
Title	Uniqueness of all fundamental noncontextuality inequalities
Author(s)	Kishor Bharti, Atul Singh Arora, Leong Chuan Kwek and Jérémie Roland

Copyright © 2021 The Authors & American Physical Society

This is the published version of the following article:

Kishor Bharti, Atul Singh Arora, Kwek, L. C., & Roland, J. (2020). Uniqueness of all fundamental noncontextuality inequalities. *Physical Review Research*, 2, Article 033010. <https://doi.org/10.1103/physrevresearch.2.033010>

Uniqueness of all fundamental noncontextuality inequalities

Kishor Bharti,^{1,*} Atul Singh Arora^{2,*} Leong Chuan Kwek,^{1,3,4} and Jérémie Roland²

¹*Centre for Quantum Technologies, National University of Singapore 117543, Singapore*

²*Centre for Quantum Information and Communication, École polytechnique de Bruxelles, CP 165, Université libre de Bruxelles, 1050 Bruxelles, Belgium*

³*MajuLab, CNRS-UNS-NUS-NTU International Joint Research Unit, Singapore UMI 3654, Singapore*

⁴*National Institute of Education, Nanyang Technological University, Singapore 637616, Singapore*



(Received 22 December 2019; revised 24 May 2020; accepted 29 May 2020; published 2 July 2020)

Contextuality is one way of capturing the nonclassicality of quantum theory. The contextual nature of a theory is often witnessed via the violation of noncontextuality inequalities—certain linear inequalities involving probabilities of measurement events. Using the exclusivity graph approach (one of the two main graph theoretic approaches for studying contextuality), it was shown [Cabello *et al.* *Phys. Rev. A* **88**, 032104 (2013); Chudnovsky *et al.* *Ann. Math.* **164**, 51 (2006)] that a necessary and sufficient condition for witnessing contextuality is the presence of an odd number of events (greater than three) which are either cyclically or anticyclically exclusive. Thus, the noncontextuality inequalities the underlying exclusivity structure of which is as stated, either cyclic or anticyclic, are fundamental to quantum theory. We show that there is a unique noncontextuality inequality for each nontrivial cycle and anticycle. In addition to the foundational interest, we expect this to aid the understanding of contextuality as a resource to quantum computing and its applications to local self-testing.

DOI: [10.1103/PhysRevResearch.2.033010](https://doi.org/10.1103/PhysRevResearch.2.033010)

I. INTRODUCTION

A. Motivation

In an attempt to conceptually understand the departure of the predictions of quantum mechanics (QM) from that of classical physics, the notion of contextuality was introduced. It is one of the most general ways of capturing this divergence [1,2]; the celebrated Bell nonlocality can be viewed as a special case of contextuality where the context is provided via spacelike separation of the parties involved [3,4]. More generally, a *context* is defined by a set of compatible observables, viz., jointly measurable observables.

Investigations into these fundamental questions have also reaped practical benefits. Bell nonlocality has found many applications in quantum key distribution [5], randomness certification [6], self-testing [7–9], and distributed computing [10], to name a few [11]. Recently, contextuality has also been applied more directly to quantum key distribution [12,13] and variants of randomness certification [14] and self-testing [15]. Further, it has been uncovered to be the resource powering the measurement based model and a class of fault tolerant models of quantum computation [16,17], among others [1,16–23].

B. Bell nonlocality or contextuality

The idea at the heart of this discussion can be traced back to Einstein, who expressed his discomfort with the probabilistic

nature of quantum mechanics by providing a striking argument against it [24] using a notion of realism (element of physical reality) for two spatially separated experiments. He believed that there must exist local hidden variables which, once supplied, make QM deterministic. Such *completions* are referred to as local hidden variable models. Bell constructed a linear inequality which is violated by QM and yet it can never be violated by any such completion [3], falsifying Einstein's belief [25]. It may be said that the *Bell inequality witnesses* the nonlocality of (any such completion of) QM.

General discussions on this topic are facilitated by correspondingly considering general probabilistic assignments to the various observable events. The set of probabilistic assignments, which admit a local hidden variable description, form a convex polytope (a bounded set the boundaries of which are defined by hyperplanes). The facet-defining Bell inequalities are the characterizing hyperplanes of the aforesaid polytope. Once formalized, this becomes a general framework for studying Bell inequalities (which can and has been refined to facilitate computations). This can, however, be further generalized if an underlying principle which is correspondingly more general than that of local realism is used. In the Bell scenario, there was a clear role of spatial separation and therefore there were at least two parties involved. It turns out that one can study nonclassicality even for a single indivisible quantum system. To this end, one uses *noncontextual completions* of probabilistic assignments where the phrase *noncontextual* emphasizes that there is a precise value assigned to each observable by the completion. This is because it is possible to define completions where the value assigned depends on the context (i.e., the set of compatible observables it is measured with), and such completions can explain the predictions of quantum mechanics. Consequently quantum

*These authors contributed equally to this work.

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

mechanics is sometimes called contextual [26]. This schism between noncontextual completions and quantum mechanics is used as the underlying principle to construct frameworks to study *noncontextuality (NC) inequalities*. The Klyachko-Can-Binicioğlu-Shumovsky (KCBS) inequality may be considered to be the simplest NC inequality [analogous to the Bell–Clauser-Horne-Shimony-Holt (CHSH) inequality in the spatial separation setting]. There are two principal graph theoretic frameworks for studying contextuality: the compatibility hypergraph approach [29–32] and the exclusivity graph approach [4]. The former uses hyperedges to encode the compatibility relations between the observables. The latter uses a slightly different physical approach and focuses on the exclusivity of measurement events. The exclusivity relations between these events are encoded using edges [33]. We show that for *any* scenario (characterized by *any* given exclusivity graph) which exhibits contextuality there is a *unique* NC inequality, corresponding to each induced subgraph of the graph possessing a certain property (which in turn are separately known to exist). Further, these NC inequalities may be seen as obvious generalizations of the KCBS inequality.

C. Fundamental noncontextuality inequalities

The key appeal of the exclusivity graph approach stems from a powerful result in graph theory—the strong perfect graph theorem [30,34]. Consider a scenario encoded by a certain exclusivity graph. The contextuality in the scenario can be witnessed by some NC inequality and appropriately constructed states or measurements if and only if the exclusivity graph associated with it contains, as an induced subgraph, an *odd* cyclic graph and/or an odd anticyclic graph of length greater than 3. The said obvious generalization of the KCBS inequality turns out to be the simplest inequality which has an underlying odd cycle as its exclusivity graph. These inequalities, together with their analog for the anticycle, may in hindsight be termed *fundamental NC inequalities* [30]. In this paper, we show that each odd cyclic graph and anticyclic graph corresponds to a unique fundamental NC inequality, justifying its name. Given any exclusivity graph which can exhibit contextuality, for each cycle and anticycle (odd) we can directly deduce that there is a unique inequality corresponding to it. There may, however, be additional inequalities corresponding to other induced subgraphs. We demonstrate this by characterizing the simplest Bell exclusivity scenario (see Sec. VI). In fact, the “other NC inequalities” in this case turn out to be the familiar CHSH-Bell inequality and a *heptagonal Bell inequality*—a Bell inequality involving seven events (see Sec. VI).

D. Relation to prior work

In the compatibility hypergraph approach, scenarios captured by odd n -cycle graphs are characterized by $2^n - 1$ nontrivial NC inequalities, which include the generalized KCBS inequality [35]. We clarify why in the exclusivity graph approach, for odd n -cycle graphs, we obtain an exponential simplification—a unique NC inequality (see Sec. VII). The relevance of anticycles is not *a priori* clear in the compatibility hypergraph approach and therefore, to the best of our knowledge, they have not been studied. However, in the exclusivity

graph approach, an easy characterization of anticycles allows us to make a much more general statement about all scenarios (due to the strong perfect graph theorem, as was noted).

The study of the simplest Bell scenario using the exclusivity graph approach was carried out in [36] and the fundamental NC inequality was shown to be a Bell inequality involving only five events, termed a *pentagonal Bell inequality* (while the CHSH-Bell inequality involves eight events). However, the one involving seven events was missing.

II. PRELIMINARIES

We summarize the exclusivity graph approach here, following the work of Amaral and Cunha [32], deferring a more complete discussion to the Appendix B. An outcome, a , and its associated measurement, M , are together called a *measurement event* (or events for brevity) and denoted by $(a|M)$. Two events are equivalent if their probability of occurrence is the same for all preparations. Let $p_j(k)$ be the probability of getting an outcome k given that a measurement j was performed. Two events, e_i and e_j , are *exclusive* if there exists a measurement M such that e_i and e_j correspond to different outcomes of M (see Definition 24 in Appendix B). With a family of events $\{e_1, e_2, \dots, e_n\}$ we associate the *exclusivity graph* $G := (V, E)$ where V is the set of vertices and E is that of edges, the vertices of which are the events, and there is an edge between the vertices if and only if the events are exclusive (see Definition 25 in Appendix B). The probabilities assigned to these events are formally given by a *behaviour* which for G is defined to be a map $p : V \rightarrow [0, 1]$ that assigns to each vertex i a probability $p(i)$ such that $p(i) + p(j) \leq 1$ for all vertices that share an edge. The map p can also be seen as a vector in $\mathbb{R}^{|V|}$ (see Definition 27 in Appendix B). Behaviors which admit a noncontextual completion, i.e., there exists a noncontextual hidden variable assignment such that if the hidden variable is traced out we recover the given behavior, are defined to be *noncontextual behaviors* (see Definition 28 in Appendix B). The set of such behaviors is denoted by $B_{\text{NC}}(G)$. We can similarly define the set of *quantum behaviors*, $B_{\text{Q}}(G)$, to be that which can be obtained by at least one quantum state and corresponding observables (see Definition 29 in Appendix B). The set of E-principle behaviors, $B_{\text{E}}(G)$, is one where the behaviors respect the exclusivity principle (also referred to as the E principle), i.e., exclusive events must have their probability sum to at most one (see Definition 31 in Appendix B). The central claim of this formalism is that $B_{\text{NC}}(G) \subseteq B_{\text{Q}}(G) \subseteq B_{\text{E}}(G)$ (see Definition 31 in Appendix B). This is a corollary of a powerful identification of each of the sets with geometrical objects studied by Lovász which we describe later. We can now define more precisely a *facet-defining NC inequality* as being a nontrivial facet of $B_{\text{NC}}(G)$ where the direction of the inequality is chosen to satisfy containment in $B_{\text{NC}}(G)$ (see Definition 32 in Appendix B). An n -cycle graph is an n vertex graph where every i th vertex is connected to the $(i + 1)$ th vertex (the addition is modulo n). We define the *fundamental cyclic noncontextuality (FCNC) inequality*, corresponding to the n -cycle graph for n odd, to be

$$K_n := \sum_{i=1}^n p_i \leq \frac{n-1}{2}. \quad (1)$$

We analogously define the *fundamental anticyclic noncontextuality inequality* for the complement of the odd n -cycle graph as

$$K_{\bar{n}} := \sum_{i=1}^n p_i \leq 2. \quad (2)$$

We recover the original KCBS inequality [37] in the special case of $n = 5$, both for the cyclic as well as the anticyclic case.

Relevant graph theoretic notions

We give a primer on the relevant graph theoretic notions which facilitates the statement and proof of our results.

Definition 1. The graph $G = (V, E)$ is defined by the set of vertices and the set of edges.

Definition 2. The orthonormal representation with respect to a graph G is defined as follows. For all $i \rightarrow |v_i\rangle$ in $\mathbb{R}^d$, $\langle v_i | v_j \rangle = 0$ whenever $(i, j) \notin E$.

Definition 3. For a vector $|v_i\rangle$ in an orthonormal representation, the cost is defined as

$$c_i = |\langle \psi | v_i \rangle|^2$$

where $|\psi\rangle \doteq (1, 0, \dots, 0)$ is a vector in $\mathbb{R}^d$.

Definition 4. The theta body corresponding to a graph G is defined to be

$$\text{TH}(G) = \{p \in \mathbb{R}^{|V|} | p_{(i)} = c_i\}$$

where c_i is the cost (see Definition 3) corresponding to $\bar{G} := (V, \bar{E})$ where $\bar{E}$ is the complement [38] of E .

Definition 5. A stable set or independent set is a subset of vertices $K \subseteq V$ such that for all $i, j \in K$ there is no edge between i and j , viz., $(i, j) \notin E$.

Definition 6. The independence number of a graph G is defined to be the cardinality of the largest independent set of G .

Definition 7. A clique is a subset of vertices $K \subseteq V$ such that for all $i, j \in K$ there is an edge between i and j , viz., $(i, j) \in E$.

Definition 8. The incidence vector of a set is defined to be a vector $\bar{p}$ (of size $|V|$) for $K \subseteq V$ such that

$$p_{(i)} = \begin{cases} 1 & \text{if } i \in K \\ 0 & \text{else.} \end{cases}$$

Example 1. Consider the five-cycle graph $V = \{1, 2, 3, 4, 5\}$, $E = \{(1, 2), (2, 3), (3, 4), (4, 5), (5, 1)\}$. $K = \{1, 3\}$ is an example of a stable set. $K' = \{1, 2\}$ is an example of a clique. The incident vector corresponding to K is $p = (1, 0, 1, 0, 0)^T$.

Definition 9. $\text{STAB}(G)$ (not to be confused with the stable set) is defined as the convex hull of the vectors $\bar{p}_{(k)}$ for all stable sets k where $\bar{p}_{(k)}$ is the incidence vector of the set k . (Note: if k were an index, $\bar{p}_{(k)}$ would refer to the k th component of the vector $\bar{p}$; here k is a set).

Definition 10. $\text{QSTAB}(G)$ is the set of vectors $x \in \mathbb{R}^{|V|}$ such that $x_i \geq 0$, $\sum_{i \in Q} x_i \leq 1$ for every clique Q .

Lemma 1 [39]. $\text{STAB}(G)$ is the convex hull of the integer solutions to the equations $x_i \geq 0$, $(x_i + x_j) \leq 1$ for every $(i, j) \in E$, where $G = (V, E)$.

Remark 1. Every set of indices which is an edge is also a clique (the other way is not necessary, obviously). This means

that the inequalities listed in Definition 10 (viz., $\sum_{i \in Q} x_i \leq 1$ for every clique Q) contain the inequalities listed in Lemma 1 [viz., $x_i + x_j \leq 1$ for every $(i, j) \in E$].

Lemma 2 [40]. $\text{STAB}(G) \subseteq \text{TH}(G) \subseteq \text{QSTAB}(G)$.

III. UNIQUENESS OF FUNDAMENTAL CYCLIC NONCONTEXTUALITY INEQUALITIES

We begin by stating our first important result.

Theorem 1. Consider an odd n -cycle exclusivity graph. The associated FCNC inequality is a unique facet-defining NC inequality.

We will need the following powerful result connecting the behaviors to geometrically well-studied objects.

Lemma 3 [4]. Let $e_1, e_2 \dots e_n$ be the (exclusive) events with an associated exclusivity graph $G = (V, E)$. Then,

$$B_{\text{NC}}(G) = \text{STAB}(G),$$

$$B_Q(G) = \text{TH}(G),$$

$$B_E(G) = \text{QSTAB}(G).$$

Let $\text{QSTAB inequalities}$ for a graph G be the set of inequalities given by $\sum_{i \in Q} x_i \leq 1$ for every clique Q of the graph. Thus, $\text{QSTAB}(G)$ may be seen as the set of vectors $x \in \mathbb{R}^{|V|}$ such that $x_i \geq 0$, and the QSTAB-inequalities associated with G are satisfied.

Before we prove Theorem 1, note that the characterization of $\text{STAB}(G)$ was given in terms of its vertices and that of $\text{QSTAB}(G)$ was in terms of its hyperplanes. The following (known) link, Lemma 4, between these representations is key to the simplification.

Lemma 4 [39]. $\text{STAB}(G)$ is the convex hull of the integer solutions to the inequalities $x_i \geq 0$, and STAB-inequalities for G , where STAB inequalities for a graph $G = (V, E)$ are defined to be the set of inequalities given by $(x_i + x_j) \leq 1$ for every $(i, j) \in E$.

Proof of Theorem 1 We consider a five-cycle graph but our techniques readily generalize to the odd n -cycle case (unless stated otherwise). The QSTAB inequalities, together with the $x_i \geq 0$ condition, can be expressed as

$$0 \leq x_i \leq 1 \quad \text{for } i = \{1, 2 \dots 5\}, \quad (3)$$

$$x_i + x_{i+1} \leq 1 \quad \text{for } i = \{1, 2 \dots 5\} \quad (4)$$

where $i + 1$ is modulo 5. Note the STAB inequalities, together with $x_i \geq 0$, turn out to be exactly the same as the aforesaid for the five-cycle graph. (The set STAB is a convex hull of integer solutions of STAB inequalities.) Each inequality is characterized by a hyperplane. The vertices must lie on the intersection of (at least) five distinct hyperplanes. From this, we can already see that the integer solutions of STAB inequalities and the QSTAB inequalities are the same. The FCNC inequality is one of the facets defining NC inequality. To see this, it suffices to observe that there are exactly five vertices of STAB, the corresponding behaviors of which saturate the said inequality (since the space is five dimensional) and remaining vertices satisfy the same inequality. For the five-cycle case, the remaining argument is trivial and we defer the proof of the n -cycle case to the end.

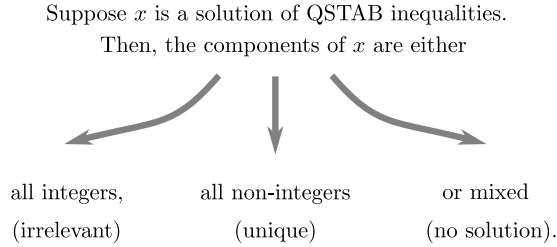


FIG. 1. To establish Theorem 1, it suffices to show that there is exactly one noninteger solution, i.e., a solution with at least one noninteger component, to QSTAB inequalities. We show that of the three kinds of solutions possible (as depicted above) the all-integer solution is irrelevant, the all-non-integer solution is unique, and no mixed solutions exist.

Note that, together with the aforementioned, if we can establish that there is only one noninteger solution of QSTAB inequalities then we have proven our result.

To this end, observe that there can only be the following three types of solutions: (1) all x_i are integers, (2) none of the x_i are integers, or (3) neither all x_i are integers nor all x_i are nonintegers (viz., at least one integer and at least one noninteger solution; see Fig. 1).

We are interested in the latter two cases. In case 2, we cannot use any of the QSTAB inequalities involving only one term [Eq. (3)]. This is because for a vertex we saturate five distinct inequalities. In this case, saturation of any of these inequalities will yield integer solutions which we are not considering. Hence, the only possibility is to use Eq. (4). Now we show that the solution is unique. Let $x_1 = q$ for any $0 < q < 1$. Saturating, we deduce $x_2 = 1 - q$, $x_3 = q$, $x_4 = 1 - q$, $x_5 = q$, and finally $x_1 = 1 - q$. This entails $x_1 = 1 - q = q$, which means $q = 1/2$ uniquely.

To complete the argument, we must show that there are no solutions in case 3. We already ruled out considering all five one term inequalities [Eq. (3)] as they yield integer solutions. Let us consider k two term inequalities [Eq. (4)] and m one term inequalities such that $m + k = 5$. The m one term inequalities, when saturated (because we consider the intersection of hyperplanes to obtain the vertices), will force the corresponding x_i s to be integers. This means that there are at least m integer x_i s. To analyze further, we consider the following game. Consider the five-cycle graph (see Fig. 2). Select m vertices of the graph (not to be confused with the vertices of QSTAB) and k edges. The vertices correspond to the variables fixed by the one term inequalities (saturated, so equalities). The edges correspond to the two term inequalities (again, saturated so equalities). Two cases can arise in such an assignment. Either each of the k edges is connected to one of the m vertices (possibly via other edges, if not directly) or there is at least one edge which is not connected to any of the m vertices (again, possibly via other edges, if not directly). These two cases are represented by the left and right graph in Fig. 2. Consider the second case. The disconnected edge (in the sense described earlier) will correspond to a two term equality involving two variables which have no other constraints. This means that the set of inequalities chosen does not uniquely determine a solution, i.e., at least one of the

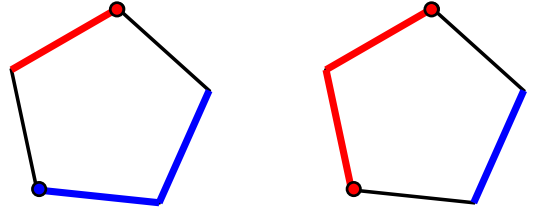


FIG. 2. There are two possible scenarios corresponding to the case where there is at least one integer and one noninteger solution (case 3 in the proof). The two term inequalities decide the values for two x_i s and have been represented as edges, and nodes (highlighted as small circles) have been used to denote the values determined by the one term inequalities. Depending on the way the combination of inequalities is selected, one gets either all x_i s as integers or a redundant set of inequalities leading to an undecidable value for x_i s.

inequalities chosen is redundant. This case is therefore irrelevant. Consider the first case now. In this case, start with any one of the m vertices. This corresponds to a one term equality which fixes the associated variable as an integer (as was noted earlier). Now the edge (if there is one) connected to this vertex directly will fix the value of the other vertex associated with the edge to be an integer. This reasoning can be repeatedly used to show that all the variables involved along the edges connected to the said initial vertex are integers. This can be repeated for every one of the m vertices. This means that all variables are assigned integer values. We have reached a contradiction which means there are no solutions of the kind assumed by case 3.

We end by showing that the FCNC inequality is facet defining (in the exclusivity graph approach). All incidence vectors (we will restrict to the ones corresponding to the stable set of the n -cycle graph, for this proof) will always satisfy the FCNC inequality because the cardinality of the stable set is bounded by the independence number (see Definition 6) of the graph, which for our case is $(n - 1)/2$ [40–42]. We will now show that there are exactly n vertices of STAB, i.e., incidence vectors which saturate the said inequality. To saturate, the incidence vector must have $(n - 1)/2$ components with entry one, and the remaining $(n + 1)/2$ components with entry zero. Note that each incidence vector satisfies the STAB inequalities, i.e., if a given component is one then its adjacent components are necessarily zero. One can convince themselves that any such vector, i.e., incidence vectors that saturate the FCNC inequality, must have two zeros adjacent (cyclically over n) while all other entries are alternatively one and zero. The total number of ways of placing two adjacent zeros, which is exactly n , then gives us the total number of incidence vectors which saturate the inequality, thereby proving that the FCNC inequality is indeed facet defining. ■

IV. UNIQUENESS OF FUNDAMENTAL ANTICYCLIC INEQUALITIES

We show that the odd anticycle admits a unique inequality. This follows easily from the following known result. For any set of nonnegative vectors X , its *antiblocker* is defined as

$$\text{abl } X = \{y \geq 0 : x \cdot y \leq 1 \ \forall x \in X\}.$$

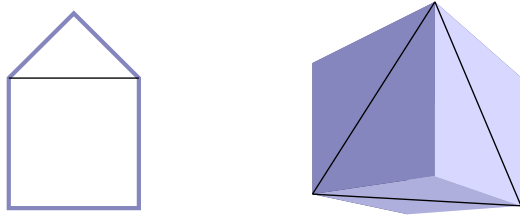


FIG. 3. Two conceivable illustrations of the QSTAB polytope (light blue) and the KCBS inequality (black) separating the said polytope from the STAB polytope. In the left image, the vertices of both polytopes are the same (except one) but there are two facets of QSTAB which are not a facet of STAB. One can show that all the facets of QSTAB are also the facets of STAB. This rules out the first image. The second image illustrates an alternative which helps us intuitively understand the higher dimension underlying geometry.

Let us denote by $\bar{G}$ the complement graph of G , viz., $\bar{G} = \{V, \bar{E}\}$ if $G = \{V, E\}$, which in particular means that if G is a cyclic graph then $\bar{G}$ is an anticyclic graph.

Lemma 5. For any graph G we have

$$\text{STAB}(\bar{G}) = \text{abl QSTAB}(G),$$

$$\text{TH}(\bar{G}) = \text{abl TH}(G),$$

$$\text{QSTAB}(\bar{G}) = \text{abl STAB}(G).$$

Note that $\text{abl } X = \text{abl convex hull } X$ because every element $y \in \text{abl } X$ will satisfy $(\theta_1 x_1 + \dots + \theta_k x_k) \cdot y \leq 1$ where $\sum_{i=1}^k \theta_i = 1$.

Theorem 2. Let G be odd n -cycle graph ($n > 3$) and consider the exclusivity graph scenario associated with $\bar{G}$. There is a unique facet-defining NC inequality, given by $\sum_{i=1}^n p_i \leq 2$.

Proof. We characterize $\text{QSTAB}(\bar{G})$ and $\text{STAB}(\bar{G})$ by using Lemma 5. Let $\{v_1, v_2 \dots v_n\}$ denote the vertices of $\text{STAB}(G)$. Each vertex v_i corresponds to a hyperplane constraining $\text{abl STAB}(G) = \text{QSTAB}(\bar{G})$. From the proof of Theorem 1 we know that $\text{QSTAB}(G)$ has exactly one more vertex, call it v_0 . Corresponding to v_0 there will be exactly one extra hyperplane constraining $\text{abl QSTAB}(G) = \text{STAB}(\bar{G})$ compared to those constraining $\text{abl STAB}(G) = \text{QSTAB}(\bar{G})$. This hyperplane is precisely $\sum_{i=1}^n p_i \leq 2$ using $v_0 = (1/2, 1/2, \dots, 1/2)$ and the definition of the antiblocker. ■

V. GEOMETRIC REPRESENTATION

Having proved the uniqueness of the fundamental noncontextuality inequalities, we briefly discuss the geometry of the associated polytopes.

Fix an odd $n > 3$. Geometrically, the FCNC inequality corresponds to a unique hyperplane cutting through QSTAB which separates all E-principle behaviors (see Definition 31 in Appendix B) uniquely into two parts, namely, noncontextual and contextual. Naïvely one might imagine the QSTAB polytope and the FCNC inequality to geometrically be illustrated by the image on the left in Fig. 3. However, it is not too hard to show that all the facets of QSTAB are also the facets of STAB, which means the naïve understanding is flawed. The image on the right in Fig. 3 better illustrates the geometry of the two convex polytopes (STAB and QSTAB). Recall that STAB for

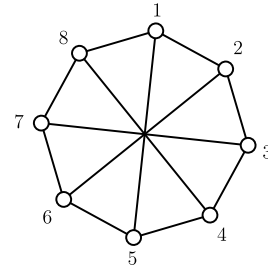


FIG. 4. The CHSH inequality is a unique nontrivial facet defining NC inequality for the (2,2,2)Bell scenario. The measurement events which are employed to get the CHSH inequality follow the exclusivity relations according to the $\text{Ci}_8[1, 4]$ graph.

a given graph is identical to the antiblocker of QSTAB for the complement graph and similarly QSTAB for a given graph is identical to the antiblocker of STAB for the complement graph (see Lemma 5). We thus recover essentially the same geometry for the anticyclic exclusivity scenario. In summary, given an n -cyclic (or anticyclic) exclusivity scenario, the associated fundamental NC inequality separates the corresponding E-principle behaviors into parts and uniquely singles out a vertex corresponding to the maximally contextual behavior.

VI. ALL BELL INEQUALITIES FOR THE SIMPLEST EXCLUSIVITY GRAPH

Any exclusivity scenario witnessing contextuality will have associated with it at least one fundamental NC inequality but there may be others; we give an example of this using the simplest Bell exclusivity scenario and report a Bell inequality in the process.

Given a list $[L]$ of integers, a graph with n vertices where every i th vertex is connected to every other $(i + l) \bmod n$ th vertex for $l \in [L]$ is called a circulant graph $\text{Ci}_n[L]$. The exclusivity graph corresponding to the measurement events for the CHSH inequality is a circulant graph. In particular, it is represented as the $\text{Ci}_8[1, 4]$ graph (see Fig. 4). This is the simplest exclusivity graph which can lead to Bell inequalities if one analyzes the corresponding stable set polytope [36].

We ran numerical tests to characterize the convex polytope $\text{STAB}(\text{Ci}_8[1, 4])$ and found three types of nontrivial facets.

(1) Pentagonal inequalities: There are eight such inequalities and they correspond to eight induced pentagons (five cycles). In the literature, these are known as pentagonal Bell inequalities [36]. Formally, the pentagonal inequalities are given by

$$\sum_{i=k}^{(k+4) \bmod 8} p_i \leq 2$$

for $k \in \{1, 2, 3, 4, 5, 6, 7, 8\}$.

(2) Heptagonal inequalities: There are eight such inequalities and correspond to induced subgraphs of $\text{Ci}_8[1, 4]$ with any seven nodes. Formally, the heptagonal NC inequalities are given by

$$\sum_{i=k}^{(k+6) \bmod 8} p_i \leq 3$$

for $k \in \{1, 2, 3, 4, 5, 6, 7, 8\}$.

(3) CHSH inequality: The CHSH inequality corresponds to the sum of the probabilities of all eight events and is formally given by

$$\sum_{i=1}^8 p_i \leq 3.$$

Note that neither the Heptagonal inequalities nor the CHSH inequality correspond to any odd cycle (or its complement).

VII. RELATIVE SIMPLIFICATION EXPLAINED

Where is this exponential simplification coming from and are we losing something in the process? To answer these questions we connect the inequality we obtained using the exclusivity graph approach to the one obtained using the compatibility hypergraph approach. We briefly introduce the compatibility hypergraph approach first (deferring a complete description to Appendix C).

A. Compatibility hypergraph approach: Overview

Instead of looking at the exclusivity of events, the compatibility hypergraph approach is based on, as the name suggests, the compatibility of measurements. The scenario here is defined by a hypergraph G , where the vertices $V(G)$ represent the measurements and the hyperedges $\mathcal{C}(G)$ of the graph represent the set of measurements which are compatible (see Definition 33 and 34 in Appendix C). The set of nodes or measurements in a hypergraph constitute a *context* (see Definition 20 in Appendix A). For each context $C \in \mathcal{C}$, a normalized probability vector p_C is defined which assigns probabilities to every possible joint outcome corresponding to the measurements belonging to that context (see Definition 20 in Appendix A). The behavior p , in this formalism, is defined to be the concatenation of all these probability vectors (one vector for each context; see Definition 37 in Appendix C). Consider a fixed measurement. The behaviors which assign the same marginal probability distribution to the outcomes corresponding to this measurement, regardless of which probability vector (and hence context) was used to evaluate the marginal, are said to be *nondisturbing behaviors*. The set of such behaviors happens to be a convex polytope and is denoted by $\mathcal{X}(G)$ (pronounced as X ; see Definition 40 in Appendix C). It is analogous (and reduces, under the appropriate restrictions) to the no-signaling distributions. Consider the set of behaviors which arise by associating a quantum observable with each measurement such that the compatibility requirements are satisfied. This set is called the *set of quantum behaviors* $\mathcal{Q}(G)$ and it happens to be convex. Finally, consider behaviors which admit a noncontextual completion (see Definition 22 in Appendix A). These behaviors can be equivalently characterized as behaviors which arise from a global joint distribution over all the observables such that the marginals yield the various probability vectors constituting the behavior (see Theorem 3 in Appendix C). These define the *set of classical behaviors* $\mathcal{C}(G)$ which, like $\mathcal{X}(G)$, happens to be a convex polytope. The facet-defining hyperplanes of $\mathcal{C}(G)$ yield NC inequalities.

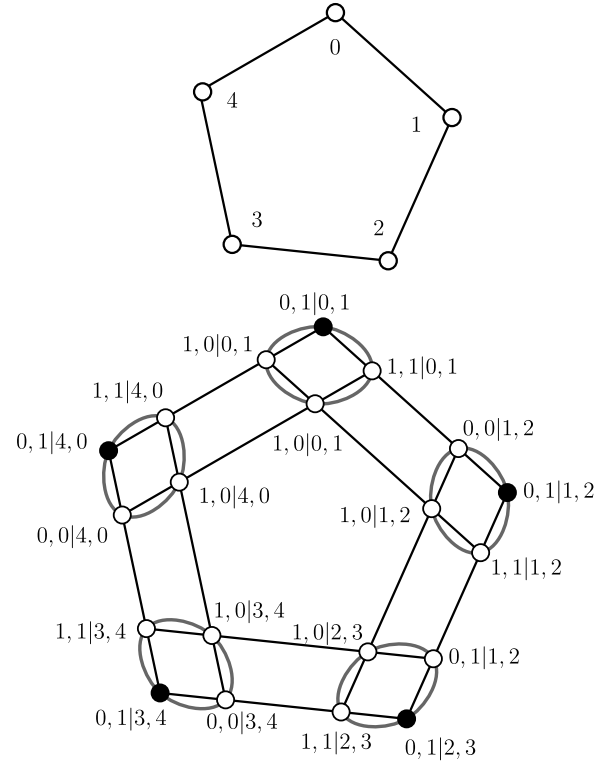


FIG. 5. The KCBS scenario can be expressed using both the compatibility hypergraph (top) and exclusivity graph (bottom) approaches. In the top figure, the nodes represent measurements, while in the bottom figure the nodes represent events.

B. KCBS inequality

For concreteness, we consider the KCBS scenario in the compatibility hypergraph formalism. Let $\{M_0, M_1, M_2, M_3, M_4\}$ be five dichotomic measurements, which form the set of vertices $V(G)$, and let $\{M_0, M_1\}, \{M_1, M_2\}, \{M_2, M_3\}, \{M_3, M_4\}, \{M_4, M_0\}$ denote the compatibility relations among them $\mathcal{C}(G)$, which define the (hyper)edges. This results in G being a five-cycle (hyper)graph [see Fig. 5 (top)]. It would be helpful to use two different conventions for labeling the outcomes of the measurements: binary (0,1) and signed (+1, -1). In the binary convention, the probability of obtaining the outcome (0,0) upon the measurement of M_i and M_{i+1} is denoted by $p_{i,i+1}(00)$. This is consistent with our notation for the probability vector p_C where C was the context, which in this case is denoted by $i, i+1$. The addition in the indices is modulo 5. Similarly in the signed convention, the outcome (+1, +1) is denoted by $p_{i,i+1}(++)$. Further, to represent the expectation value of the product of measurements $M_i M_{i+1}$ we use the notation

$$\begin{aligned} \langle i, i+1 \rangle &= p_{i,i+1}(++) + p_{i,i+1}(--) \\ &\quad - p_{i,i+1}(+-) - p_{i,i+1}(-+). \end{aligned} \quad (5)$$

Note that we discuss expectation values only in the signed convention. By studying the polytope of classical behaviors $\mathcal{C}(G)$ one can find the facet-defining hyperplanes which yield the NC inequalities. These turn out to be [35]

$$\gamma_0 \langle 0, 1 \rangle + \gamma_1 \langle 1, 2 \rangle + \gamma_2 \langle 2, 3 \rangle + \gamma_3 \langle 3, 4 \rangle + \gamma_4 \langle 4, 0 \rangle \leq 3$$

where $\gamma_i \in \{\pm 1\}$ such that the number of γ s with a negative sign is odd. To see how these relate to the unique KCBS inequality we obtain using the exclusivity graph approach, consider the possible events corresponding to a given context [the (hyper)edge]. All these events are mutually exclusive because they correspond to different outcomes of a given set of measurement [see Fig. 5 (bottom)]. This exclusivity is denoted by the ellipse in the graph. Now consider the measurement M_0 . Let $(01|M_0M_1)$ denote the event where M_0 and M_1 are measured and the outcomes are zero and one, respectively (see Definitions 23 and 24 in Appendix B). Using this notation, observe that the events $(01|M_0M_1)$, $(00|M_0M_1)$, $(11|M_4M_0)$, and $(01|M_4M_0)$ are mutually exclusive. The first two are exclusive as they correspond to different outcomes of M_1 , the second and third are exclusive because they correspond to different outcomes of M_0 , and the third and fourth are exclusive because they correspond to different outcomes of M_4 . A similar argument works for all the other pairs. This exclusivity is denoted by a straight line. One can verify that all the events on a given straight line are exclusive. If we let $e_i = (01|M_iM_{i+1})$ then a five-cycle exclusivity graph can be extracted from the aforementioned by defining the set of vertices to be the set of events $\{e_0, e_1, \dots, e_4\}$ and the set of edges to be their exclusivity relations $\{\{e_0, e_1\}, \{e_1, e_2\}, \dots, \{e_4, e_0\}\}$. We had denoted the probability of these events by p_i to write the KCBS inequality, which in our current notation becomes

$$p_{0,1}(01) + p_{1,2}(01) + p_{2,3}(01) + p_{3,4}(01) + p_{4,0}(01) \leq 2. \quad (6)$$

Since the exclusivity graph approach does not require the explicit specification of the measurements which lead to exclusivity, merely their existence, it is able to extract the essential nature of the problem without creating redundancies caused by the labeling. If the complete exclusivity graph was used then the exclusivity graph formalism should yield effectively [43] the same NC inequalities as the compatibility hypergraph approach, recreating the said redundancies. We now explicitly combine the following two NC inequalities obtained using two different exclusivity graphs,

$$\sum_{i=0}^4 p_{i,i+1}(01) \leq 2, \quad \sum_{i=0}^4 p_{i,i+1}(10) \leq 2,$$

into an NC inequality obtained using the compatibility hypergraph formalism:

$$-\sum_{i=0}^4 \langle i, i+1 \rangle \leq 3, \quad (7)$$

which corresponds to taking all $\gamma_i = -1$. Using $p_{i,i+1}(00) = p_{i,i+1}(++)$ (which holds by assumption; they were just different labels for the same outcomes) in Eq. (5), and probability conservation (probability vectors p_C are normalized; in a given context, the probabilities sum to one), we deduce

$$\langle i, i+1 \rangle = 1 - 2[p_{i,i+1}(01) + p_{i,i+1}(10)]. \quad (8)$$

To obtain Eq. (7) we sum the two inequalities

$$\begin{aligned} & \sum_{i=0}^4 [p_{i,i+1}(01) + p_{i,i+1}(10)] \leq 4 \\ \Leftrightarrow & \sum_{i=0}^4 \underbrace{[2p_{i,i+1}(01) + p_{i,i+1}(10) - 1]}_{=-\langle i, i+1 \rangle} \leq 8 - 5 = 3. \end{aligned}$$

While at first sight it might appear that the inequality obtained using the compatibility hypergraph is weaker as it is a linear combination of two exclusivity graph based inequalities, this conclusion is incorrect. This is because we can do better. We can obtain from a single compatibility hypergraph based inequality a corresponding exclusivity graph based inequality. We show this explicitly for $\sum_{i=0}^4 p_{i,i+1}(01) \leq 2$. The marginal $p_i(1) = \sum_{l=0}^1 p_{i,j}(1l)$ for every i, j belonging to a hyperedge. We start with noting that $p_{i,i+1}(10) + p_{i,i+1}(01) = p_i(1) - p_{i,i+1}(11) + p_{i,i+1}(01)$ but we can also write $p_i(1) = p_{i-1,i}(01) + p_{i-1,i}(11)$ (this is a consequence of the no-disturbance requirement). Consequently, using Eq. (8), we have

$$\begin{aligned} & 1 - \langle i, i+1 \rangle \\ &= 2[p_{i-1,i}(01) + p_{i-1,i}(11) - p_{i,i+1}(11) + p_{i,i+1}(01)] \\ &\Rightarrow 5 - \sum_{i=0}^4 \langle i, i+1 \rangle = 4 \sum_{i=0}^4 p_{i,i+1}(01) \end{aligned}$$

and substituting Eq. (7) we obtain $4 \sum_{i=0}^4 p_{i,i+1}(01) \leq 5 + 3$, which is just Eq. (6).

The NC inequalities for the compatibility hypergraph structure corresponding to anticycles are not known in the literature. A numerical investigation for the classical polytope for the same indicates the presence of exponentially many nontrivial NC inequalities. However, in the exclusivity graph approach, the nontrivial NC inequalities corresponding to anticycles are unique (see Theorem 2). The reason for the simplification is the same as that for the cyclic case.

VIII. DISCUSSION AND CONCLUSION

We showed that all fundamental NC inequalities are unique for their corresponding odd cycle (or anticycle) exclusivity scenario. This is an exponential simplification compared to the analogous compatibility hypergraph scenarios. Thereafter, we discussed the geometry of the associated polytopes. Every exclusivity scenario which exhibits contextuality must admit at least one fundamental NC inequality; however, a behavior might exhibit contextuality by violating some other NC inequality. We reported a Bell inequality in the process of exemplifying this phenomenon. Finally, we also analyzed the source of the simplification by relating the KCBS scenario in the exclusivity graph approach with the compatibility hypergraph approach.

We outline two directions for further exploration.

(1) The first direction is the linking of our results to the resource theory of contextuality. The simplification for cyclic and anticyclic exclusivity scenarios indicates that there is a unique way to capture the amount of contextuality, to wit,

TABLE I. The anticycle case appears naturally in the exclusivity graph approach due to the strong perfect graph theorem. We show that both the cyclic and anticyclic case in this approach admit a unique NC inequality while Araújo *et al.* [35] obtained exponentially many NC inequalities in the compatibility hypergraph approach. The characterization of the anticycle case for the latter approach is an open problem.

	Exclusivity	Compatibility
Cycles	This paper	Araújo <i>et al.</i> [35]
Anticycles	This paper	?

the distance of the contextual behavior from the hyperplane corresponding to fundamental NC inequalities. The strong perfect graph theorem could potentially allow a generalization to all exclusivity scenarios.

(2) The second direction is the exploration of anticycles in the compatibility hypergraph approach. We characterized the NC polytopes corresponding to the cycles and anticycles in the exclusivity graph approach and compared the former with its compatibility hypergraph analog. The characterization of the anticycle case in the latter approach is still open (see Table I).

We note that self-testing using the NC inequality for the cyclic case was shown in [15] and it was recently extended to the anticycle case as well [44]. Our results, together with the aforementioned works on self-testing, show that this characterization is complete in the following sense: in the cyclic and anticyclic case, there is a unique (fundamental) NC inequality and this can be self-tested. Further, by the strong perfect graph theorem, every exclusivity scenario which can exhibit contextuality must contain at least one such fundamental NC inequality.

ACKNOWLEDGMENTS

We are thankful to Michael Kleider for the CON2VERT and VERT2CON MATLAB packages that gave us the first numerical glimpse of our result. A.S.A. and J.R. acknowledge financial support from the Belgian Fonds de la Recherche Scientifique (FNRS) under Grants No. F.4515.16 (QUICTIME) and No. R.50.05.18.F (QuantAlgo). A.S.A. further acknowledges the FNRS for support through Grant No. F3/5/5-MCF/XH/FC-16749 FRIA. K.B. acknowledges the Centre for Quantum Technologies (CQT) Graduate Scholarship. K.B. and L.C.K. are grateful to the National Research Foundation and the Ministry of Education, Singapore for financial support.

APPENDIX A: PROBABILISTIC MODELS: STATES AND MEASUREMENTS

The results discussed here are based on the work of Amaral and Cunha [32]. In any experimental scenario there are two types of interventions possible, either preparation or operation. Preparation is used in the intuitive sense of the word, that is, preparing the system in a given state, for instance using a laser to initialize the state of an atom. More explicitly, we make the following assumptions about the theory.

(1) Interventions are of two types: Preparation and operation.

(2) Experiments are reproducible: For each operation, there may be several different outcomes, each occurring with a well-defined probability for a given preparation.

Definition 11. State. Two preparations are defined to be equivalent if they give the same probability distribution for all available operations. We will refer to an equivalence class of preparations as a *state*.

Definition 12. State space. The set of all states is referred to as the *state space* of the system.

Remark 2. The state space is convex.

Definition 13. Pure states. All extremal points of the state space are defined to be pure states.

Definition 14. Measurements. Measurements are operations with more than one outcome.

Remark 3. Unitary evolution is an example of an operation which is not a measurement.

Definition 15. Probabilistic model. We call any mathematical description of a physical system which provides the following a *probabilistic model*: (1) objects to represent [(a) state, (b) operations, and (c) measurements] and (2) a rule to calculate the probabilities of the possible outcomes of any arbitrary measurement given any arbitrary state.

Definition 16. Probability theory. A *probability theory* is a collection of probabilistic models.

Definition 17. Outcome repeatable measurements. A measurement j is defined to be an outcome repeatable measurement if every time one performs this measurement on a system and an outcome k is obtained a subsequent measurement of j on the same system gives the outcome k again with probability 1.

Definition 18. $p_j(k)$. The probability of getting an outcome k given that a measurement j has been performed will be denoted by $p_j(k)$.

All the measurements henceforth will be assumed to be outcome repeatable.

Definition 19. Compatible measurements, refinement, and coarse graining. A set of measurements $\{j_1, j_2 \dots j_n\}$ is *compatible* if there is another measurement j with outcomes $\{1, \dots m\}$ and functions $\{f_1, f_2 \dots f_n\}$ such that the possible outcomes of j_s are the same as $\{f_s(1), f_s(2) \dots f_s(m)\}$ for each s and

$$p_{j_s}(l) = \sum_{k \in f_s^{-1}(l)} p_j(k)$$

where j is called a *refinement* of $\{j_1, j_2 \dots j_n\}$ and each j_s is called a *coarse graining* of j .

If a set of measurements is compatible it is called a *set of compatible measurements*.

Completion of a probabilistic model

Definition 20. Context. A set of compatible measurements is defined to be a context.

Our objective now is to construct a general mathematical framework which can describe the completion of a probabilistic model, i.e., give a model which is no longer probabilistic but reduces to the same probabilistic model if certain variables are ignored.

Definition 21. Completion. Consider a probabilistic model P where S represents the set of pure states and X represents the set of measurements. The *completion* of this probabilistic model, denoted by P' , consists of a set of measurements X' , which are in one-to-one correspondence with X , and a set of pure states S' , which are in one-to-one correspondence with $\Lambda \times S$ for some set Λ . P' must satisfy the following requirements. For all $\rho \in S$ and all contexts $c = \{j_1, j_2 \dots j_n\}$, P' should specify a probability distribution over Λ given by $p(\lambda)$ and a probability distribution $p_{j_k}^{(\lambda, \rho, c)} : \mathbb{R} \rightarrow \{0, 1\}$ for each λ, ρ, c, j_k such that

$$p_c^\rho(i_1, i_2 \dots i_m) = \sum_{\lambda \in \Lambda} p(\lambda) \prod_{k=1}^m p_{j_k}^{(\lambda, \rho, c)}(i_k)$$

where $p_c^\rho(i_1, i_2 \dots i_m)$ is the probability assigned by P to the measurement of $j_1, j_2 \dots j_n$ (encoded in c) yielding the outcomes $i_1, i_2 \dots i_m$, respectively, for the state ρ .

Remark 4. We expect the completion P' to specify S' as (λ, ρ) for all $\lambda \in \Lambda$ and $\rho \in S$. Let us assume for simplicity that $X' = X$. Now for every context $c = \{j_1, j_2 \dots j_n\}$ (i.e., set of compatible measurements from X) the completion P' will predict with certainty the outcome of measuring any $j_i \in c$, for a given (λ, ρ) . This prediction is allowed to depend on the set c itself to accommodate “contextual completions.” We will see later that noncontextual (and functionally consistent) completions contradict the predictions of quantum mechanics.

Let X be a set of measurements. Let $\{j_1, j_2 \dots j_m\} \subset X$ be a set of compatible measurements.

Definition 22. Noncontextual completion. Let $c_1 = \{j_1, j_2, \dots j_m\}$, $c_2 = \{j'_1, j'_2, \dots j'_m\}$ be two contexts (note that j_i and j'_k may not be compatible for $i, k > 1$). A completion P' of a probabilistic model P is called *noncontextual* if $p_{j_i}^{(\lambda, \rho, c_1)}(i) = p_{j_i}^{(\lambda, \rho, c_2)}(i)$ for all contexts c_1 and c_2 of the aforesaid form.

APPENDIX B: THE EXCLUSIVITY GRAPH APPROACH

1. Formalizing scenarios

Definition 23. Measurement event. We denote a *measurement event* by $(a_1 a_2 \dots a_n | M_1 M_2 \dots M_n)$ where a_i is the measurement outcome associated with M_i , and M_i is compatible with M_j for all $i, j \in \{1, 2 \dots n\}$. Two measurement events $(a_1 a_2 \dots a_n | M_1 M_2 \dots M_n)$ and $(a | M)$ are *equivalent* if for all states (see Definition 11) their probabilities of occurrence of these events are equal.

For brevity, we will use the word *event* in lieu of equivalent measurement events whenever there is no ambiguity.

Definition 24. Exclusive event. Two events e_i and e_j are defined to be *exclusive* if there exists a measurement M such that e_i and e_j correspond to different outcomes of M , i.e., $e_i = (a_i | M)$ and $e_j = (a_j | M)$ such that $a_i \neq a_j$.

Definition 25. Exclusivity graph. For a family of events $\{e_1, e_2 \dots e_n\}$ we associate a simple undirected graph, $G := (V, E)$, with vertex set V and edge set E such that two vertices $i, j \in V$ share an edge if and only if e_i and e_j are exclusive events. G is called an *exclusivity graph*.

Definition 26. Probability vector. For a given exclusivity graph $G = (V, E)$ and a probability theory, the *probability vector* is a vector $p \in \mathbb{R}^{|V|}$ such that $p_{(i)} = \text{prob}(e_i)$ where $\text{prob}(e_i)$ is the probability assigned by the probability theory to the event e_i .

Definition 27. Behavior. A *behavior* for an exclusivity graph $G = (V, E)$ is a map $p : V \rightarrow [0, 1]$ which assigns to each vertex $i \in V$ a probability $p(i)$ such that $p(i) + p(j) \leq 1$, for all vertices that share an edge, i.e., $(i, j) \in E(G)$. Due to the isomorphism between the map $p : V \rightarrow [0, 1]$ and the vector $\vec{p} \in \{0, 1\}^{|V|}$ we will associate with the i th component of $\vec{p}$ the value $p(i)$, i.e., $\vec{p}_{(i)} = p(i)$. (Sometimes we will even drop the vector sign.)

Remark 5. We do not use $p_{(i)} = p_M(i)$ because M is not explicitly *a priori* known so cluttering the notation does not help.

Definition 28. Noncontextual behavior. A behavior p is called a *deterministic noncontextual behavior* if $p : V \rightarrow \{0, 1\}$, i.e., $p_{(i)} \in \{0, 1\}$ for all i and there exists a noncontextual completion of the corresponding probabilistic model P . The set of noncontextual behaviors is defined to be the convex hull of deterministic noncontextual behaviors and is denoted by $B_{\text{NC}}(G)$.

Remark 6. Defining the behavior this way implicitly imposes functional consistency. This is because we require a noncontextual completion of deterministic behaviors to start with and later take its convex combination. This imposes the exclusivity condition at the level of the hidden variable model, which in turn is a manifestation of functional consistency.

Definition 29. Quantum behavior. A behavior for an exclusivity graph G is called a *quantum behavior* if there exists a quantum state ρ and projectors $\Pi_1, \dots, \Pi_n$ acting on a Hilbert space $\mathcal{H}$ such that $p_{(i)} = \text{Tr}(\rho \Pi_i)$ for all $i \in V$ and $\text{Tr}(\Pi_i \Pi_j) = 0$ for vertices that share an edge, i.e., $(i, j) \in E$.

The convex set of all quantum behaviors is denoted by $B_Q(G)$.

Definition 30. Exclusivity principle. Given a subset $\{e_i\}$ of events which are pairwise exclusive we say that the exclusivity principle is obeyed by a probabilistic model if $\sum_i \text{prob}(e_i) \leq 1$ for all such subsets. We will sometimes refer to this as the E principle.

Definition 31. E-principle behavior. A behavior p for an exclusivity graph G is said to be an E-principle behavior if the associated probabilistic model satisfies the exclusivity principle, i.e., $\text{prob}(e_i) = p_{(i)}$ satisfies the E principle.

The set of E-principle behaviors will be denoted by $B_E(G)$.

Let $e_1, e_2 \dots e_n$ denote a family of measurement events.

Remark 7. The set $B_{\text{NC}}(G)$ is a (convex) polytope, i.e., can be expressed as a solution of a finite number of linear inequalities.

Definition 32. NC inequality, facet defining. Let p be a behavior and $\gamma_i, \beta \in \mathbb{R}$. A linear inequality, $\sum \gamma_i p_{(i)} \leq \beta$, is called an *NC inequality* if its satisfaction is a necessary condition for membership to the set $B_{\text{NC}}(G)$. Equivalently, to claim nonmembership in the set $B_{\text{NC}}(G)$, it is sufficient to show a violation of the said linear inequality.

An NC inequality is called *facet defining* if it defines a nontrivial facet of $B_{\text{NC}}(G)$.

2. Impossible completions: Linking geometry and quantum mechanics

Lemma 6 [4]. Let $e_1, e_2 \dots e_n$ be the (exclusive) events associated with an exclusivity graph $G = (V, E)$. Then,

$$\begin{aligned} B_{\text{NC}}(G) &= \text{STAB}(G), \\ B_Q(G) &= \text{TH}(G), \\ B_E(G) &= \text{QSTAB}(G). \end{aligned}$$

Corollary 1. For a given exclusivity graph G we have

$$B_{\text{NC}}(G) \subseteq B_Q(G) \subseteq B_E(G).$$

APPENDIX C: COMPATIBILITY HYPERGRAPH APPROACH

1. Formalizing scenarios

Definition 33. Compatibility scenario. A compatibility scenario is specified by the tuple $\Upsilon := (X, \mathcal{C}, O)$ where the following hold.

- (1) O is a finite set.
- (2) X is a finite set of random variables from O to $\mathcal{P}(O)$.
- (3) $\mathcal{C}$ is a collection of subsets of X such that their union is equal to X and the intersection of any two subsets is never equal to one of the subsets. Each of these subsets will be called a context.

It might be useful to keep the following in mind to get an intuitive understanding. The set X can be thought of as the measurements, the set O as the outcome of these measurements, and the set $\mathcal{C}$ as containing maximal contexts.

Definition 34. Compatibility hypergraph. The compatibility hypergraph corresponding to a scenario $\Upsilon := (X, \mathcal{C}, O)$ is a hypergraph the nodes of which are elements in X and hyperedges are contexts in $\mathcal{C}$. We denote the compatibility hypergraph for the scenario by $\mathcal{H} = (X, \mathcal{C})$.

We use calligraphic letters to notationally distinguish the (hyper)graphs associated with the compatibility hypergraph from those associated with the exclusivity graph approach.

Definition 35. Compatibility graph. The compatibility graph for a scenario is the two-skeleton of the corresponding hypergraph, i.e., $\mathcal{H} = (X, \mathcal{C})$. We will denote the compatibility graph by $\mathcal{G} = (X, E)$. Given two elements i and j in X , they share an edge in E if and only if $\{i, j\} \subseteq C$ for some context $C \in \mathcal{C}$.

Definition 36. Measurement event. A measurement event corresponds to a single run of an experiment where the measurements in a context $C \in \mathcal{C}$ are jointly performed with outcomes in $O^{|C|}$.

Henceforth, for notational simplicity, we use O^C instead of $O^{|C|}$.

Definition 37. Behavior and behavior vector. Given a scenario $\Upsilon := (X, \mathcal{C}, O)$, a behavior is a family of probability distributions defined over O^C , defined as

$$B = \left\{ p_C : O^C \rightarrow [0, 1] \mid \sum_{s \in O^C} p_C(s) = 1, C \in \mathcal{C} \right\}.$$

One can stack the $p_C(s)$ for all $s \in O^C$ and $C \in \mathcal{C}$ to form a column vector of probabilities for a behavior B . Such probability vectors are called *behavior vectors*.

Remark 8. The set of possible behaviors forms a polytope with $\prod_{C \in \mathcal{C}} |O^C|$ nodes, where the extreme points (nodes) correspond to deterministic points, i.e., $p_C(s)$ equal to zero or one. This can be deduced from the convexity of probability distributions.

Definition 38. Restriction map. Given a context $C \in \mathcal{C}$ with outcomes in O^C and a set $U \subset C$, a restriction map r_U^C is given by

$$r_U^C : O^C \rightarrow O^U,$$

$$s = (a_i)_{M_i \in C} \mapsto s|_U = (a_i)_{M_i \in U}.$$

Definition 39. Marginal distribution (for a context). The marginal distribution for a probability distribution over a context $C \in \mathcal{C}$ corresponding to a set $U \subset C$ is defined as

$$p_U^C : O^U \rightarrow [0, 1],$$

$$p_U^C(s) = \sum_{s' \in O^C : r_U^C(s') = s} p_C(s'),$$

where r_U^C is a restriction map from O^C to O^U .

Definition 40. Nondisturbing set, $\mathcal{X}(\Upsilon)$. Given a scenario $\Upsilon := (X, \mathcal{C}, O)$, the set of behaviors is called a nondisturbing set, if for any given behavior and two different contexts C_1 and C_2 we have $p_{C_1 \cap C_2}^{C_1} = p_{C_1 \cap C_2}^{C_2}$.

Definition 41. Global section. Given a scenario $\Upsilon := (X, \mathcal{C}, O)$, a global section for X is a probability distribution over O^X , denoted by $p_X : O^X \rightarrow [0, 1]$.

Definition 42. Global section for a behavior. Given a scenario $\Upsilon := (X, \mathcal{C}, O)$ and a behavior B , a global section p_X is called a global section for the behavior B if

$$p_C^X = p_C \quad \forall \quad C \in \mathcal{C}.$$

Definition 43. Noncontextual behavior. A behavior which admits a global section is called a noncontextual behavior.

Remark 9. Noncontextual completions were defined independently.

Theorem 3. Abramsky and Brandenburger [2]. Given a scenario $\Upsilon := (X, \mathcal{C}, O)$ and a behavior $B \in \mathcal{X}(\Upsilon)$, the aforementioned behavior B has a global section if and only if there exists a noncontextual completion recovering B .

2. Probability distributions and physical theories

a. Classical realizations and noncontextuality

Remark 10. Given a scenario Υ , the set of classical behaviors will be denoted by $\mathcal{C}(\Upsilon)$. Note that this notation is distinct from the $\mathcal{C}$ which was used to define the compatibility cover for the scenario Υ .

Claim 1. Given a scenario Υ and a behavior B , the following statements are equivalent.

- (1) B has a global section.
- (2) B is classical.

Remark 11. There exists a noncontextual completion recovering B .

b. Quantum realization

Remark 12. The set of quantum behaviors will be denoted by $\mathcal{Q}(\Upsilon)$.

Theorem 4. $\mathcal{Q}(\Upsilon)$ is a convex set.

Remark 13. Restricting the dimension of realization in the quantum case yields a nonconvex set.

c. Noncontextuality inequalities

Remark 14. Given a scenario Υ and a behavior B , how can one determine if there exists a noncontextual completion recovering B ? This motivates us to define linear inequalities, the violation of which (for a behavior B) guarantees that B is contextual, i.e., there is no noncontextual completion recovering B . It is important to note that the set of noncontextual behaviors forms a polytope, which means that characterization of the same can be given by intersection of finitely many hyperplanes and half spaces. NC inequalities for a scenario correspond to the facets of the classical (or noncontextual) polytope. Note that the set of nondisturbing behaviors $\mathcal{X}(\Upsilon)$ is also a polytope.

Given a polytope, the representation in terms of half spaces and hyperplanes is often called H representation. The same polytope can also be described as a convex hull of finitely many vertices of the polytope, referred to as V representation.

Definition 44. NC inequality. The NC inequality is a linear inequality

$$S := \sum_{s \in \mathcal{O}^c, C \in \mathcal{C}} \gamma_C(s) P_C(s) \leq b$$

where $\gamma_C(s)$ and b are real numbers such that the inequality is satisfied for the behaviors in the noncontextual polytope $\mathcal{C}(\Upsilon)$. Often b is called the noncontextual hidden variable bound because the noncontextual behaviors [the behaviors in $\mathcal{C}(\Upsilon)$] respect the bound.

Remark 15. There may exist behaviors in $\mathcal{X}(\Upsilon)$ which violate one or many NC inequalities for the scenario Υ . Such behaviors are often called *contextual behaviors*. An NC inequality is called *tight* if there exists a noncontextual behavior saturating the inequality. Furthermore, it is called *facet defining* if it corresponds to one of the facets of the noncontextual polytope. Given a behavior B , its membership in $\mathcal{C}$ is equivalent to checking if all the facet defining NC inequalities are satisfied.

d. Geometry of the case $\mathcal{H} = \mathcal{G}$ (compatibility hypergraph = compatibility graph)

Remark 16. If we assume that every context has at most two measurements, then the compatibility hypergraph for the scenario is given by its two-skeleton. Furthermore, if we assume that every measurement has two outcomes, it leads to description of classical and no-disturbing sets as familiar polytopes from graph theory.

Description of the nondisturbing quantum and noncontextual behaviors.

Remark 17. The nondisturbing set $\mathcal{X}(\Upsilon)$ lies in $\mathbb{R}^{4|E(G)|}$ because every edge (or equivalently context) corresponds to two binary measurements.

Definition 45. Notation. Fix $\{M_i, M_j\} \in \mathcal{C}$.

(1) $p_{ij}(ab)$ is the probability of outcome a and b for the joint measurement of i and j , respectively.

(2) $p_i(a) = \sum_b p_{ij}(ab)$.

(3) $p_j(b) = \sum_a p_{ij}(ab)$.

Claim 2. $p_{ij}(ab)$ can be determined from $p_{ij}(11)$ and $p_i(1)$ due to the constraints on nondisturbing behaviors.

Definition 46. $\phi : \mathbb{R}^{4|E(G)|} \rightarrow \mathbb{R}^{|V(G)|+|E(G)|}$, $B \mapsto q = (q_i, q_{kj})_{i \in V(G); (k,j) \in E(G)}$ where $q_i = p_i(1)$ and $q_{ij} = p_{ij}(11)$.

Remark 18. To return from the q space to the B space, we use

$$p_{ij}(10) = q_i - q_{ij},$$

$$p_{ij}(01) = q_j - q_{ij},$$

$$p_{ij}(00) = 1 - q_i - q_j + q_{ij}.$$

Remark 19. The map is injective.

Definition 46. Correlation vector $v(S)$, correlation polytope. Correlation vector $v(S)$ is defined as follows. Given $S \subset V(G)$, $v(S) \in \mathbb{R}^{|V|+|E|}$ is defined as

$$v(S)_i = \begin{cases} 1 & i \in S \\ 0 & \text{else} \end{cases} \quad \forall i \in V,$$

$$v(S)_{ij} = \begin{cases} 1 & i, j \in S \\ 0 & \text{else} \end{cases} \quad \forall \{i, j\} \in E.$$

The *correlation polytope* is defined to be the convex hull of all correlation vectors.

Theorem 5. $\phi[\mathcal{C}(\Upsilon)] = \text{COR}(G)$.

Definition 47. (For completeness) Rooted correlation semi-metric polytope. $\text{RCMET}(G)$ of a graph G is the set of vectors

$$q = (q_i, q_{jk}) \in \mathbb{R}^{|V(G)|+|E(G)|}$$

such that

$$q_{ij} \geq 0,$$

$$q_i - q_{ij} \geq 0,$$

$$1 - q_i - q_j + q_{ij} \geq 0.$$

Theorem 6. $\phi[\mathcal{X}(G)] = \text{RCMET}(G)$.

e. Cut polytope

Definition 48. Cut vector, cut polytope. Given a graph G and $c \in \{0, 1\}^{|V(G)|}$ the cut vector is $x(c) \in \mathbb{R}^{|E(G)|}$ such that

$$x(C)_{ij} = c_i \oplus c_j.$$

The *cut-01 polytope*, $\text{CUT}^{01}(G)$, is the convex hull of all cut vectors of G .

Definition 49. ± 1 cut vectors. Given a graph G and $c \in \{0, 1\}^{|V(G)|}$ the cut vectors are defined as $y(c) \in \mathbb{R}^{|E(G)|}$ such that

$$y(c)_{ij} = c_i c_j.$$

The *cut ± 1 polytope*, $\text{CUT}^{\pm 1}(G)$, is the convex hull of all ± 1 cut vectors of G .

Definition 50. Suspension graph (∇G) . The suspension graph ∇G of G is the graph with vertex set $V(G) \sqcup \{e\}$ and edge set $E(G) \sqcup \{(e, i), i \in V(G)\}$.

Remark 20. ∇G is the graph obtained by adding an extra vertex and joining all the vertices of G to it.

Theorem 7. $\text{CUT}^{01}(\nabla G) = \psi[\text{COR}(G)]$ where $\psi : \mathbb{R}^{|V(G)|+|E(G)|} \rightarrow \mathbb{R}^{|V(G)|+|E(G)|}$ defined by

$$q \mapsto x$$

where

$$x_{ij} = 1 - q_i - q_j + 2q_{ij}, \quad \{i, j\} \in E(G),$$

$$x_{ei} = q_i, \quad i \in V(G).$$

Remark 21. $x_{ei} = \langle j \rangle = q_j = p_j(1)$ where $\langle j \rangle$ is the expectation value corresponding to the observable j . Similarly, $x_{ij} = p_{ij}(00) + p_{ij}(11)$, which is the probability of getting the same outcome.

Claim 3. $\text{CUT}^{\pm 1}$ and CUT^{01} are related by a bijective linear map $\alpha : \text{CUT}^{01}(G) \rightarrow \text{CUT}^{\pm}(G)$ defined by

$$x \mapsto y$$

where

$$y_{ij} = 2x_{ij} - 1.$$

Remark 22. If we relabel the outcomes as $0 \rightarrow +1$ and $1 \rightarrow -1$ then we can write

$$y_{ei} = -\langle i \rangle = -[p_i(1) - p_i(-1)]$$

and

$$y_{ij} = -\langle ij \rangle$$

$$= -[p_{ij}(1, 1) + p_{ij}(-1, -1) - p_{ij}(-1, 1) - p_{ij}(1, -1)].$$

Now we discuss an example.

Example 2. Bell scenario. The Bell scenario corresponds to the case where context is generated via spacelike separation of the involved parties. We explain the defining components of the Bell scenario $\Upsilon := (X, \mathcal{C}, O)$ hereafter. We assume the number of parties to be n . The set X consists of various disjoint subsets $X_1, X_2, \dots, X_n$. The subset X_i consists of measurement operators for party i . All contexts $C \in \mathcal{C}$ are of the form $C = \{M_1, M_2, \dots, M_n\}$ where $M_i \in X_i$. The compatibility graph for the scenario is a complete n -partite graph. The Bell scenario corresponding to n parties with m measurements per party where each measurement has o outcomes is often denoted as (n, m, o) .

-
- [1] S. Kochen and E. P. Specker, in *The Logico-Algebraic Approach to Quantum Mechanics* (Springer, New York, 1975), pp. 293–328.
 - [2] S. Abramsky and A. Brandenburger, *New J. Phys.* **13**, 113036 (2011).
 - [3] J. S. Bell, *Phys. Phys. Fiz.* **1**, 195 (1964).
 - [4] A. Cabello, S. Severini, and A. Winter, *Phys. Rev. Lett.* **112**, 040401 (2014).
 - [5] A. K. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).
 - [6] R. Colbeck, *arXiv:0911.3814*.
 - [7] B. S. Tsirel'son, *J. Sov. Math.* **36**, 557 (1987).
 - [8] S. J. Summers and R. Werner, *J. Math. Phys.* **28**, 2440 (1987).
 - [9] S. Popescu and D. Rohrlich, *Phys. Lett. A* **166**, 293 (1992).
 - [10] R. Cleve and H. Buhrman, *Phys. Rev. A* **56**, 1201 (1997).
 - [11] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, *Rev. Mod. Phys.* **86**, 419 (2014).
 - [12] J. Singh, K. Bharti, and Arvind, *Phys. Rev. A* **95**, 062333 (2017).
 - [13] A. Cabello, V. D'Ambrosio, E. Nagali, and F. Sciarrino, *Phys. Rev. A* **84**, 030302(R) (2011).
 - [14] M. Um, X. Zhang, J. Zhang, Y. Wang, S. Yangchao, D.-L. Deng, L.-M. Duan, and K. Kim, *Sci. Rep.* **3**, 1627 (2013).
 - [15] K. Bharti, M. Ray, A. Varvitsiotis, N. A. Warsi, A. Cabello, and L.-C. Kwek, *Phys. Rev. Lett.* **122**, 250403 (2019).
 - [16] M. Howard, J. Wallman, V. Veitch, and J. Emerson, *Nature (London)* **510**, 351 (2014).
 - [17] R. Raussendorf, *Phys. Rev. A* **88**, 022322 (2013).
 - [18] S. Mansfield and E. Kashefi, *Phys. Rev. Lett.* **121**, 230401 (2018).
 - [19] N. Delfosse, P. A. Guerin, J. Bian, and R. Raussendorf, *Phys. Rev. X* **5**, 021003 (2015).
 - [20] H. Pashayan, J. J. Wallman, and S. D. Bartlett, *Phys. Rev. Lett.* **115**, 070501 (2015).
 - [21] J. Bermejo-Vega, N. Delfosse, D. E. Browne, C. Okay, and R. Raussendorf, *Phys. Rev. Lett.* **119**, 120505 (2017).
 - [22] L. Catani and D. E. Browne, *Phys. Rev. A* **98**, 052108 (2018).
 - [23] R. W. Spekkens, *Phys. Rev. Lett.* **101**, 020401 (2008).
 - [24] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).
 - [25] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenbergh, R. F. Vermeulen, R. N. Schouten, C. Abellán *et al.*, *Nature (London)* **526**, 682 (2015).
 - [26] It is interesting to note that noncontextual completions which don't satisfy a property known as *functional consistency* can also explain the predictions of quantum mechanics. See [27,28] for details.
 - [27] A. Peres, *Quantum Theory: Concepts and Methods*, Fundamental Theories of Physics (Springer, New York, 2006), Vol. 57.
 - [28] A. S. Arora, K. Bharti, and Arvind, *Phys. Lett. A* **383**, 833 (2018).
 - [29] P. Kurzyński, R. Ramanathan, and D. Kaszlikowski, *Phys. Rev. Lett.* **109**, 020404 (2012).
 - [30] A. Cabello, L. E. Danielsen, A. J. López-Tarrida, and J. R. Portillo, *Phys. Rev. A* **88**, 032104 (2013).
 - [31] B. Amaral and M. T. Cunha, *arXiv:1709.04812v2*.
 - [32] B. Amaral and M. T. Cunha, *On Graph Approaches to Contextuality and their Role in Quantum Theory*, SpringerBriefs in Mathematics (Springer International Publishing, 2018).
 - [33] See Sec. VII for a comparison of these two approaches.
 - [34] M. Chudnovsky, N. Robertson, P. Seymour, and R. Thomas, *Ann. Math.* **164**, 51 (2006).
 - [35] M. Araújo, M. T. Quintino, C. Budroni, M. T. Cunha, and A. Cabello, *Phys. Rev. A* **88**, 022118 (2013).
 - [36] M. Sadiq, P. Badziąg, M. Bourennane, and A. Cabello, *Phys. Rev. A* **87**, 012128 (2013).

- [37] A. A. Klyachko, M. A. Can, S. Binicioğlu, and A. S. Shumovsky, *Phys. Rev. Lett.* **101**, 020403 (2008).
- [38] With respect to the set of all edges, excluding self-loops.
- [39] M. Grötschel, L. Lovasz, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization (Algorithms and Combinatorics)* (Springer, New York, 2011), p. 273.
- [40] D. E. Knuth, *Electron. J. Comb.* **1**, 1 (1994).
- [41] I. Bengtsson, *Foundations of Probability and Physics* (Springer, New York, 2009), Vol. 5.
- [42] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, *Phys. Rep.* **506**, 1 (2011).
- [43] Related by a linear transformation.
- [44] K. Bharti, M. Ray, A. Varvitsiotis, A. Cabello, and L.-C. Kwek, [arXiv:1911.09448v1](https://arxiv.org/abs/1911.09448v1).