsilon$ and hence $p_{\mathrm{mix},\epsilon} = p^j_{\mathrm{mix},\epsilon}$. Thus we have shown that whenever $\epsilon > 4/7$, $\exists j$ such that $p^j_{\mathrm{mix},\epsilon}$ is nonclassical and hence $p^j_{\mathrm{mix},\epsilon}$ is local for all j implies that $\epsilon \leqslant 4/7$ which concludes the proof. ■

By symmetry, there is an analog of proposition 9 with $p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ replaced by $p^{\mathrm{R},i}_\epsilon$ for any i, so we have the following corollary.

Corollary 6. $\tilde{p}^{i,j}_{\mathrm{mix},\epsilon} := \frac{1}{2} p^{\mathrm{R},i}_\epsilon + \frac{1}{2} p^{\mathrm{R},j}_\epsilon$ is local $\forall j \neq i$ if and only if $\epsilon \leqslant 4/7$.

Theorem 7 (Bemporad *et al.* 2001 [37]). Let $\mathcal{P}$ and $\mathcal{Q}$ be polytopes with vertex sets V and W, respectively, i.e., $\mathcal{P} = \mathrm{Conv}(V)$ and $\mathcal{Q} = \mathrm{Conv}(W)$. Then $\mathcal{P} \bigcup \mathcal{Q}$ is convex if and only if the line segment $[v, w]$ is contained in $\mathcal{P} \bigcup \mathcal{Q}\ \forall v \in V$ and $w \in W$.

Let $\mathcal{P}_j = \mathrm{Conv}(\{p^{\mathrm{R},j}_\epsilon\} \bigcup \{p^{\mathrm{L},k}\}_k)$ and $\mathbb{P}$ be the set of polytopes $\mathbb{P} := \{\mathcal{P}_j\}_j$. We use the above theorem to prove the final result that establishes propositions 7 and 8.

Lemma 8. Let V_j be the vertex set of the polytope $\mathcal{P}_j \in \mathbb{P}$ and $V := \bigcup_j V_j$. Then, $\bigcup_{\mathcal{P}_j \in \mathbb{P}} \mathcal{P}_j = \mathrm{Conv}(\bigcup_i V_j) = \mathrm{Conv}(V) = \Pi_\epsilon\ \forall \epsilon \leqslant 4/7$.

Proof. By corollary 6, for $i \neq j$ we have that $\frac{1}{2}(p^{\mathrm{R},i}_\epsilon + p^{\mathrm{R},j}_\epsilon) \in \mathcal{L}^{(2,2,3,3)} = \mathcal{P}_i \bigcap \mathcal{P}_j\ \forall \epsilon \leqslant 4/7$. This implies that $\alpha p^{\mathrm{R},i}_\epsilon + (1-\alpha) p^{\mathrm{R},j}_\epsilon \in \mathcal{P}_i \bigcup \mathcal{P}_j \quad \forall \epsilon \leqslant 4/7, \quad \alpha \in [0,1]$, i.e., the line segment $[p^{\mathrm{R},i}_\epsilon, p^{\mathrm{R},j}_\epsilon]$ is completely contained in the union of the corresponding polytopes $\mathcal{P}_i \bigcup \mathcal{P}_j$. Note that all other line segments $[v_i, v_j]$ with $v_i \in V_i$ and $v_j \in V_j$ are contained in $\mathcal{P}_i \bigcup \mathcal{P}_j$ by construction since at least one of v_i or v_j would be a local-deterministic vertex. Therefore, by theorem 7, $\mathcal{P}_i \bigcup \mathcal{P}_j$ is convex $\forall i, j$ and $\epsilon \leqslant 4/7$. We can then apply proposition 9 and theorem 7 to the convex polytopes $\mathcal{P}_i \bigcup \mathcal{P}_j$ and $\mathcal{P}_k$ and show that $\mathcal{P}_i \bigcup \mathcal{P}_j \bigcup \mathcal{P}_k$ is convex $\forall i, j, k$ and $\epsilon \leqslant 4/7$. Proceeding in this way, we conclude that $\bigcup_{\mathcal{P}_i \in \mathbb{P}} \mathcal{P}_i$ is convex $\forall \epsilon \leqslant 4/7$, and hence $\bigcup_{\mathcal{P}_j \in \mathbb{P}} \mathcal{P}_j = \mathrm{Conv}(\bigcup_i V_j) = \mathrm{Conv}(V) = \Pi_\epsilon$. ■

[12] $j = 1$ is excluded since $p^1_{\mathrm{mix},\epsilon} = p^{(2,2,3,3)}_{\mathrm{iso},\epsilon}$ which is nonclassical for $\epsilon > 1/2$.

[1] J. Barrett, Information processing in generalized probabilistic theories, Phys. Rev. A **75**, 032304 (2007).

[2] A. K. Ekert, Quantum Cryptography Based on Bell's Theorem, Phys. Rev. Lett. **67**, 661 (1991).

[3] D. Mayers and A. Yao, Quantum cryptography with imperfect apparatus, in *Proceedings of the 39th Annual Symposium on Foundations of Computer Science (FOCS-98)* (IEEE Computer Society, Los Alamitos, CA, USA, 1998), pp. 503–509.

[4] J. Barrett, L. Hardy, and A. Kent, No Signaling and Quantum Key Distribution, Phys. Rev. Lett. **95**, 010503 (2005).

[5] R. Colbeck, Quantum and relativistic protocols for secure multi-party computation, Ph.D. dissertation, University of Cambridge, 2007, also available as arXiv:0911.3814.

[6] R. Colbeck and A. Kent, Private randomness expansion with untrusted devices, J. Phys. A **44**, 095305 (2011).

[7] S. Pironio *et al.*, Random numbers certified by Bell's theorem, Nature (London) **464**, 1021 (2010).

[8] J. S. Bell, On the Einstein-Podolsky-Rosen paradox, Phys. Phys. Fiz. **1**, 195 (1964).

[9] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[10] D. Kaszlikowsli, L. C. Kwek, J. L. Chen, M. Zukowski, and C. H. Oh, Clauser-Horne inequality for three-state systems, Phys. Rev. A **65**, 032118 (2002).

[11] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Bell Inequalities for Arbitrarily High-Dimensional Systems, Phys. Rev. Lett. **88**, 040404 (2002).

[12] I. Pitowski, *Quantum Probability – Quantum Logic*, Lecture Notes in Physics Vol. 321 (Springer-Verlag, Berlin, Heidelberg, 1989).

[13] L. Masanes, Tight Bell inequality for d-outcome measurements correlations, Quantum Inf. Comput. **3**, 345 (2002).

[14] J.-D. Bancal, N. Gisin, and S. Pironio, Looking for symmetric Bell inequalities, J. Phys. A **43**, 385303 (2010).

[15] T. Cope and R. Colbeck, Bell inequalities from no-signaling distributions, Phys. Rev. A **100**, 022114 (2019).

[16] S. L. Braunstein and C. M. Caves, Information-Theoretic Bell Inequalities, Phys. Rev. Lett. **61**, 662 (1988).

[17] M. Weilenmann and R. Colbeck, Inability of the entropy vector method to certify nonclassicality in linelike causal structures, Phys. Rev. A **94**, 042112 (2016).

[18] V. Vilasini and R. Colbeck, Analyzing causal structures using Tsallis entropies, Phys. Rev. A **100**, 062108 (2019).

[19] R. Chaves, Entropic inequalities as a necessary and sufficient condition to noncontextuality and locality, Phys. Rev. A **87**, 022102 (2013).

[20] M. Wajs, P. Kurzyński, and D. Kaszlikowski, Information-theoretic Bell inequalities based on Tsallis entropy, Phys. Rev. A **91**, 012114 (2015).

[21] B. Tsirelson, Some results and problems on quantum Bell-type inequalities, Hadronic J. Suppl. **8**, 329 (1993).

[22] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, Found. Phys. **24**, 379 (1994).

[23] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, J. Stat. Phys. **52**, 479 (1988).

[24] Z. Daróczy, Generalized information functions, Information and Control **16**, 36 (1970).

[25] S. Furuichi, Information theoretical properties of Tsallis entropies, J. Math. Phys. **47**, 023302 (2004).

[26] M. Weilenmann and R. Colbeck, Analysing causal structures with entropy, Proc. R. Soc. London A **473**, 20170483 (2017).

[27] T. Fritz and R. Chaves, Entropic inequalities and marginal problems, IEEE Trans. Inf. Theory **59**, 803 (2013).

[28] D. Collins and N. Gisin, A relevant two qubit Bell inequality inequivalent to the CHSH inequality, J. Phys. A **37**, 1775 (2004).

[29] T. Christof and A. Loebel, POlyhedron Representation Transformation Algorithm (PORTA) (1997), http://comopt.ifi.uni-heidelberg.de/software/PORTA/index.html.

[30] R. Chaves and T. Fritz, Entropic approach to local realism and noncontextuality, Phys. Rev. A **85**, 032113 (2012).

[31] D. Avis, H. Imai, I. Tsuyoshi, and Y. Sasaki, Deriving tight Bell inequalities for 2 parties with many 2-valued observables from facets of cut polytopes, arXiv:quant-ph/0404014.

[32] E. Wolfe, D. Schmid, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Quantifying Bell: The resource theory of nonclassicality of common-cause boxes, Quantum **4**, 280 (2020).

[33] M. Weilenmann and R. Colbeck, Non-Shannon inequalities in the entropy vector approach to causal structures, Quantum **2**, 57 (2018).

[34] R. Horst and N. V. Thoai, DC programming: Overview, J. Optim. Theory Appl. **103**, 1 (1999).

[35] M. Zukowski, D. Kaszlikowski, A. Baturo, and J.-Å. Larsson, Strengthening the Bell theorem: Conditions to falsify local realism in an experiment, arXiv:quant-ph/9910058.

[36] R. Colbeck and V. Vilasini, LPASSUMPTIONS (MATHEMATICA package), https://github.com/rogercolbeck/LPAssumptions.

[37] A. Bemporad, K. Fukuda, and F. D. Torrisi, Convexity recognition of the union of polyhedra, Comput. Geometry **18**, 141 (2001).