

Inequalities witnessing coherence, nonlocality, and contextuality

Rafael Wagner,^{1,2,*} Rui Soares Barbosa,^{1,†} and Ernesto F. Galvão^{1,3,‡}

¹*INL – International Iberian Nanotechnology Laboratory, Braga, Portugal*

²*Centro de Física, Universidade do Minho, Braga, Portugal*

³*Instituto de Física, Universidade Federal Fluminense, Niterói – RJ, Brazil*

(Dated: 24th January 2024)

Quantum coherence, nonlocality, and contextuality are key resources for quantum advantage in metrology, communication, and computation. We introduce a graph-based approach to derive classicality inequalities that bound local, noncontextual, and coherence-free models, offering a unified description of these seemingly disparate quantum resources. Our approach generalizes recently proposed basis-independent coherence witnesses, and recovers all noncontextuality inequalities of the exclusivity graph approach. Moreover, violations of certain classicality inequalities witness preparation contextuality. We describe an algorithm to find all such classicality inequalities, and use it to analyze some of the simplest scenarios.

Introduction.—Non-classical resources provided by quantum theory are key to quantum advantage for information processing [1–8]; see [9–11] for comprehensive reviews of applications. Many different nonclassical features of quantum mechanics have been identified, studied, witnessed, and quantified [4, 12–27]. It is natural to wonder to what extent different quantum resources can be characterized in a unified way. Here we address this question by proposing a single formalism that yields inequalities bounding three different notions of classicality: noncontextual, local, and coherence-free models.

A number of modern approaches to contextuality have successfully incorporated nonlocality as a special case [28–31]. The relationship between this unified notion of non-classical correlations and coherence, however, has been harder to establish. One roadblock is that most approaches to characterize coherence presuppose the choice of a fixed reference basis [11]. Recently, different approaches have been proposed to study a basis-independent notion of coherence [26, 32], dubbed *set coherence* in Ref. [26]. A recent approach, on which the present work builds, derives witnesses of basis-independent coherence using only relational information between states in the form of two-state overlaps [32]. Still, so far there has been no clear identification between non-locality and contextuality on one hand, and coherence on the other. There are examples of models that mimic quantum coherence *without* displaying contextuality or nonlocality, such as the toy models from Refs. [33, 34], while on the other hand incoherent states – even maximally mixed states – can of course be used to witness state-independent quantum contextuality [35, 36]. Theory-independent approaches have been used to compare relevant types of nonclassical resources [23, 27, 37], but an understanding of the special case of coherence and contextuality is still lacking. A better understanding of the relationship between

these two fundamental manifestations of nonclassicality has both important foundational impact and potential technological applications.

Building on the study of coherence using two-state overlaps [32], we propose a framework that associates to any (simple) graph G a probability polytope C_G of edge weightings. Vertices of the graph G represent probabilistic processes, while edges of G correspond to correlations between neighbouring processes. We show that the faces of the polytope C_G describe bounds on noncontextual, local, and coherence-free models, depending on the interpretation of vertices of the graph G as preparations and measurements. The description of three notions of classicality under a single framework represents a significant conceptual advance towards clarifying the source of quantum computational advantage.

The classical polytope C_G .—Let $G = (V(G), E(G))$ be an undirected graph, which we call the *event graph*. We consider edge weightings $r: E(G) \rightarrow [0, 1]$, which assign a weight $r_e = r_{ij}$ to each edge $e = \{i, j\}$ of G . We regard these weightings as points forming a polytope, the unit hypercube, $r \in [0, 1]^{E(G)}$. To define the *classical polytope* $C_G \subseteq [0, 1]^{E(G)}$, take each vertex $i \in V(G)$ to represent a random variable A_i with values belonging to an alphabet Λ , and suppose these are jointly distributed. This determines an edge weighting r where each weight r_{ij} is the probability that the processes corresponding to vertices i and j output equal values, i.e.

$$r_{ij} = P(A_i = A_j).$$

An edge weighting r is in the classical polytope C_G if it arises in this fashion from jointly distributed random variables $(A_i)_{i \in V(G)}$. Each weight r_{ij} is then a measure of the correlation between the output values of A_i and of A_j . In the case of dichotomic values $\Lambda = \{+1, -1\}$, this quantity is related to the expected value of the product by $\langle A_i A_j \rangle = 2r_{ij} - 1$ [38]. An (alternative) formal description of C_G is given in detail in Appendix A.

Inequalities defining C_G .—The inequalities defining the polytope C_G impose logical conditions determining the set of classical edge weightings. The existence of non-trivial facets of C_G can be illustrated with the example

* rafael.wagner@inl.int

† rui.soaresbarbosa@inl.int

‡ ernesto.galvao@inl.int

of Fig. 1–(a), the 3-vertex complete graph K_3 , with edge weights r_{12}, r_{23}, r_{13} . We cannot have e.g.

$$r_{12} = 1, \quad r_{23} = 1, \quad r_{13} = 0,$$

as this would contradict transitivity of equality on the deterministic values corresponding to each of the three vertices: $A_1 = A_2 = A_3 \neq A_1$. In Ref. [32] it was shown that the only non-trivial inequalities for the n -cycle event graph C_n are

$$-r_e + \sum_{e' \neq e} r_{e'} \leq n - 2, \quad \text{for each } e \in E(C_n). \quad (1)$$

Incidentally, these inequalities have been known at least since the work of Boole [39–41].

We now give a high-level description of an algorithm to completely characterize C_G for general event graphs G . We start by enumerating the vertices of C_G . These are all the ‘deterministic’ labellings of the edges of G with values in $\{0, 1\}$ that are logically consistent with transitivity of equality. The facets of C_G can then be found using standard convex geometry tools [42].

Whether a given deterministic edge labelling is consistent – and therefore a vertex of C_G – can be checked in linear time on the size of G by a graph traversal. However, it is unnecessary to generate all $2^{|E(G)|}$ -many labellings and discard the inconsistent ones. Instead, one can directly generate only the consistent ones by searching through underlying value assignments to the vertices of G . Despite being much more efficient for most graphs, this also quickly becomes unavoidably intractable due to the exponentially-increasing number of vertices of the polytope C_G . We deepen this discussion in Appendix A.

Using the method just outlined, we find all facets of C_G for some small graphs, including all graphs shown in Fig. 1. Interestingly, already for K_4 (Fig. 1–(c)) a new type of facet appears which is different from the cycle inequalities in Eq. (1). These new facets of K_4 are described by the inequalities of the form

$$(r_{12} + r_{13} + r_{14}) - (r_{23} + r_{34} + r_{24}) \leq 1, \quad (2)$$

(and others obtained by label permutations).

In Appendix B, we prove that some constructions of graphs by combining smaller graphs do not give rise to new facet inequalities, trimming the class of graphs worth analyzing. In Appendix C, we list all facet inequalities of the classical polytopes for the complete graphs K_4 , K_5 , and K_6 . We also give numerically-found examples of quantum violations – witnessing basis-independent coherence in the sense described in the next section – of all non-trivial facets of K_4 and K_5 . All the new inequalities and quantum violations found, together with the code used to obtain them, which is applicable to analyze C_G for an arbitrary graph G , are made available in an associated Git repository [43]. In Appendix D, we generalize the inequalities of Eq. (2) to complete graphs of $n \geq 4$ vertices, and prove that these define facets of the classical polytopes

C_{K_n} for all such n . This yields an infinite family h_n of new classicality inequalities not previously described in the literature. The first three new inequalities from this family (h_4, h_5, h_6) have recently been experimentally violated, serving to benchmark quantum photonic devices [44].

We now proceed to describe how the inequalities obtained for the abstract scenarios considered above establish bounds both on coherence-free models and on noncontextual/local models. Each type of operational scenario suggests an interpretation for edge weights, and naturally imposes further constraints on them, resulting in cross-sections of the polytope C_G . These cross-sections recover known noncontextuality/locality polytopes, as well as basis-independent coherence witnesses.

C_G bounds coherence-free models.—Most commonly, coherence is defined for a quantum state with respect to a fixed basis, as the presence of nonzero off-diagonal elements in its density matrix (in that basis) [45, 46]. Recently, Refs. [26, 32] proposed a *basis-independent* notion of coherence as a property of a *set of states*: this is said to be *coherent* when the states in the set are not simultaneously diagonalizable, i.e. when there is no basis in which all their density matrices are diagonal, or equivalently, if the states in the set do not pairwise commute. Otherwise, the set is said to be *coherence-free*, or *incoherent*.

In Ref. [32], basis-independent coherence witnesses were described using only pairwise overlaps $r_{ij} = \text{Tr}(\rho_i \rho_j)$ among a set of quantum states, focusing on witnesses provided by violations of the cycle inequalities in Eq. (1). We explain the interpretation of the facet inequalities of C_G as basis-independent coherence witnesses, generalizing the results of Ref. [32] to *any* event graph G .

Let G be any graph with n vertices. Consider a general separable state of n quantum systems of the same type (e.g. qudits), each associated to a vertex of the graph. Each edge of G is given a weight equal to the overlap between the two states of its incident vertices. These overlaps can be estimated using the well-known SWAP test [47]. In Ref. [32] it was shown that the facet-inequalities of C_G describe necessary conditions on the set of overlaps, i.e. on edge weightings of G , for the set of single-system states to be coherence-free, that is, all of them diagonal in a common single-system basis. This is so because for such a coherence-free set of states, the overlap r_{ij} equals the probability of obtaining equal outcomes in independent measurements of the states associated to vertices i and j using the observable that projects onto the reference basis.

C_G bounds local and noncontextual models.—The faces of C_G can also be understood as bounds on noncontextual models [48, 49]. A simple first approach consists in having vertices of G represent measurements, while edges identify two-measurement contexts, i.e. pairs of observables that can be measured simultaneously. The weight of an edge corresponds to the probability, with respect to a given global state, that the two incident measurements yield equal outcomes. A necessary and sufficient condition for

the existence of a noncontextual model whose behaviour is consistent with a given edge weighting is the existence of a global probability distribution (on outcome assignments to all measurements) whose marginals recover the correct outcome probabilities. This is the content of the Fine–Abramsky–Brandenburger theorem [29, 50, 51].

Such a global distribution, when it exists, can also be interpreted as a classical coherence-free model. This dual role of global probability distributions is the link connecting coherence-free models and noncontextual models, and allowing violations of facet inequalities of C_G to witness either property, depending on the interpretation of the scenario at hand.

In general, this simple approach, interpreting vertices as measurements and edges as equality of outcome in two-measurement contexts, is not sufficient to capture contextuality in full generality [31, Section 2.5.3]. Even restricting to contextuality scenarios whose maximal contexts have size two, the facets of C_G are not necessarily facet, or even tight, noncontextuality inequalities, except in the case of dichotomic measurements [52, Theorem 38], where equality of outcomes fully determines the measurement statistics. An important example is the Clauser–Horne–Shimony–Holt (CHSH) inequality.

Encoding some contextuality scenarios requires the imposition of further constraints, which geometrically determine cross-sections on the classical polytope C_G . These constraints may, for example, represent operational symmetries of the measurement scenario, e.g. making two vertices equal, or may encode given conditions on the compatibility of observables. One example is the exclusivity constraint present in the Cabello–Severini–Winter (CSW) graph approach [28].

We now show how both CHSH and the original 3-setting Bell inequality can be obtained from cycle inequalities, before describing a more systematic approach that recovers all noncontextuality inequalities obtainable from the exclusivity graph approach [28, 31].

We remark that we treat Bell nonlocality as an instance of contextuality, in which measurement compatibility is ensured by space-like separation between various parties who locally measure a shared multipartite system. This view of nonlocality as a special case of contextuality is well established, e.g., in Refs. [29, 30], although there are important subtle differences when considering free transformations in a resource-theoretic setup [53].

Example: CHSH inequality from the 4-cycle graph C_4 .—It is easy to check from Eq. (1) that the 4-cycle graph C_4 with edges $r_{12}, r_{23}, r_{34}, r_{14}$ (see Fig. 1–(b)) has 4 nontrivial facets given by the inequality

$$r_{12} + r_{23} + r_{34} - r_{14} \leq 2, \quad (3)$$

and label permutations thereof. We translate this into the CHSH [54] Bell scenario, with Alice locally measuring one of two rank-1 projectors A_1 or A_2 , and Bob locally measuring either B_1 or B_2 , on the singlet state $|\psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$. As a contextuality scenario, the CHSH graph C_4 is a graph with no clique with more than

two vertices, and the only non-trivial noncontextuality inequality is given in terms of correlations. From the event graph perspective, each vertex can be understood as a two-outcome measurement at either Alice or Bob. It is easy to check that the overlap between two single-qubit projectors A, B is the probability of obtaining different outcomes [55] when measuring those projectors on each part of the singlet state: $r_{AB} = p_{\neq}^{AB} = 1 - p_{=}^{AB}$. Using this interpretation, the facet of C_{C_4} given by Eq. (3) can be rewritten as

$$p_{\neq}^{A_1 B_1} + p_{\neq}^{A_2 B_1} + p_{\neq}^{A_2 B_2} - p_{\neq}^{A_1 B_2} \leq 2, \quad (4)$$

which is a well-known way to write the CHSH inequality [56]. This same procedure can be used to obtain chained Bell inequalities [57, 58] from cycle inequalities.

Example: Original Bell inequality from the 3-cycle graph C_3 .—If on the C_4 graph we have just analyzed we impose the constraint that one of the edge weights equal 1, we recover the non-trivial facets for the 3-cycle C_3 , namely $r_{12} + r_{23} - r_{13} \leq 1$ and label permutations. The embedded tetrahedron with these 3 facets delimits the local correlations in the original two-party Bell inequality [49], featuring three settings at each party, and assuming perfect anticorrelation for pairs of aligned settings. For a geometrical description of the ellipope of quantum correlations, see Ref. [59].

Example: CHSH inequality from the 5-vertex wheel graph W_5 .—An alternative way of interpreting an event graph as a contextuality scenario involves having a single vertex, the handle, represent a quantum state, and all the others represent measurement operators. Take the 5-vertex wheel graph W_5 of Fig. 1–(d) as an instructive example. A simple calculation shows that if we impose $r_{12} = r_{34}$ and $r_{23} = r_{14}$, then adding together four 3-cycle inequalities for this graph recovers the CHSH inequality in the form of Eq. (4). The quantum realization of this graph scenario has the central vertex 5 representing a singlet state, with the other vertices representing the four projectors measured jointly by Alice and Bob. The imposed constraints reflect the fact that opposing edges represent the same quantity, the overlap between the two projectors locally measured by one of the parties.

Recovering all noncontextuality inequalities of the exclusivity graph formalism.—The second approach to obtaining the CHSH inequality does not rely on particular properties of the singlet state. The use of a handle vertex to represent a state can be generalized to other scenarios, as we now describe.

In the *exclusivity graph approach* to contextuality one considers a graph H whose vertices represent measurement events[60] (in a quantum realization, projection operators), and edges connect mutually exclusive events (in the quantum setting, orthogonal projectors). In this formalism, the noncontextual behaviours are described by a well-known construction, the stable polytope of the graph H , denoted $\text{STAB}(H)$ [31]. This is reviewed in detail in Appendix E. In brief, the vertices or extreme points of the polytope $\text{STAB}(H)$ are (the characteristic

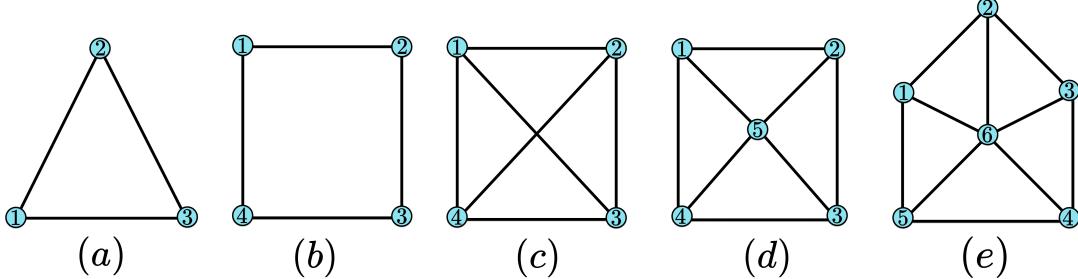


Figure 1. **Event graphs corresponding to bounds on classical models.** Each of these graphs can be used to obtain the following nonclassicality inequalities: (a) constrained CHSH inequality; (b),(d) CHSH Bell locality inequality; (c) new K_4 classicality inequality from Eq. (2), and (e) Klyachko, Can, Binicioğlu, and Schumovsky (KCBS) noncontextuality inequality.

functions of) subsets of $V(H)$ that do not contain any pair of adjacent vertices. More intuitively, perhaps, they correspond to truth-value assignments to the measurement events, i.e. functions $V(H) \rightarrow \{0, 1\}$, such that no two exclusive events are deemed true, i.e. no two adjacent vertices are assigned the value 1.

We can understand this setup in terms of our formalism as follows. We define an event graph H_* obtained from the exclusivity graph H by adding a new vertex connected to all other vertices. This new vertex is used to represent a handle state ψ . Formally, H_* is given by $V(H_*) := V(H) \sqcup \{\psi\}$ and $E(H_*) := E(H) \cup \{\{v, \psi\} \mid v \in V(H)\}$. The structure of the exclusivity graph H is then used to force a constraint on edge weightings of H_* , namely that all edges already present in H be assigned zero weight. The resulting cross-section $C_{H_*}^0 := \{r \in C_{H_*} \mid \forall e \in E(H). r_e = 0\}$ of the polytope C_{H_*} , which moreover is a subpolytope, then carries information about the noncontextual behaviours in $\text{STAB}(H)$. Formally, in Appendix E, we exhibit an isomorphism between the polytopes $\text{STAB}(H)$ and $C_{H_*}^0$ for any exclusivity graph H . As a consequence, we show that the *facet-defining noncontextuality inequalities bounding noncontextual behaviours for H are precisely the facet-defining inequalities of $C_{H_*}^0$* . Moreover, these inequalities can be obtained from the inequalities defining facets of the whole classical polytope C_{H_*} by removing (i.e. setting to zero) the variables r_e with $e \in E(H)$.

Example: KCBS noncontextuality inequality.—We illustrate this mapping between formalisms with the noncontextuality inequality obtained by Klyachko, Can, Binicioğlu, and Schumovsky (KCBS) [61], and expressed in the CSW formalism in Ref. [28].

Starting with the 5-cycle graph $H = C_5$ interpreted as an exclusivity graph, then H_* is the 6-vertex wheel graph W_6 of Fig. 1-(e). The central vertex represents a quantum state, while neighbouring vertices in the outer 5-cycle represent mutually exclusive measurement events (quantum mechanically: orthogonal projectors) so as to impose $r_{vw} = 0$ for neighbouring v and w in this outer subgraph. The KCBS noncontextuality inequality is a bound on weightings of the edges connected to the central

vertex:

$$\sum_{v=1}^5 r_{v6} \leq 2. \quad (5)$$

Note that each edge weight r_{v6} in Eq. (5) is the probability of successful projection of the central vertex state onto the projector associated with vertex v .

In our framework, this inequality is obtained from a facet-defining inequality of C_{W_6} ,

$$-r_{12} - r_{23} - r_{34} - r_{45} - r_{15} + r_{16} + r_{26} + r_{36} + r_{46} + r_{56} \leq 2,$$

by imposing the exclusivity (or orthogonality) condition of null edge weights on the 5-cycle outer subgraph.

Cycle inequalities witness preparation contextuality.—Besides considering different approaches to Kochen–Specker noncontextuality, one can also consider different *notions* of noncontextuality. One such proposal, put forth by Spekkens in Ref. [62], is that of preparation (generalized) noncontextuality [7, 63–66]. We consider once more a quantum realization of the event graph representing vertices as states and edges as two-state overlaps. In Appendix F we prove that *violations of the inequalities for the classical polytope of the cycle event graph C_n are witnesses of preparation contextuality*. This result is shown for a class of prepare-and-measure operational scenarios [63, 64], which includes quantum theory viewed as an operational theory. In contrast to quantum theory, the well-known noncontextual toy theory of Ref. [33] does not violate these event graph inequalities, if vertices of the event graph are taken to represent toy theory states.

Discussion and future directions.—We proposed a new graph-theoretic approach that unifies the study of three different quantum resources, namely contextuality, nonlocality, and coherence. Non-classicality inequalities are obtained as facets of a polytope C_G of edge weightings associated with an *event graph* G , with suitable constraints that depend on the chosen interpretation of vertices as quantum states or measurements, as required by each scenario.

Connections with the theory of contextuality were presented with respect to different approaches and definitions. In particular, we recovered all inequalities of the

CSW exclusivity graph approach [28], and we explicitly derived CHSH and KCBS inequalities as examples. We also showed that for cycle graphs the classical polytope bounds Spekkens preparation noncontextuality.

It would be interesting to understand whether these results can be made more robust. In particular, we observed that the noncontextuality inequalities for exclusivity graphs H are obtained from the inequalities of a classical polytope $C_{H,*}$ by assigning weight zero to some edges. But many of these inequalities of $C_{H,*}$ allow for deviations from such null weights without leaving the classical polytope $C_{H,*}$. This suggests that perhaps those inequalities could still be interpreted as a robust form of noncontextuality inequalities, where exclusivity is relaxed.

Future research directions include characterizing this framework in the landscape of general probabilistic theories (GPTs) and understanding how this approach bounds relational unitary invariants involving three or more states, such as Bargmann invariants [67]. It would also be interesting to relate violation of our inequalities with advantage in quantum protocols, as recently done by some of us in [68] for the task of quantum interrogation.

Acknowledgements.—We would like to thank Marcelo Terra Cunha, John Selby, David Schmid, and Raman Choudhary for helpful discussions. We also thank Roberto D. Baldijão for critically reviewing an early version of this work.

We acknowledge financial support from FCT – Fundação para a Ciência e a Tecnologia (Portugal) through PhD Grant SFRH/BD/151199/2021 (RW) and through CEECINST/00062/2018 (RSB and EFG). This work was supported by the ERC Advanced Grant QU-BOSS, GA no. 884676.

Appendix A: Characterizing the classical polytope

In Ref. [32], Galvão and Brod derived the facet-defining inequalities of the classical polytope C_{C_n} of the n -cycle event graph C_n , as discussed in the text. The construction uses an argument based on Boole's inequalities for logically consistent processes [39]. In the main text we discuss that, in fact, *any* event graph, and not only cycle graphs, can be used to bound classicality of different forms.

In this section, we consider the computational problem of characterizing the classical polytope C_G for any event graph G . We propose a simple algorithm for computing all its vertices and facets. This proceeds by first calculating the list of vertices of C_G , i.e. its V-representation, and then finding its facet-defining inequalities, i.e. its H-representation, using standard convex geometry tools. As discussed in the main text, this last step is computationally efficient on the size of the polytope. However, the overall efficiency of the procedure is intrinsically limited by the fact that the number of vertices and facets of C_G grows exponentially on the size of G . The brunt of this section is dedicated to computing the set of vertices of C_G .

After setting out the formal definitions, we characterize the edge $\{0, 1\}$ -labellings $E(G) \rightarrow \{0, 1\}$ that respect logical consistency conditions and thus correspond to the vertices of C_G . This characterization yields an efficient procedure for checking whether such an edge labelling is a vertex of C_G , whose complexity we analyze.

However, when the goal is to generate all vertices of C_G , it is needlessly wasteful to generate all the $2^{|E(G)|}$ -many edge $\{0, 1\}$ -labellings and then filter them one by one. Instead, we present a procedure that generates the edge labellings that are vertices of C_G by generating vertex labellings underlying them, thus limiting the search through the space $\{0, 1\}^{\bar{E}(G)}$ of edge labellings. Even though it might output the same vertex more than once, the method works well, especially for dense graphs. It is optimal for the complete graphs K_n , which as we will see in Appendix C are our main examples of interest. We observe that the number of vertices of the polytope C_{K_n} is given by a well-known combinatorial sequence, known as the Bell numbers [69], which count the number of partitions of a set, precisely the space that is searched by this procedure. Finally, we discuss an alternative method that might be more efficient for sparse graphs.

Basic definitions.—We start with the relevant definitions.

Definition 1 (Graph). A *graph* $G = (V(G), E(G))$ consists of a finite set $V(G)$ of vertices and a set $E(G)$ of edges, which are two-element subsets of $V(G)$, i.e. sets of the form $\{v, w\}$ where $v, w \in V(G)$ are distinct vertices.

Note that the graphs we consider in this text are so-called *simple* graphs: they are undirected (since $\{v, w\} = \{w, v\}$), have at most one edge between any two vertices v and w , and have no loops (i.e. have no edges from a vertex to itself). In one well-delimited passage, however, we will need to consider *possibly loopy graphs*, which may have loops. This corresponds to dropping the requirement that v and w be distinct in the definition above. A possibly loopy graph is said to be *loop-free* if it has no loops, i.e. if it is a bona fide (simple) graph.

Definition 2 (Labellings and colouring). A *vertex labelling* by a set Λ , or a *vertex Λ -labelling* for short, is a function $\lambda: V(G) \rightarrow \Lambda$ assigning to each vertex a label from Λ . It is called a *colouring* if $\{v, w\} \in E(G)$ implies $\lambda(v) \neq \lambda(w)$. The graph G is said to be k -colourable for $k \in \mathbb{N}$ when it admits a colouring by a set of size k .

Similarly, an *edge labelling* by a set Λ , or an *edge Λ -labelling* for short, is a function $\alpha: E(G) \rightarrow \Lambda$ assigning a label from Λ to each edge. When $\Lambda = [0, 1]$, we call this an *edge weighting*.

Definition 3 (Chromatic number). The *chromatic number* of a graph G , written $\chi(G)$, is the smallest $k \in \mathbb{N}$ such that G is k -colourable.

In the classical, deterministic situation modelled by our framework, we consider a vertex labelling of a graph G by an arbitrary labelling set Λ . However, operationally, we

do not have access to the vertex labels, but only to the information of whether the labels of neighbouring edges are equal or different.

Definition 4. Given any vertex labelling $\lambda: V(G) \rightarrow \Lambda$, its *equality labelling* ϵ_λ is the edge $\{0, 1\}$ -labelling given by:

$$\epsilon_\lambda: E(G) \rightarrow \{0, 1\}$$

$$\epsilon_\lambda \{v, w\} := \delta_{\lambda(v), \lambda(w)} = \begin{cases} 1 & \text{if } \lambda(v) = \lambda(w) \\ 0 & \text{if } \lambda(v) \neq \lambda(w) \end{cases}.$$

We are interested in characterizing the edge $\{0, 1\}$ -labellings that arise as equality labellings of vertex labellings.

Definition 5. An edge $\{0, 1\}$ -labelling $\alpha: E(G) \rightarrow \{0, 1\}$ is said to be Λ -realizable if it is the equality labelling of some vertex Λ -labelling, i.e. if $\alpha = \epsilon_\lambda$ for some $\lambda: V(G) \rightarrow \Lambda$. If Λ has size $k \in \mathbb{N}$, we say that α is k -realizable.

We write $\text{Eq}(G)$ for the set of realizable edge labellings of G (with any Λ), and $\text{Eq}_k(G)$ for the set of k -realizable ones. We have that $\text{Eq}_k(G) \subseteq \text{Eq}_{k'}(G)$ whenever $k \leq k'$, and $\text{Eq}(G) = \cup_{k \in \mathbb{N}} \text{Eq}_k(G)$. Moreover, $\text{Eq}(G) = \text{Eq}_{|V(G)|}(G)$ because a vertex labelling uses at most one distinct label per vertex of the graph.

We often refer to these realisable edge $\{0, 1\}$ -labellings as the *classical* edge labellings. By the inclusion $\{0, 1\} \subseteq [0, 1]$, we can think of any edge $\{0, 1\}$ -labelling as a (deterministic) edge weighting. This gives an alternative description of the classical polytope C_G in the main text.

Definition 6. Given a graph G , its *classical polytope* $C_G \subseteq [0, 1]^{|E(G)|}$ is the convex hull of the set $\text{Eq}(G)$ seen as a set of points in $[0, 1]^{|E(G)|}$.

Characterizing the vertices of C_G .—We now consider the question of determining whether a given edge $\{0, 1\}$ -labelling is realizable (as the equality labelling of some vertex labelling).

Given $\alpha: E(G) \rightarrow \{0, 1\}$, define a relation \sim_α on the set of vertices of G whereby $v \sim_\alpha w$ if and only if there is a path from v to w through edges labelled by 1, i.e. there is a sequence $e_1, \dots, e_n \in E(G)$ such that $v \in e_1, w \in e_n, e_i \cap e_{i+1} \neq \emptyset$, and $\alpha(e_i) = 1$. This is easily seen to be an equivalence relation.

It yields the following characterization of the classical edge labellings.

Proposition 7. An edge labelling $\alpha: E(G) \rightarrow \{0, 1\}$ is realizable (i.e. classical) if and only if for all edges $\{v, w\} \in E(G)$, $v \sim_\alpha w$ implies $\alpha(\{v, w\}) = 1$.

In other words, an edge labelling $\alpha: E(G) \rightarrow \{0, 1\}$ fails to be realizable precisely when there is an edge $\{v, w\} \in E(G)$ such that $v \sim_\alpha w$ and $\alpha(\{v, w\}) = 0$. In terms of the underlying vertex labellings, such a situation would violate the transitivity of equality.

A slightly different perspective is given by using α to construct a new graph that ‘collapses’ G through paths labelled by 1. Note that this construction yields a possibly loopy graph.

An edge $\{0, 1\}$ -labelling α partitions the edges of G into two sets. This determines two graphs $G_{\alpha=0}$ and $G_{\alpha=1}$, both with the same vertex set as G , but each retaining only the edges of G with the corresponding label, i.e. for each $\lambda \in \{0, 1\}$,

$$\begin{aligned} V(G_{\alpha=\lambda}) &:= V(G) \\ E(G_{\alpha=\lambda}) &:= \{e \in E(G) \mid \alpha(e) = \lambda\} \end{aligned}$$

A possibly loopy graph G/α is then defined as follows:

- its vertices are connected components of $G_{\alpha=1}$, or equivalently, the equivalence classes of \sim_α ;
- there is an edge between two connected components A and B of $G_{\alpha=1}$ whenever there exist vertices $v \in A, w \in B$, such that $\{v, w\} \in E(G_{\alpha=0})$.

Lemma 8. Let $\alpha: E(G) \rightarrow \{0, 1\}$ and Λ be any set. There is a one-to-one correspondence between Λ -realizations of α and Λ -colourings of G/α .

Proof. Let $\lambda: V(G) \rightarrow \Lambda$ such that $\alpha = \epsilon_\lambda$. If $v \sim_\alpha w$, then $\lambda(v) = \lambda(w)$, by propagating equality along the path labelled by 1. Hence, the map $\kappa: V(G/\alpha) \rightarrow \Lambda$ given by $\kappa([v]) := \lambda(v)$ is well defined. Now, an edge $e \in E_{G/\alpha}$ is of the form $e = \{[v], [w]\}$ for some $v, w \in V(G)$ such that $\alpha(\{v, w\}) = 0$. Since $\alpha = \epsilon_\lambda$, this means that $\lambda(v) \neq \lambda(w)$, hence $\kappa([v]) \neq \kappa([w])$. Thus, κ is a colouring.

Conversely, given a colouring $\kappa: V_{G/\alpha} \rightarrow \Lambda$, set $\lambda(v) := \kappa([v])$. Let $e = \{v, w\} \in E(G)$. If $\alpha(e) = 1$, then $[v] = [w]$, hence $\lambda(v) = \lambda(w)$ because κ is a colouring. If $\alpha(e) = 0$, then $\{[v], [w]\} \in E_{G/\alpha}$, hence $\lambda(v) \neq \lambda(w)$. In either case, $\alpha(e) = \epsilon_\lambda(e)$.

The two processes just described are inverses of one another. \square

Corollary 9. An edge $\{0, 1\}$ -labelling is Λ -realizable if and only if the possibly loopy graph G/α is Λ -colourable. In particular, it is realizable (i.e. classical) if and only if G/α is loop-free.

Proposition 10. Checking whether an edge $\{0, 1\}$ -labelling for a graph G is realizable can be done in time $O(n + m)$ where $n = |V(G)|$ and $m = |E(G)|$. Checking k -realizability in a given $k \geq 3$ is NP-complete.

Proof. For the first part, transverse the graph $G_{\alpha=1}$ using a depth-first search (DFS). When visiting each vertex, run through all the departing edges of $G_{\alpha=0}$ to see if any is linked to an already visited vertex in the connected component of $G_{\alpha=1}$ currently being traversed. If any is found, reject α .

For the second part, use corollary 9 to reduce to graph colouring: a graph G is k -colourable if and only if the constant 0 edge labelling is realizable. \square

The procedure outlined in the proof above is described below in more detail using pseudo-code.

Input: graph G with $V(G) = \{1, \dots, N\}$.
edge-labelling $\alpha: E(G) \rightarrow \{0, 1\}$

Output: whether α is realizable, hence a vertex of the polytope C_G .

```

global variable  $d_i$  for each  $i \in V(G)$ 
global variable  $c_i$  for each  $i \in V(G)$ 

procedure MAIN()
   $d_i \leftarrow \text{false}$  for all  $i \in V(G)$ 
  for  $i \in V(G)$  do
    if  $\neg d_i$  then
       $c_j \leftarrow \text{false}$  for all  $j \in V(G)$ 
      SEARCH ( $i$ )
    end if
  end for
  terminate with output true

procedure SEARCH( $i$ )
   $d_i, c_i \leftarrow \text{true}$ 
  for  $j \in \text{NEIGHBOURS } (i)$  do
    if  $\alpha(\{i, j\}) = 0 \wedge c_j$  then
      terminate with output false
    else if  $\alpha(\{i, j\}) = 1 \wedge \neg d_j$  then
      SEARCH ( $j$ )
    end if
  end for

```

Computing all the vertices of C_G .—We conclude that it is computationally easy to check whether a given edge $\{0, 1\}$ -labelling, i.e. a given deterministic edge weighting, is classical. Nevertheless, determining the whole set of vertices of the classical polytope is computationally hard since the number of edge labelling to be tested grows exponentially with the number of edges of the graph.

It is interesting to note that for the complete event graph K_n of n vertices the number of classical edge labellings, i.e. vertices of the classical polytope C_{K_n} , is given by a well-known sequence, the Bell or exponential numbers [69, 70]. The n -th Bell number is the number of partitions, or equivalence relations, of a set of size n . It is clear that edge $\{0, 1\}$ -labellings of K_n are in one-to-one correspondence with symmetric reflexive relations on the set of vertices $\{1, \dots, n\}$, where the label of an edge $\{v, w\}$ determines whether the pairs (v, w) and (w, v) are in the relation. Among these, the classical edge labellings correspond to the equivalence relations (which additionally satisfy transitivity), with the underlying vertex labelling determining a partition of the vertices. For a general graph G , it is still true that the classical edge labellings arise from partitions, or equivalence relations, on the set of vertices, determined by the underlying vertex labelling. The difference is that an edge labelling does not carry enough information to characterize a relation fully. So, in particular, different vertex partitions may give rise to the same classical edge labelling.

We can use this observation to propose a different method for generating all vertices of C_G by constructing vertex-labellings of G . The procedure is given below in pseudo-code.

Input: graph G with $V(G) = \{1, \dots, N\}$.
Output: vertices of the polytope C_G .

```

global variable  $\lambda_i$  for each  $i \in V(G)$ 
global variable  $\alpha_e$  for each  $e \in E(G)$ 

procedure MAIN()
  GENERATE (1, 1)
end procedure

procedure GENERATE ( $i, next$ )
  if  $i = N + 1$  then
    output  $(\alpha_e)_{e \in E(G)}$ 
  else
    for  $x < next$  do
      UPDATE ( $i, x$ )
      GENERATE ( $i + 1, next$ )
    end for
    UPDATE ( $i, next$ )
    GENERATE ( $i + 1, next + 1$ )
  end if
end procedure

procedure UPDATE ( $i, x$ )
   $\lambda_i \leftarrow x$ 
  for  $j < i$  with  $\{i, j\} \in E(G)$  do
     $\alpha_{\{i, j\}} \leftarrow$  if  $\lambda_j = x$  then 1 else 0
  end for
end procedure

```

The procedure above has the disadvantage that it might output the same vertex of the polytope multiple times. This is because, as already discussed, different partitions of the vertices of G can give rise to the same edge labelling. The problem is especially noticeable for sparse graphs.

An alternative method for generating the vertices of C_G , which might be more efficient in the case of sparser graphs, is to directly search through $\{0, 1\}^{E(G)}$ while checking for consistency on the fly, in order to trim the search space so that only the realizable edge labellings are constructed. This can be done by keeping a representation of the current vertex partition (induced by the edges labelled 1 in the edge labelling being constructed), for example using a union-find data structure, together with a record of forbidden merges between partition components (induced by the edges labelled 0s in the edge labelling being constructed). The disadvantage is that the upkeep of this representation, necessary for checking consistency on the fly, cannot be done in constant time. This incurs an overhead at each step in the search.

Appendix B: Characterizing classical polytopes by graph decompositions

In this section, we prove some general facts that relate the classical polytopes of different graphs. In particular, we show that some methods of combining graphs to build larger graphs do not give rise to new classicality inequalities. Or, seen analytically rather than synthetically, that some graphs G can be decomposed into smaller

component graphs in a way that reduces the question of characterizing C_G to that of characterizing the polytopes of these components. These observations help trim down the class of graphs that is worth analyzing in the search for new classicality inequalities. As a by-product, we characterize the class of graphs for which all edge weightings are classical as being that of trees, an analogue of Vorob'ev's theorem [71] in this framework.

Proposition 11. *Let G_1 and G_2 be graphs, and write $G_1 + G_2$ for their disjoint union. Then*

$$C_{G_1+G_2} = C_{G_1} \times C_{G_2} = \{(r_1, r_2) \mid r_1 \in G_1, r_2 \in G_2\}.$$

Proof. Given vertex labellings $\lambda_i: V(G_i) \rightarrow \Lambda_i$ for each $i = 1, 2$, one obtains a function

$$\lambda_1 + \lambda_2: V(G_1) \sqcup V(G_2) \rightarrow \Lambda_1 \sqcup \Lambda_2$$

which is a vertex labelling of $G_1 + G_2$ since $V(G_1 + G_2) = V(G_1) \sqcup V(G_2)$. The corresponding equality edge labelling, $\epsilon_{\lambda_1 + \lambda_2}: E(G_1 + G_2) \rightarrow \{0, 1\}$, is precisely the function

$$[\epsilon_{\lambda_1}, \epsilon_{\lambda_2}]: E(G_1) \sqcup E(G_2) \rightarrow \{0, 1\}$$

given by

$$e \mapsto \begin{cases} \epsilon_{\lambda_1}(e) & \text{if } e \in E(G_1) \\ \epsilon_{\lambda_2}(e) & \text{if } e \in E(G_2) \end{cases},$$

implying the result. \square

In particular, vertices of the polytope $C_{G_1+G_2}$ are in bijective correspondence with pairs consisting of one vertex from each of the polytopes C_{G_i} , while the facets of $C_{G_1+G_2}$ are in bijective correspondence with the union of the facets of C_{G_1} and the facets of C_{G_2} . That is, the inequalities defining $C_{G_1+G_2}$ are those defining C_{G_1} plus those defining C_{G_2} . Taking the disjoint union of event graphs thus creates no new classicality inequalities. As a consequence, we might as well focus solely on studying the classical polytopes of connected graphs.

The result above considers the construction of a new graph by placing two graphs side by side. But similar results can be obtained for more complicated ways of combining graphs, namely gluing along a vertex or along an edge.

Definition 12 (Gluing). Given graphs G_1 and G_2 , and tuples of vertices

$$\begin{aligned} \mathbf{v}_1 &= (v_1^1, \dots, v_1^k) \in V(G_1)^k, \\ \mathbf{v}_2 &= (v_2^1, \dots, v_2^k) \in V(G_2)^k, \end{aligned}$$

the *gluing of G_1 and G_2 along \mathbf{v}_1 and \mathbf{v}_2* , written $G_1 +_{\mathbf{v}_1=\mathbf{v}_2} G_2$, is the graph obtained by taking the disjoint union $G_1 + G_2$ and identifying the vertices v_1^j and v_2^j for $j = 1, \dots, k$. Explicitly: its vertices are

$$V(G_1 +_{\mathbf{v}_1=\mathbf{v}_2} G_2) := O_1 \sqcup O_2 \sqcup N,$$

where $O_i := V(G_i) \setminus \{v_i^1, \dots, v_i^k\}$ is the set of vertices of G_i not being identified and $N = \{v^1, \dots, v^k\}$ is a set of ‘new’ vertices (i.e. not in either G_i); its edges are

$$E(G_1 +_{\mathbf{v}_1=\mathbf{v}_2} G_2) := E_1 \sqcup E_2,$$

where E_i is equal to $E(G_i)$ but with occurrences of v_i^j replaced by the new v^j .

Proposition 13. *Let G_1 and G_2 be graphs, $v_1 \in V(G_1)$ and $v_2 \in V(G_2)$, then $C_{G_1+v_1=v_2} G_2 = C_{G_1} \times C_{G_2}$.*

Proof. We proceed as in the proof of proposition 11, using the same notation, but then take a quotient of the merged alphabet $\Lambda_1 \sqcup \Lambda_2$ identifying two labels, one from each component: $\lambda_1(v_1) \in \Lambda_1$ with $\lambda_2(v_2) \in \Lambda_2$. This yields a well-defined labelling for $G_1 +_{v_1=v_2} G_2$ where the new vertex v is labelled by the element resulting from this identification. This does not affect the equality edge-labellings, and so we obtain the same result. \square

Read analytically, if G is a graph with a cut vertex V , i.e. a vertex whose removal disconnects the graph into two components with vertex sets V_1 and V_2 , then its polytope can be characterized in terms of the polytopes of the induced subgraph on $V_1 \cup \{v\}$ and $V_2 \cup \{v\}$. In particular, the facet-defining inequalities of C_G are those of each of these two components.

As an aside, this result is the missing ingredient for fully characterizing the event graphs that cannot display any nonclassicality, i.e. for which all edge weightings $E(G) \rightarrow [0, 1]$ are classical. This could be seen as an analogue of Vorob'ev's [71] theorem in our framework.

Corollary 14. *A graph G is such that $C_G = [0, 1]^{E(G)}$ if and only if it is a tree.*

Proof. For ‘only if’ part, if G has a cycle then any edge labelling $E(G) \rightarrow \{0, 1\}$ that restricts to $(1, \dots, 1, 0)$ on said cycle is not in C_G . For the ‘if’ part, apply proposition 13 multiple times, following the construction of a tree as a sequence of gluings along a vertex of copies of K_2 , whose classical polytope is $[0, 1]$. \square

We now move to consider gluing along an edge.

Proposition 15. *Let G_1 and G_2 be graphs, $v_1, w_1 \in V(G_1)$ and $v_2, w_2 \in V(G_2)$ such that $e_i := \{v_i, w_i\} \in E(G_i)$. Writing*

$$G := G_1 +_{(v_1, w_1)=(v_2, w_2)} G_2,$$

we have

$$\begin{aligned} C_G &= \left\{ r \in [0, 1]^{E(G)} \mid r|_{E(G_1)} \in C_{G_1}, r|_{E(G_2)} \in C_{G_2} \right\} \\ &\cong \{(r, s) \mid r \in C_{G_1}, s \in C_{G_2}, r_{e_1} = s_{e_2}\} \\ &\cong (C_{G_1} \times [0, 1]^{E(G_2) \setminus \{e_2\}}) \cap ([0, 1]^{E(G_2) \setminus \{e_1\}} \times C_{G_2}), \end{aligned}$$

where for the last line we assume that C_{G_1} is written with e_1 as its last coordinate and C_{G_2} with e_2 as its first coordinate.

Table I. Quantum violations for facet inequalities of C_{K_5}

Class	Violation	Inequality Representative for the Class	Dimension
11–40	1/4	$-r_{12} + r_{15} + r_{25} \leq 1$	2
41–60	1/3	$+r_{15} + r_{25} + r_{35} - (r_{12} + r_{13} + r_{23}) \leq 1$	3
61–65	0.243	$+r_{12} + r_{13} + r_{14} + r_{15} - (r_{23} + r_{24} + r_{25} + r_{34} + r_{35} + r_{45}) \leq 1$	4
66–75	0.312	$+r_{12} + r_{14} + r_{15} + r_{23} + r_{34} + r_{35} - (r_{13} + r_{24} + r_{25} + r_{45}) \leq 2$	3
76–87	0.795	$+r_{12} + r_{15} + r_{23} + r_{34} + r_{45} - (r_{13} + r_{14} + r_{24} + r_{25} + r_{35}) \leq 2$	2
88–92	0.344	$+2r_{12} + 2r_{23} + 2r_{24} + 2r_{25} - (r_{13} + r_{14} + r_{15} + r_{34} + r_{35} + r_{45}) \leq 3$	4
93–152	0.688	$+r_{13} + r_{14} + 2r_{24} + r_{34} + 2r_{45} - (2r_{12} + 2r_{25} + 2r_{35}) \leq 3$	3
153–212	0.7306	$+2r_{12} + 2r_{14} + 2r_{15} + r_{23} + r_{35} - (2r_{13} + 2r_{24} + r_{25} + 2r_{45}) \leq 3$	2
213–242	0.855	$+2r_{13} + 2r_{14} + 2r_{23} + 2r_{24} + 3r_{35} + 3r_{45} - (2r_{12} + 4r_{15} + 4r_{25} + r_{34}) \leq 5$	3

Proof. The proof is similar to that of proposition 13, but now we are forced to make two identifications between elements of Λ_1 and of Λ_2 in $\Lambda_1 \sqcup \Lambda_2$. When λ_1 and λ_2 are such that $\epsilon_{\lambda_1}(e_1) = \epsilon_{\lambda_2}(e_2)$, i.e. such that

$$\lambda_1(v_1) = \lambda_1(w_1) \Leftrightarrow \lambda_2(v_2) = \lambda_2(w_2),$$

then this yields a well-defined vertex labelling of G and the result follows. \square

Note that the result is not quite as strong as propositions 11 and 13. While the inequalities of C_{G_1} plus those of C_{G_2} form a complete set of inequalities for the classical polytope of the resulting graph $G_1 +_{(v_1, w_1)=(v_2, w_2)} G_2$, this is not necessarily a minimal set.

Proposition 16. *Let G be a graph and G' be a subgraph of G on the same set of vertices, i.e. $V(G') = V(G)$ and $E(G') \subseteq E(G)$. Then C_G is a subpolytope of $C_{G'} \times [0, 1]^{E(G) \setminus E(G')}$.*

Proof. We need to show that the vertices of C_G constitute a subset of the vertices of $C_{G'} \times [0, 1]^{E(G) \setminus E(G')}$, i.e. that $\text{Eq}(G) \subseteq \text{Eq}(G') \times \{0, 1\}^{E(G) \setminus E(G')}$. Given a classical edge labelling of G , i.e. an edge labelling of the form ϵ_λ for some vertex labelling $\lambda: V(G) \rightarrow \Lambda$, we can regard λ as a vertex labelling of G' and conclude that its equality labelling is simply the restriction of $\epsilon_\lambda: E(G) \rightarrow \{0, 1\}$ to the subset $E(G')$ of its domain. \square

In particular, C_{K_n} is a subpolytope of C_G for any event graph G with n vertices.

Appendix C: Classical polytope facets and quantum violations for small graphs

In this section, we study the facet-defining inequalities of some small graphs. In particular, we analyze and classify the facet-defining inequalities for the classical polytopes C_G corresponding to complete event graphs of 4 and 5 vertices ($G = K_4$ and $G = K_5$, respectively). We also find quantum violations of these inequalities with pure states that are sampled from the set of quantum states. For sampling we used the Python library QuTip [72].

Ref. [32] gave a complete characterization of the classical polytope of the graph $K_3 = C_3$, the smallest graph with non-trivial inequalities, together with a characterization of its maximal quantum violations, as well as applications. More generally, Ref. [32] gave the complete set of inequalities for the classical polytope of the cycle graphs C_n , which take the very simple form in Eq. (1). Here, we move to consider graphs with more than three edges and which are not cycles.

Facet-defining inequalities for small complete graphs.—The facet-defining inequalities of the classical polytope of the graph C_4 (the 4-cycle) are those of the form given by the CHSH inequality mentioned in the main text. If we add one more edge to this graph, the corresponding polytope ends up being described by 3-cycle inequalities alone. Therefore, the first interesting graph yielding non-trivial and non-cycle inequalities is K_4 , the complete graph of 4 vertices. The classical polytope of this graph has facets defined by 3- and 4-cycle inequalities, together with facets defined by the new inequalities described in Eq. (2) in the main text, i.e. those of the form

$$(r_{12} + r_{13} + r_{14}) - (r_{23} + r_{34} + r_{24}) \leq 1.$$

This inequality has a structure that is present for all K_n graphs, as will be discussed in Appendix D. Since complete graphs have all possible edges, these are the graphs that impose the largest number of non-trivial constraints on edge assignments, as per proposition 16. Therefore, it is natural to look at those graphs first.

We addressed the complete characterization of the classical polytopes of complete graphs, proceeding as far as the computational complexity of the problem allowed. In particular, we found complete sets of facet-defining inequalities for C_{K_5} and C_{K_6} . The polytope C_{K_5} has 52 vertices and 242 facets. These facets fall are divided into 9 symmetry classes. Representative inequalities from each of these classes are shown in table I. The polytope C_{K_6} has 203 vertices and requires 50,652 inequalities. A list of inequalities and Python code used to obtain them can be found in Ref. [43].

Quantum violations.—We looked for quantum violations of each inequality class of C_{K_5} obtained by pure states in Hilbert spaces of dimensions 2, 3, and 4. The violations found are included in table I. The inequality in

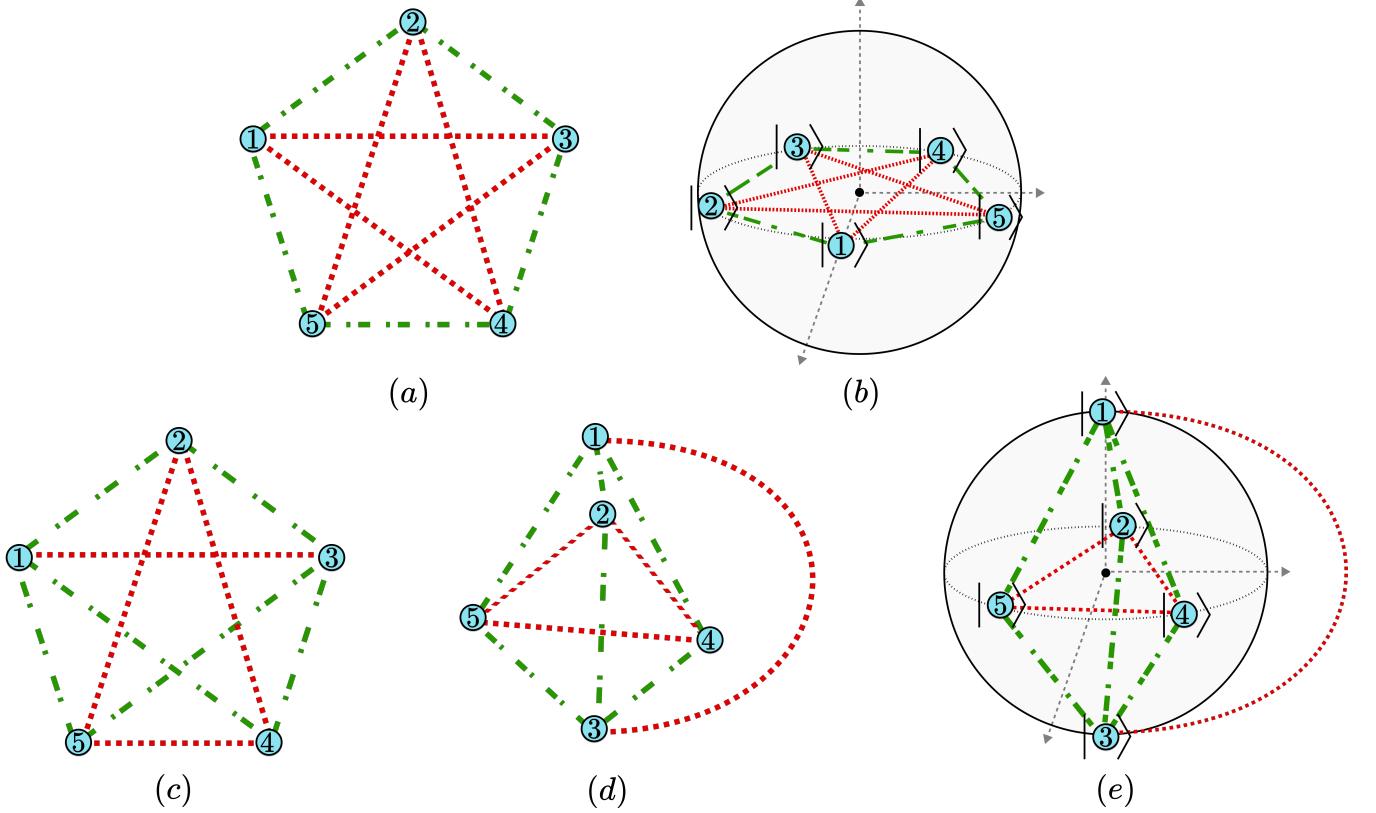


Figure 2. Qubit states violating classicality inequalities. (a) depicts the classicality inequality $r_{12} + r_{23} + r_{34} + r_{45} + r_{15} - r_{13} - r_{14} - r_{24} - r_{25} - r_{35} \leq 2$ with edges corresponding to positive terms in green (dash-dotted lines) and to negative terms in red (dashed-only lines). (b) shows a set of five pure states equally spaced over a great circle of the Bloch sphere, which violates this inequality attaining a value of $5\sqrt{5}/4 > 2$. (c) depicts the classicality inequality $r_{12} + r_{14} + r_{15} + r_{23} + r_{34} + r_{35} - r_{13} - r_{24} - r_{25} - r_{45} \leq 2$ as in (a). (d) depicts the same inequality with the graph displayed in a different geometric configuration, mirroring that of a set of states in the Bloch sphere that largely violates it. (e) represents that set of five pure states in the Bloch sphere: three states equally spaced around the equator plus the two eigenstates of the Pauli matrix σ_z ; this set of states attains a value of $9/4 > 2$ for the inequality.

the third row is apparently not violated by either qubit or qutrit states. The largest violation found among all the inequalities was 0.855, for the inequality in the last row of the table. The sets of quantum states yielding the violations found are presented in Ref. [43].

For some classes of inequalities, we also found violations using pure qubit states that display interesting symmetries in the Bloch sphere. We present those violations in Fig. 2. For instance, consider the inequality in the fifth row of table I. It can be violated with the quantum states

$$|\psi_k\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle + e^{2\pi i k/5} |1\rangle \right) \quad (\text{C1})$$

with $k = 0, \dots, 4$. This quantum realization attains a value of $5\sqrt{5}/4$ and hence a violation of $5\sqrt{5}/4 - 2 \approx 0.79508$. Another interesting violation with qubits is for the inequality in the fourth row of the table. There, a maximal qubit violation is achieved by the states depicted in Fig. 2: choosing $|\psi_2\rangle, |\psi_4\rangle, |\psi_5\rangle$ equally distributed on the equator of the Bloch sphere, i.e. separated by angles of $2\pi/3$, implying that $r_{24} = r_{25} = r_{45} = 1/4$, and choosing $|\psi_1\rangle = |0\rangle, |\psi_3\rangle = |1\rangle$, implying that $r_{13} = 0$ and all

remaining overlaps are equal to $1/2$. This set of vectors attains the value $6/2 - 3/4 = 9/4$ and hence a violation of $9/4 - 2 = 1/4$. These symmetrically arranged qubit states are also the states used in the construction of the elegant joint measurement of Ref. [73]. However, we could find a higher violation of the same inequality using qutrits, as shown in the table.

We will see in appendix D that the inequality in Eq. (2) generalizes to an infinite family of inequalities for the polytope of K_n . The quantum violation found for this non-cycle K_4 inequality used the following four qutrit states:

$$\begin{aligned} |\psi_1\rangle &= |0\rangle \\ |\psi_2\rangle &= \sqrt{\frac{5}{9}} |0\rangle + \sqrt{\frac{4}{9}} |1\rangle \\ |\psi_3\rangle &= \sqrt{\frac{5}{9}} |0\rangle - \sqrt{\frac{1}{9}} |1\rangle + i\sqrt{\frac{1}{3}} |2\rangle \\ |\psi_4\rangle &= \sqrt{\frac{5}{9}} |0\rangle - \sqrt{\frac{1}{9}} |1\rangle - i\sqrt{\frac{1}{3}} |2\rangle \end{aligned}$$

This set of states attains a value of $4/3$ and hence violation of $4/3 - 1 = 1/3$. This corresponds to the second class of inequalities of C_{K_5} in table I.

We remark once more that the above violations are *not* necessarily optimal. They were not found using e.g. techniques of semidefinite programming over the quantum set. We found this landscape of violations by sampling quantum states and calculating the value of the left-hand side of the inequality, which is suboptimal. An important remark is that the quantum violation for the 3-cycle inequality class (first row in table I) is *provably maximal*, as shown in Ref. [32].

Appendix D: Infinite family of classical polytope facets

Equation (2) in the main text shows a facet-defining inequality of the polytope C_{K_4} that is not of the previously known form of inequalities derived from cycles in Ref. [32] (which were enough, incidentally, to characterise the classical polytope of the graph $K_3 = C_3$). In this section, we generalize it to an infinite family of new classicality inequalities. More concretely, we present a construction of a facet-defining inequality of the classical polytope C_{K_n} for any $n \geq 2$. Moreover, each inequality on this family cannot be obtained from combining prior members of the family. For $n = 4$, this recovers the just-mentioned inequality from Eq. (2), while for $n = 3$ it naturally reduces to the 3-cycle inequality.

Fix a natural number $n \geq 2$. Write $V_n = \{1, \dots, n\}$ for the vertices of K_n , and let E_n denote the set of edges of K_n , i.e. all two-element subsets of V_n . Consider a partition of E_n into the subsets $G_n, R_n \subseteq E_n$ given as

$$\begin{aligned} G_n &:= \{\{1, i\} \mid i = 2, \dots, n\} \\ R_n &:= E_n \setminus G_n. \end{aligned}$$

The edges in R_n determine a complete subgraph of K_n with one fewer vertex, i.e. a subgraph isomorphic to K_{n-1} . In turn, the edges in G_n form a subgraph isomorphic to $K_{1, n-1}$, a star graph with n vertices. We use this specific partition of E_n to define a generalized version of the inequality from Eq. (2):

$$h_n(r) := \sum_{e \in G_n} r_e - \sum_{e \in R_n} r_e \leq 1. \quad (\text{D1})$$

We first show that this is indeed a classicality inequality for the complete graph K_n .

Proposition 17. *For any $n \geq 2$, the classical polytope C_{K_n} of the complete event graph K_n is contained in the half-space defined by the inequality h_n from Eq. (D1), i.e. all classical edge weightings of K_n satisfy the inequality.*

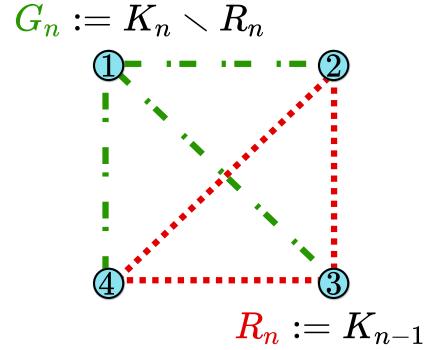


Figure 3. **Depiction of the sets R_n and G_n for a given complete graph K_n .** The set R_n is always a complete subgraph (isomorphic to) K_{n-1} of K_n . Here we considered $n = 4$ as an example.

Proof. It suffices to check that the inequality is satisfied by any vertex of the polytope C_{K_n} . Recall that the vertices of this polytope correspond to classical edge $\{0, 1\}$ -labellings of the graph K_n , that is, those realisable as the equality labelling of some vertex labelling.

So, let $\lambda: V_n \rightarrow \Lambda$ be any vertex labelling and $r \in [0, 1]^{E_n}$ be the vertex of the classical polytope corresponding to its equality edge labelling. That is, for all $e = \{i, j\} \in E_n$,

$$r_{ij} = \epsilon_\lambda(\{i, j\}) = \begin{cases} 1 & \text{if } \lambda(i) = \lambda(j) \\ 0 & \text{if } \lambda(i) \neq \lambda(j) \end{cases}.$$

Consider the set of vertices in $\{2, \dots, n\}$ that are labelled the same as vertex 1,

$$\begin{aligned} S_\lambda &= \{i \in \{2, \dots, n\} \mid \lambda(i) = \lambda(1)\} \\ &= \{i \in \{2, \dots, n\} \mid r_{1i} = 1\} \end{aligned}$$

By construction, an edge in G_n , which is of the form $\{1, i\}$, is labelled 1 or 0 depending on whether i is in S_λ or not. Moreover, by transitivity of equality, if $i, j \in S_\lambda$ then $\lambda(i) = \lambda(j)$, meaning that the edge $\{i, j\}$ is also labelled 1. Writing $k := |S_\lambda|$, one can therefore bound the left-hand side of Eq. (D1):

$$\begin{aligned} \sum_{e \in G_n} r_e - \sum_{e \in R_n} r_e &= \sum_{i \in S_\lambda} r_{1i} - \sum_{e \in R_n} r_e \\ &= k - \sum_{e \in R_n} r_e \\ &\leq k - \sum_{i, j \in S_\lambda} r_{ij} \\ &= k - \binom{k}{2} \\ &= 1 - \binom{k-1}{2} \\ &\leq 1 \end{aligned}$$

where for the corner case $k = 0$ this still holds putting $\binom{-1}{2} = 1$ \square

We now state the central result of this section.

Theorem 18. *The inequality h_n from Eq. (D1) defines a facet of the classical polytope C_{K_n} of the complete event graph K_n for any $n \geq 2$.*

Proof. We establish this result by finding the set of vertices of the polytope C_{K_n} that belongs to – and therefore determines – this facet. In fact, it suffices to find a set of points F in the space (of edge weightings) such that: (i) all the points in F belong to the polytope C_{K_n} , (ii) all the points in F saturate the inequality, i.e. belong to the hyperplane determined by it, (iii) the set F is affinely independent, and (iv) F contains as many points as the dimension D of the polytope, so that it generates an affine subspace of dimension $D - 1$. In our proof, the chosen points are moreover vertices of the polytope, as they are edge $\{0, 1\}$ -labellings.

We construct a set F of polytope vertices. This consists of two kinds of edge labellings: those that assign 1 to exactly one edge of G_n (and 0 to all other edges of E_n) and those that assign 1 precisely to a triangle consisting of two edges from G_n and another from R_n . More formally, we define a family of edge $\{0, 1\}$ -labellings indexed by subsets of size 1 or 2 of the vertex set $\{2, \dots, n\}$, as follows: for each $i = 2, \dots, n$, define the edge $\{0, 1\}$ -labelling $r^{(i)}$ with $r_{1i}^{(i)} = 1$ and $r_e^{(i)} = 0$ for all other edges e ; for each pair $i, j = 2, \dots, n$ with $i \neq j$ define the edge $\{0, 1\}$ -labelling $r^{(i,j)}$ with $r_{1i}^{(i,j)} = r_{1j}^{(i,j)} = r_{ij}^{(i,j)} = 1$ and $r_e^{(i,j)} = 0$ for all other edges e . The set F is then given by

$$F := \left\{ r^{(i)} \mid i = 2, \dots, n \right\} \cup \left\{ r^{(i,j)} \mid i, j = 2, \dots, n, i \neq j \right\}.$$

Fig. 4 depicts the construction of the set F for the case of $n = 5$. We now check conditions (i)–(iv) to establish the desired result.

For condition (i), we use proposition 7 to show that all the edge labellings in the set S are classical and thus vertices of the polytope C_{K_n} . Indeed, no cycle can have exactly one edge with label 0. In the case of the labellings of the form $r^{(i)}$, this is immediate as there is only one edge not labelled 0. For the labellings of the form $r^{(i,j)}$, no triangle (i.e. subgraph isomorphic to C_3) has exactly one edge labelled 0: if one chooses two edges labelled 1 then the remaining edge that completes the 3-cycle also has label 1. Moreover, any larger cycle can have at most two edges labelled 1. Alternatively, we can show this by constructing an underlying vertex labelling: for $r^{(i)}$ pick $\lambda: V_n \rightarrow \Lambda$ with $\lambda(1) = \lambda(i)$ and all other vertex labelled differently; for $r^{(i,j)}$ pick λ with $\lambda(1) = \lambda(i) = \lambda(j)$ and the other vertices labelled differently.

Condition (ii) is directly checked: for each $i = 2, \dots, n$ we have

$$\sum_{e \in G_n} r_e^{(i)} - \sum_{e \in R_n} r_e^{(i)} = r_{1i}^{(i)} - 0 = 1 - 0 = 1,$$

and for each pair $i, j = 2, \dots, n$ with $i \neq j$,

$$\sum_{e \in G_n} r_e^{(i,j)} - \sum_{e \in R_n} r_e^{(i,j)} = r_{1i}^{(i,j)} + r_{1j}^{(i,j)} - r_{ij}^{(i,j)} = 2 - 1 = 1.$$

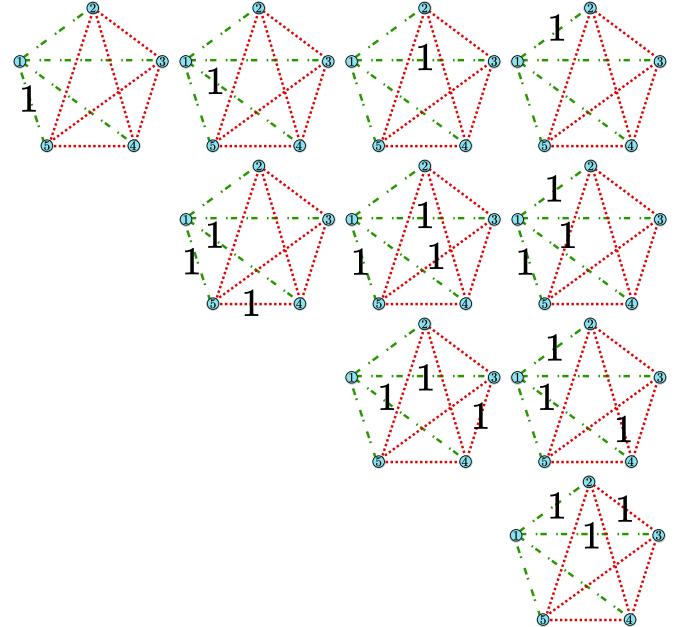


Figure 4. The construction of the set F for the K_5 graph. Each edge is labelled 1 where explicitly noted, otherwise it is labelled 0 (to keep the figures easy to read). The first row shows the four labellings of the form $r^{(i)}$ with only one edge labelled 1 from G_5 . The remaining rows show the labellings of the form $r^{(i,j)}$, which assign label 1 to exactly one triangle consisting of two edges from G_5 and the connecting edge from R_5 .

For condition (iii), affine independence can be verified by inspecting the matrix whose columns are the vectors corresponding to the edge-labellings in F . Ordering the components of each vector (corresponding to the edges of K_n) in lexicographic order and listing $r^{(i)}$ followed by $r^{(i,j)}$ also in that order, the matrix is arranged to be triangular with diagonal entries all equal to 1, hence its determinant is equal to 1, implying linear independence of the vectors.

Finally, for condition (iv), as all these labellings are distinct, one can count the number of elements of S from the way they were constructed:

$$|F| = \binom{n-1}{1} + \binom{n-1}{2} = \binom{n}{2} = \frac{n(n-1)}{2}.$$

We conclude that it is the same as the dimension of the ambient space (of edge labellings) where the polytope lives, and thus also of the polytope itself. \square

Appendix E: Event graphs and Kochen–Specker contextuality

In this section, we establish a formal connection between our framework and (Kochen–Specker) contextuality. The central result (theorem 22) shows how our event graph formalism recovers all noncontextuality in-

equalities obtainable from the Cabello–Severini–Winter (CSW) exclusivity graph approach [28].

To achieve this, we encode a contextuality setup, represented in CSW by an exclusivity graph H , by imposing exclusivity constraints on a related event graph H_* . This process amounts to taking a cross-section yielding a subpolytope of the classical polytope C_{H_*} . We show that the resulting facet inequalities bound noncontextual models for H .

In fact, we prove something *stronger*. We describe an explicit isomorphism between the noncontextual polytope associated by CSW to the exclusivity graph H and this cross-section subpolytope of the classical polytope C_{H_*} associated by our approach to the event graph H_* . In particular, these polytopes have the same non-trivial facet-defining inequalities. These are obtainable from the inequalities that define the full (unconstrained) classical polytope of the event graph H_* by setting some coefficients to zero. Theorem 22 thus establishes a tight correspondence between our event graph approach and a broad, well-established framework for contextuality.

In what follows, we introduce the relevant definitions regarding the exclusivity graph approach, the associated event graphs, and the constraints to be imposed on them, before proving the new results.

The exclusivity graph approach.—In the CSW framework from Ref. [28], contextuality scenarios are described by so-called exclusivity graphs. Hence this formalism is also known as the exclusivity graph approach; see also [31, Chapter 3] for a recent and comprehensive discussion.

The vertices of an exclusivity graph H represent measurement events, and its edges indicate exclusivity between events, where two events are exclusive that can be distinguished by a measurement procedure.

Even though the CSW framework is theory-independent, it is helpful for clarity of exposition to consider its instantiation in quantum theory, in order to better convey the underlying intuitions. In quantum theory, measurement events are represented by projectors (PVM elements) on a Hilbert space, or equivalently, by closed subspaces of the Hilbert space. Exclusivity is captured by orthogonality, which characterizes when two projectors may appear as elements of the same PVM, i.e. events from *the same* measurement procedure. Given a set of projectors $\{\Pi_v\}_{v \in V}$ on a fixed Hilbert space, the corresponding contextuality scenario is thus described by its orthogonality graph. This graph has set of vertices V and has an edge $\{u, v\}$ if and only if the projectors Π_u and Π_v are orthogonal to each other, i.e. when $\Pi_u \Pi_v = 0$.

In this approach, a non-negative vertex weighting $\gamma: V(H) \rightarrow \mathbb{R}_{\geq 0}$ on the exclusivity graph H determines a noncontextuality inequality on the probabilities $P(v)$ of measurement events $v \in V(H)$:

$$\sum_{v \in V(H)} \gamma(v) P(v) \leq \alpha(H, \gamma),$$

where $\alpha(H, \gamma)$ is the independence number of the vertex-weighted graph. In the quantum case, this yields a non-

contextuality condition on the statistics predicted by a given quantum state ψ :

$$\sum_{v \in V(H)} \gamma(v) \langle \psi | \Pi_v | \psi \rangle \leq \alpha(H, \gamma).$$

Such noncontextuality inequalities above determine the polytope of noncontextual behaviours for any exclusivity graph H . This polytope, known as the stable set polytope of H , $\text{STAB}(H)$, is most readily defined by its V-representation, which we now present, following [31, Chapter 3].

Definition 19. Let H be a graph. A subset $S \subseteq V(H)$ of vertices is called a *stable set* if no two vertices of S are adjacent in H , i.e. for all $v, w \in S$, $\{v, w\} \notin E(H)$. Write $\mathcal{S}(H)$ for the set of stable sets of H .

To any subset of vertices $W \subseteq V(H)$ corresponds its characteristic map, the vertex $\{0, 1\}$ -labelling $\chi_W: V(H) \rightarrow \{0, 1\}$ given by:

$$\chi_W(v) := \begin{cases} 1 & \text{if } v \in W, \\ 0 & \text{if } v \notin W. \end{cases}$$

Through the inclusion $\{0, 1\} \subseteq [0, 1]$, one regards a vertex $\{0, 1\}$ -labelling (equivalently, a subset of vertices) as a point in $[0, 1]^{V(H)} \subseteq \mathbb{R}^{V(H)}$. Those arising from stable sets $S \in \mathcal{S}(H)$ correspond to the deterministic noncontextual models, which determine the whole convex set of noncontextual behaviours.

Definition 20. The *stable set polytope* of a graph H , denoted $\text{STAB}(H)$, is the convex hull of the points $\chi_S \in [0, 1]^{V(H)}$ with S ranging over all stable sets of H ,

$$\text{STAB}(H) := \text{ConvHull}\{\chi_S \mid S \in \mathcal{S}(H)\}.$$

To get the intuition underlying this description, one may think of a vertex $\{0, 1\}$ -labelling $\chi_W: V(H) \rightarrow \{0, 1\}$ as a deterministic assignment of truth values to all measurement events (vertices of the exclusivity graph). In this interpretation, the subset of vertices $W \subseteq V(H)$ is the set of measurement events that are assigned *true*. The stability condition indicates that no two adjacent vertices of the exclusivity graph H are labelled with 1, that is, two exclusive measurement events cannot be simultaneously true. This captures the exclusivity condition at the deterministic level, thus yielding the deterministic noncontextual models.

From exclusivity graphs to constrained event graphs.—We relate this approach to our framework by constructing a new (event) graph H_* from any (exclusivity) graph H . This is obtained by adding a new vertex ψ with an edge connecting it to all the vertices of H . See Fig. 5 for an instance of this construction for the KCBS scenario and Fig. 6 for a more generic description. The construction is formally described in definition 21 below.

The relevance of the new vertex ψ is well known; it is usually called the ‘handle’ and it appears in the literature

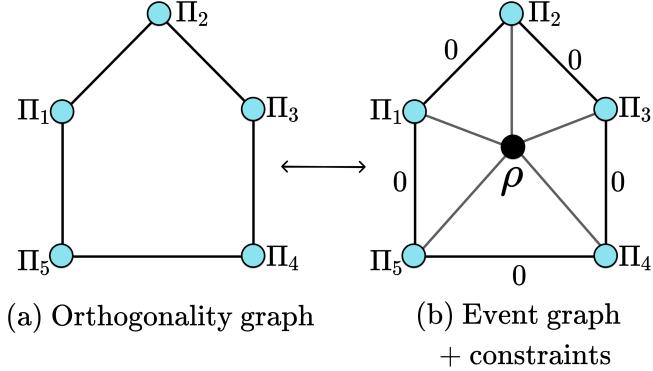


Figure 5. **Equivalence described by theorem 22 linking contextuality à la CSW to event graphs.** The behaviours on an exclusivity graph are in bijective correspondence with edge weightings (overlap assignments) in the related event graph subject to constraints. In particular, the *noncontextual* behaviours for the exclusivity graph correspond bijectively to the *classical* edge weightings in the event graph with constraints.

on the graph approaches [31, 74, 75]. Its name comes from the geometric arrangement of the vectors providing the maximal quantum violation of the KCBS inequality of Eq. (5): the quantum state resembles the handle of an umbrella made of the vectors that describe measurement events.

Definition 21. Let H be a graph. Define a new graph H_* by

$$\begin{aligned} V(H_*) &:= V(H) \sqcup \{\psi\} \\ E(H_*) &:= E(H) \cup \{\{\psi, v\} \mid v \in V(H)\}. \end{aligned}$$

Moreover, define $C_{H_*}^0$ to consist of the classical edge weightings of H_* that assign value 0 to all edges in H ,

$$C_{H_*}^0 := \{r \in C_{H_*} \mid \forall e \in E(H). r_e = 0\}.$$

The set $C_{H_*}^0$ is, by construction, a cross-section of the classical polytope C_{H_*} of the event graph H_* , being its intersection with the $|V(H)|$ -dimensional subspace defined by the equations $\bigwedge_{e \in E(H)} r_e = 0$. Moreover, it is a subpolytope of C_{H_*} , i.e. the convex hull of a subset of its vertices. These vertices are the classical edge $\{0, 1\}$ -labellings that assign label 0 to edges in H . In terms of the underlying vertex labellings (from which classical edge labellings arise as equality labellings), the requirement is that any two vertices adjacent in H must be labelled differently.

Recovering the noncontextual polytope.—The edge set of the graph H_* can be partitioned into two sets: the edges already present in H and the new edges of the form $\{\psi, v\}$ for $v \in V(H)$. The latter are in one-to-one correspondence with vertices of H . So, there is a bijection $E(H_*) \cong E(H) \sqcup V(H)$.

When considering the polytope $C_{H_*} \subseteq [0, 1]^{E(H_*)}$ we adopt the convention of ordering the coordinates with the

edges already in H listed first, so that

$$\mathbb{R}^{E(H_*)} \cong \mathbb{R}^{E(H) \sqcup V(H)} \cong \mathbb{R}^{E(H)} \times \mathbb{R}^{V(H)}.$$

The subpolytope $C_{H_*}^0$ is thus written as the set of points of C_{H_*} of the form $(\mathbf{0}_H, r)$ where $\mathbf{0}_H$ is the zero vector in $\mathbb{R}^{E(H)}$ (corresponding to the edges inherited from H) and r is a weighting of the remaining (new) edges. In particular, the vertices of $C_{H_*}^0$ are precisely the classical $\{0, 1\}$ -labellings of H_* that assign the label 0 to all the edges in H .

We can now prove our main result, showing that C_{H_*} is indeed (isomorphic to) the polytope of noncontextual behaviours for H .

Theorem 22. *For any (exclusivity) graph H , there is an isomorphism of polytopes*

$$C_{H_*}^0 \cong \text{STAB}(H)$$

between the stable set polytope (of noncontextual models) of H and the subpolytope of the classical polytope of event graph H_ constrained by the exclusivity conditions. More explicitly, this is given by the identification*

$$C_{H_*}^0 = \{\mathbf{0}_H\} \times \text{STAB}(H)$$

where $\mathbf{0}_H$ is the zero vector in $\mathbb{R}^{E(H)}$.

Proof. To establish the result, we consider the vertices of these polytopes. Per the above discussion, we have $E(H_*) \cong E(H) \sqcup V(H)$. Consequently, there is a bijection between vertex $\{0, 1\}$ -labellings of H (equivalently, subsets of $V(H)$), on the one hand, and edge $\{0, 1\}$ -labellings of H_* that assign label 0 to all the edges in $E(H)$, on the other. Explicitly, to each subset of vertices $W \subseteq V(H)$ corresponds the edge-labelling of H_*

$$[\mathbf{0}_H, \chi_W]: E(H_*) \cong E(H) \sqcup V(H) \longrightarrow \{0, 1\},$$

as depicted in Fig. 6.

We show that this bijection restricts to a bijection between the *classical* assignments in both cases. Concretely, a subset of vertices $S \subseteq V(H)$ is *stable*, hence (its characteristic map $\chi_S: V(H) \longrightarrow \{0, 1\}$ is) a vertex of the polytope $\text{STAB}(H)$, if and only if the corresponding edge labelling $[\mathbf{0}_H, \chi_S]$ of H_* is classical, hence a vertex of the polytope $C_{H_*}^0$ and thus of $C_{H_*}^0$.

We establish the two directions of this equivalence simultaneously, recalling the characterisation of classical edge labellings from proposition 7. Consider H_* with edge labelling $[\mathbf{0}_H, \chi_S]$. The labelling fails to be classical if and only if there is an edge with label 0 between two vertices linked by a path consisting of edges with label 1. Since all the edges between vertices in H have label 0, the only way to build such a path of 1-labelled edges is via the handle ψ : e.g. $\{u, \psi\}, \{\psi, v\}$ where both u and v must belong to S . So, two vertices u and v of H_* are linked by a 1-labelled path if and only if they both belong to $S \cup \{\psi\}$. Therefore, the labelling is classical if and only

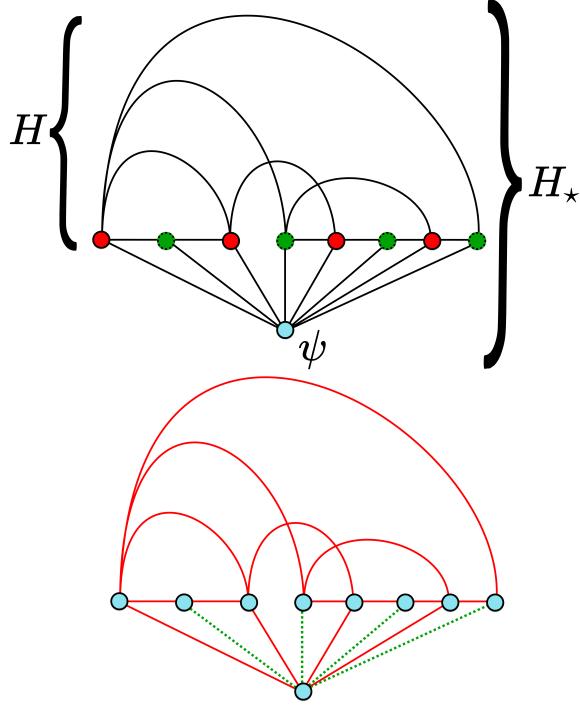


Figure 6. Translation between vertex labellings of H (that are characteristic maps of stable sets, hence vertices of $\text{STAB}(H)$) and constrained edge labellings of H_* (that are classical, hence vertices of $C_{H_*}^0$.) The top figure depicts a graph H , standing for a generic exclusivity graph, and its extension H_* by adjoining the handle ψ and new edges $\{\psi, v\}$ for all $v \in V(H)$. The vertices of H that are shown in green (dashed) form a stable set $S \in \mathcal{S}(H)$. Its characteristic map $\chi_S: V(H) \rightarrow \{0, 1\}$ assigns 1 to the green (dashed) vertices and 0 to the red (solid) vertices of H . The bottom figure shows how such a vertex $\{0, 1\}$ -labelling is translated to an edge $\{0, 1\}$ -labelling of H_* , assigning 0 to all the edges of H (and vice-versa) as described in the proof of theorem 22. Green (dashed) edges are labelled 1 and red (solid) edges are labelled 0, in accordance with the vertex labellings from χ_S , complemented by the labels induced by exclusivity constraints, $\mathbf{0}_H$, as described in the text. S being stable is equivalent to the resulting edge labelling of H_* being classical.

if there is no edge with label 0 between vertices in this set $S \cup \{\psi\}$. To further simplify this condition, note that edges between ψ and a vertex from S have label 1 by construction of the second component of $[\mathbf{0}_H, \chi_S]$, while from the first component, all edges between vertices in H have label 0. The classicality condition is thus equivalent to there being no edges in H between vertices in S , which is precisely to say that S is stable. \square

Recovering all noncontextuality inequalities.— We established theorem 22 in terms of the vertices of the polytopes, i.e. by working with their V-representations. We now consider the relationship between their H-representations, i.e. their facet-defining inequalities.[76]

Of course, there is also a bijection between the facets of

$\text{STAB}(H)$ and those of $C_{H_*}^0$. Given the particularly simple description of the isomorphism, whereby $C_{H_*}^0$ is written as a product of polytopes, we can write this correspondence explicitly. It turns out that the facet-defining inequalities of the subpolytope $C_{H_*}^0$ are precisely the same as the facet-defining inequalities of the stable polytope of H . Moreover, these can be obtained from the inequalities defining the (unconstrained) polytope C_{H_*} of the event graph H_* by setting some coefficients to zero. We thus recover the full set of noncontextuality inequalities from our event graph formalism.

To see this, recall that if P and Q are two convex polytopes with H-representations $P = \{x \mid A_1 x \leq b_1\}$ and $Q = \{y \mid A_2 y \leq b_2\}$ then their product has H-representation

$$P \times Q = \{(x, y) \mid A_1 x \leq b_1 \text{ and } A_2 y \leq b_2\}.$$

Here, the notation $A z \leq b$ describes a set of linear inequalities on z in matrix form, with the symbol \leq standing for component-wise inequality \leq between real numbers.

Applying this to

$$\begin{aligned} C_{H_*}^0 &= \{\mathbf{0}_H\} \times \text{STAB}(H) \\ &= \{(x, y) \mid x \in \{\mathbf{0}_H\}, y \in \text{STAB}(H)\}. \end{aligned}$$

we obtain that the H-representation of $C_{H_*}^0$ is the conjunction of the H-representations of $\{\mathbf{0}_H\}$ and of $\text{STAB}(H)$. The former consists simply of the equations $r_e = 0$ for each $e \in E(H)$, zeroing out the first components, which corresponds to the weights of edges already in H . Thus the non-trivial inequalities bounding $C_{H_*}^0$ are thus the same as the inequalities bounding $\text{STAB}(H)$.

Since $C_{H_*}^0$ is obtained from C_{H_*} by intersecting with the subspace that zeroes the components corresponding to edges in $E(H)$, a complete set of inequalities for $C_{H_*}^0$ can be obtained from the facet-defining inequalities of C_{H_*} by disregarding those components, i.e. setting the corresponding coefficients to zero.

This process is illustrated by the derivation of the KCBS inequality presented in the main text. There, the exclusivity graph is the 5-cycle, with neighbouring vertices representing orthogonal projectors. The graph H_* is then the 6-vertex wheel graph W_6 of Fig. 1-(e). As shown in the main text, the KCBS noncontextuality inequality $\sum_a \gamma_a |\langle \psi | a \rangle|^2 \leq \alpha(H, \gamma)$ arises as a $C_{H_*}^0$ inequality, being obtained from a classicality inequality for the event graph W_6 (a facet-defining inequality of C_{H_*}) by setting to zero the coefficients relating to edges already in H .

Appendix F: Event graphs and preparation contextuality

In this section, we relate our approach to Spekkens's notion of preparation contextuality. This may be understood as providing a *theory-independent* perspective on the use of our formalism to witness quantum coherence. There, the vertices of event graphs were interpreted as

quantum states and the edges as two-state overlaps. A similar treatment can be carried out for a certain class of operational theories which support a notion of confusability, with vertices interpreted as (abstract) preparation procedures.

Operational probabilistic theories.— Spekkens's notion of generalized contextuality is associated to operational probabilistic theories [77–79]. The description of an operational theory starts with a set of basic (operational) physical processes: in the simplest scenarios, one considers preparations and measurements. One considers experiments consisting of a preparation P followed by a measurement M that returns an outcome k . A probability rule associates a probability $p(k | M, P)$ of obtaining outcome k when performing measurement M after the preparation P . More precisely, it associates a probability distribution over outcomes k to each choice of preparation P and measurement M . For a dichotomic measurement M , i.e. one with only two possible outcomes 0 and 1, we simplify notation and write $p(M | P)$ for $p(1 | M, P)$. A crucial – if sometimes overlooked – aspect is that the full set of procedures includes also classical probabilistic mixtures (i.e. convex combinations) of basic procedures, with the probability rule extended accordingly (i.e. linearly).

Given an operational theory, one defines an equivalence relation identifying indistinguishable procedures. Following Ref. [62], two preparation procedures are *operationally equivalent*, written $P \simeq P'$, if and only if for all measurements M and possible outcomes k ,

$$p(k | M, P) = p(k | M, P').$$

A similar definition applies to measurement procedures, but this will not be needed in what follows.

When one treats quantum theory as an operational theory, quantum states $|\phi\rangle$ correspond to equivalence classes of operational procedures. For instance, a state $|0\rangle$ may represent preparing a ground state of a nitrogen atom, or preparing the horizontal polarization in photonic qubits. We relax this terminology and refer to ‘the preparation P associated with a state $|\phi\rangle$ ’, even though strictly speaking P is only an instance of an equivalence class of procedures. Such relaxation is safe for our purposes. In effect, it corresponds to treating pure quantum states as the basic procedures. The interesting operational equivalences relevant for preparation contextuality go beyond these, holding between classical mixtures of basic procedures. For example, in quantum theory, the preparation procedure corresponding of an equal mixture of pure qubit states $|0\rangle$ and $|1\rangle$ is operationally equivalent to that corresponding to an equal mixture of states $|+\rangle$ and $|-\rangle$. Indeed, both these classical mixtures define the same qubit mixed state, the totally mixed state.

LSSS operational constraints.— We wish to generalize the situation in which our graph-theoretic framework is used to witness quantum coherence. There, vertices of an event graph G are interpreted as representing vectors $\{|\phi_i\rangle\}_{i \in V(G)}$ in some Hilbert space \mathcal{H} , i.e. pure quantum states. Edge weights then correspond to two-

state quantum overlaps, $|\langle\phi_i|\phi_j\rangle|^2$. Such overlaps can be accessed empirically by e.g. measuring one of the states on a measurement basis that includes the other.

Abstracting from this, we consider a situation in which we associate a preparation procedure P_i to each vertex $i \in V(G)$ of a given graph G . But in order to emulate the setup above for more general operational theories, it is necessary to impose some additional operational constraints. These constraints distil the aspects of quantum theory that make this work, allowing (a theory-independent version of) two-state overlaps. We shall refer to them as the Lostaglio–Senno–Schmid–Spekkens (LSSS) operational constraints, after Refs. [63, 64]. Note that these constraints apply to preparation procedures; we need not assume any operational equivalences for measurement procedures. Therefore, the scenarios under consideration aim to probe preparation contextuality only.

First, for any preparation P_i , we assume that there is a corresponding ‘test measurement’ M_i with outcomes $\{0, 1\}$ satisfying the operational statistics $p(M_i | P_i) = 1$. In quantum theory, if P_i is the preparation associated with state $|\phi_i\rangle$ then M_i is realised by the projective measurement $\{|\phi_i\rangle\langle\phi_i|, \mathbb{1} - |\phi_i\rangle\langle\phi_i|\}$ where the first projector corresponds to the outcome $k = 1$.

Moreover, for any edge $\{i, j\} \in E(G)$, whose incident vertices have preparations P_i and P_j , we assume that there exists another pair of preparations P_{i^\perp} and P_{j^\perp} satisfying $p(M_i | P_{i^\perp}) = 0$, $p(M_j | P_{j^\perp}) = 0$, and the operational equivalence $\frac{1}{2}P_i + \frac{1}{2}P_{i^\perp} \simeq \frac{1}{2}P_j + \frac{1}{2}P_{j^\perp}$. In quantum theory, this is always satisfied: given a pair of pure states $|\phi_i\rangle$ and $|\phi_j\rangle$, one picks $|\phi_{i^\perp}\rangle$ to be the vector orthogonal to $|\phi_i\rangle$ living in the two-dimensional space spanned by $\{|\phi_i\rangle, |\phi_j\rangle\}$, and similarly for $|\phi_{j^\perp}\rangle$.

The probabilities $p(M_i | P_j)$ are usually called the *confusability* [7, 64], because they may be interpreted as the probability of guessing incorrectly that the preparation performed had been P_i instead of P_j . These probabilities provide a theory-independent, operational treatment of two-state overlaps, which reduces to the familiar notion in the case of quantum theory viewed as an operational theory:

$$p(M_i | P_j) \stackrel{QT}{=} \text{Tr}(|\phi_i\rangle\langle\phi_i||\phi_j\rangle\langle\phi_j|) = |\langle\phi_i|\phi_j\rangle|^2.$$

Therefore, we use these confusability probabilities to provide edge weights $r_{ij} = p(M_i | P_j)$ in our framework. In summary, an assignment of preparation procedures to the vertices of G such that the LSSS operational constraints are satisfied for the pairs of preparations associated to each edge determines an edge weighting $r: E(G) \rightarrow [0, 1]$.

Preparation noncontextuality.— When faced with an operational theory, a natural question is whether it admits a (noncontextual) hidden-variable explanation, that is, whether it can be realised by a noncontextual ontological model. In general, an ontological model consists of a measurable space $(\Lambda, \mathcal{F}_\Lambda)$ of *ontic* states equipped with ontological interpretations for preparation and measurement procedures: preparation procedures P determine

probability measures μ_P on Λ , whereas measurement procedures M determine measurable functions ξ_M mapping each ontic state $\lambda \in \Lambda$ to (a distribution on) outcomes. Note that the interpretation of classical mixtures of procedures must be determined linearly from that of basic procedures, e.g. $\mu_{\frac{1}{2}P+\frac{1}{2}Q} = \frac{1}{2}\mu_P + \frac{1}{2}\mu_Q$. The composition of the interpretations of a preparation and a measurement (going via the ontic space Λ) is required to recover the empirical or operational predictions, i.e.

$$p(\cdot | M, P) = \int_{\Lambda} \xi_M d\mu_P,$$

or with variables,

$$p(k | M, P) = \int_{\Lambda} \xi_M(k | \lambda) d\mu_P(\lambda).$$

Such a realization by an ontological model is said to be noncontextual if operationally equivalent procedures are given the same interpretation. For preparations, the requirement is that two operationally equivalent preparation procedures determine the same probability measure on Λ . We refrain from going into detail on the general definition, as the characterization that follows suffices.

In Refs. [63, 64], it was shown that the LSSS constraints imply that any preparation noncontextual model explaining preparation procedures P_i as probability measures μ_i on Λ must satisfy

$$p(M_i | P_j) = 1 - \|\mu_i - \mu_j\|_{\text{TV}}, \quad (\text{F1})$$

where $\|\cdot - \cdot\|_{\text{TV}}$ denotes the total variation distance between probability measures, given for an arbitrary measurable space $(\Lambda, \mathcal{F}_{\Lambda})$ by

$$\|\mu_i - \mu_j\|_{\text{TV}} = \sup_{E \in \mathcal{F}_{\Lambda}} |\mu_i(E) - \mu_j(E)|.$$

In the case when Λ is discrete (which is effectively all we actually need), this distance is related to the l_1 norm[80]:

$$\|\mu_i - \mu_j\|_{\text{TV}} = \frac{1}{2} \|\mu_i - \mu_j\|_1 = \frac{1}{2} \sum_{\lambda \in \Lambda} |\mu_i(\lambda) - \mu_j(\lambda)|.$$

We can take that as a *definition* of preparation noncontextual edge weightings.

Definition 23. Let G be an event graph. An edge weighting $r: E(G) \rightarrow [0, 1]$ is said to be *preparation noncontextual* if the edge weights are of the form in the right-hand side of Eq. (F1), i.e. $r_{ij} = 1 - \|\mu_i - \mu_j\|_{\text{TV}}$, for some choice of an (ontic) measurable space Λ and of probability measures μ_i on Λ for each vertex $i \in V(G)$.

Cycle inequalities witness preparation contextuality.— We now show how in the case of cycle graphs the inequalities derived from our framework serve as witnesses of preparation contextuality for operational theories satisfying the LSSS constraints.

The technical result is stated in the following proposition; it follows from the triangle inequality.

Proposition 24. Any inequality bounding the set C_{C_n} cannot be violated by a preparation noncontextual edge weighting (definition 23).

Proof. For simplicity, we use addition modulo n when labelling the vertices of the cycle graph C_n , meaning that $i = i + n$. From the triangle inequality of the norm $\|\cdot\|_{\text{TV}}$ it follows that

$$\begin{aligned} & \|\mu_i - \mu_{i+n-1}\|_{\text{TV}} \\ &= \|\underbrace{\mu_i - \mu_{i+1} + \mu_{i+1} - \cdots - \mu_{i+n-2} + \mu_{i+n-2}}_{n-2 \text{ zeros}} - \mu_{i+n-1}\|_{\text{TV}} \\ &\leq \|\mu_i - \mu_{i+1}\|_{\text{TV}} + \cdots + \|\mu_{i+n-2} - \mu_{i+n-1}\|_{\text{TV}}. \end{aligned}$$

Therefore, writing $\|\mu_{i,j}\|_{\text{TV}} := \|\mu_i - \mu_j\|_{\text{TV}}$ for clarity,

$$\|\mu_{i,i+n-1}\|_{\text{TV}} - \|\mu_{i,i+1}\|_{\text{TV}} - \cdots - \|\mu_{i+n-2,i+n-1}\|_{\text{TV}} \leq 0.$$

We must now add 1 to each term to recover the noncontextual overlaps of Eq. (F1). We have n terms, but since the first term has a different sign, two of these 1s will cancel, leaving $n - 2$ added to both sides of the inequality:

$$\begin{aligned} & -1 + \|\mu_{i,i+n-1}\|_{\text{TV}} + 1 - \|\mu_{i,i+1}\|_{\text{TV}} \\ &+ \cdots + 1 - \|\mu_{i+n-2,i+n-1}\|_{\text{TV}} \leq n - 2. \end{aligned}$$

Recalling that $r_{ij} = 1 - \|\mu_{i,j}\|_{\text{TV}}$, we recover a cycle inequality for any chosen vertex i :

$$-r_{i,i+n-1} + r_{i,i+1} + \cdots + r_{i+n-2,i+n-1} \leq n - 2.$$

□

We may see this result from two perspectives. We can take a *theory-dependent perspective* and look for what information we can extract assuming quantum theory as the relevant operational theory; this proposition then shows that pure quantum states that violate the n -cycle inequalities can be used to construct a proof of quantum preparation contextuality. The construction is done by constructing states and measurements that represent a realization of the prepare-and-measure scenario described by the LSSS constraints. In summary, violations of these inequalities serve as *witnesses of quantum preparation contextuality*.

In light of this result, the experiment of Ref. [81] can be understood as an experimental test that witnessed preparation contextuality of quantum theory; however since the purpose was not to witness preparation contextuality the authors have not experimentally probed the relevant operational equivalences, and have not ruled out loopholes for such a test.

We can also take a *theory-independent perspective*. If a given operational theory satisfying the LSSS constraints for some cycle graph admits a preparation noncontextual ontological model, then the confusabilities $r_{ij} = p(M_i | P_j)$ are bounded by the cycle inequalities. For instance, the Spekkens Toy Theory [33] satisfies the LSSS constraints for any pair of preparation procedures. Since it admits a noncontextual ontological model, it cannot violate the cycle inequalities.

-
- [1] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson, “Contextuality supplies the ‘magic’ for quantum computation,” *Nature* **510**, 351–355 (2014).
- [2] Alastair A. Abbott, Cristian S. Calude, Jonathan Conder, and Karl Svozil, “Strong Kochen-Specker theorem and incomputability of quantum randomness,” *Physical Review A* **86**, 062109 (2012).
- [3] Robert Raussendorf, “Contextuality in measurement-based quantum computation,” *Physical Review A* **88**, 022322 (2013).
- [4] Samson Abramsky, Rui Soares Barbosa, and Shane Mansfield, “Contextual fraction as a measure of contextuality,” *Physical Review Letters* **119**, 050504 (2017).
- [5] Shane Mansfield and Elham Kashefi, “Quantum advantage from sequential-transformation contextuality,” *Physical Review Letters* **121**, 230401 (2018).
- [6] Debashis Saha and Anubhav Chaturvedi, “Preparation contextuality as an essential feature underlying quantum communication advantage,” *Physical Review A* **100**, 022108 (2019).
- [7] Matteo Lostaglio, “Certifying quantum signatures in thermodynamics and metrology via contextuality of quantum linear response,” *Physical Review Letters* **125**, 230603 (2020).
- [8] William M. Kirby and Peter J. Love, “Classical simulation of noncontextual Pauli Hamiltonians,” *Physical Review A* **102**, 032418 (2020).
- [9] Costantino Budroni, Adán Cabello, Otfried Gühne, Matthias Kleinmann, and Jan-Åke Larsson, “Kochen-specker contextuality,” *Reviews of Modern Physics* **94**, 045007 (2022).
- [10] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner, “Bell nonlocality,” *Reviews of Modern Physics* **86**, 419–478 (2014).
- [11] Alexander Streltsov, Gerardo Adesso, and Martin B. Plenio, “Colloquium: Quantum coherence as a resource,” *Reviews of Modern Physics* **89**, 041003 (2017).
- [12] Andreas Winter and Dong Yang, “Operational resource theory of coherence,” *Physical Review Letters* **116**, 120404 (2016).
- [13] John H. Selby and Ciarán M. Lee, “Compositional resource theories of coherence,” *Quantum* **4**, 319 (2020).
- [14] Barbara Amaral, “Resource theory of contextuality,” *Philosophical Transactions of the Royal Society A* **377**, 20190010 (2019).
- [15] Cristhiano Duarte and Barbara Amaral, “Resource theory of contextuality for arbitrary prepare-and-measure experiments,” *Journal of Mathematical Physics* **59**, 062202 (2018).
- [16] Elie Wolfe, David Schmid, Ana Belén Sainz, Ravi Kunjwal, and Robert W. Spekkens, “Quantifying Bell: the resource theory of nonclassicality of common-cause boxes,” *Quantum* **4**, 280 (2020).
- [17] David Schmid, Denis Rosset, and Francesco Buscemi, “The type-independent resource theory of local operations and shared randomness,” *Quantum* **4**, 262 (2020).
- [18] Kang-Da Wu, Tulja Varun Kondra, Swapna Rana, Carlo Maria Scandolo, Guo-Yong Xiang, Chuan-Feng Li, Guang-Can Guo, and Alexander Streltsov, “Operational resource theory of imaginarity,” *Physical Review Letters* **126**, 090401 (2021).
- [19] Eric Chitambar and Gilad Gour, “Quantum resource theories,” *Reviews of Modern Physics* **91**, 025001 (2019).
- [20] Samson Abramsky, Rui Soares Barbosa, Martti Karvonen, and Shane Mansfield, “A comonadic view of simulation and quantum resources,” in *34th Annual ACM/IEEE Symposium on Logic in Computer Science (LiCS 2019)* (IEEE, 2019) pp. 1–12.
- [21] Rui Soares Barbosa, Martti Karvonen, and Shane Mansfield, “Closing Bell: Boxing black box simulations in the resource theory of contextuality,” in *Samson Abramsky on Logic and Structure in Computer Science and Beyond*, Outstanding Contributions to Logic, Vol. 25, edited by Alessandra Palmigiano and Mehrnoosh Sadrzadeh (Springer Cham, 2023).
- [22] Rafael Wagner, Roberto D. Baldijão, Alisson Tezzin, and Bárbara Amaral, “Using a resource theoretic perspective to witness and engineer quantum generalized contextuality for prepare-and-measure scenarios,” arXiv preprint arXiv:2102.10469 [quant-ph] (2021).
- [23] Bartosz Regula, “Convex geometry of quantum resource quantification,” *Journal of Physics A: Mathematical and Theoretical* **51**, 045303 (2017).
- [24] Bartosz Regula, “Tight constraints on probabilistic convertibility of quantum states,” *Quantum* **6**, 817 (2022).
- [25] Thomas Theurer, Nathan Killoran, Dario Egloff, and Martin B. Plenio, “Resource theory of superposition,” *Physical Review Letters* **119**, 230401 (2017).
- [26] Sébastien Designolle, Roope Uola, Kimmo Luoma, and Nicolas Brunner, “Set coherence: Basis-independent quantification of quantum coherence,” *Physical Review Letters* **126**, 220404 (2021).
- [27] Bartosz Regula, “Probabilistic transformations of quantum resources,” *Physical Review Letters* **128**, 110505 (2022).
- [28] Adán Cabello, Simone Severini, and Andreas Winter, “Graph-theoretic approach to quantum correlations,” *Physical Review Letters* **112**, 040401 (2014).
- [29] Samson Abramsky and Samson Brandenburger, “The sheaf-theoretic structure of non-locality and contextuality,” *New Journal of Physics* **13**, 113036 (2011).
- [30] Antonio Acín, Tobias Fritz, Anthony Leverrier, and Ana Belén Sainz, “A combinatorial approach to nonlocality and contextuality,” *Communications in Mathematical Physics* **334**, 533–628 (2015).
- [31] Barbara Amaral and Marcelo Terra Cunha, *On graph approaches to contextuality and their role in quantum theory*, SpringerBriefs in Mathematics (Springer Cham, 2018).
- [32] Ernesto F. Galvão and Daniel J. Brod, “Quantum and classical bounds for two-state overlaps,” *Physical Review A* **101**, 062110 (2020).
- [33] Robert W. Spekkens, “Evidence for the epistemic view of quantum states: A toy theory,” *Physical Review A* **75**, 032110 (2007).
- [34] Lorenzo Catani, Matthew Leifer, David Schmid, and Robert W. Spekkens, “Why interference phenomena do not capture the essence of quantum theory,” arXiv preprint arXiv:2102.10469 [quant-ph] (2021).
- [35] Adán Cabello, “Experimentally testable state-independent quantum contextuality,” *Physical Review Letters* **101**, 210401 (2008).

- [36] Elias Amselem, Magnus Rådmark, Mohamed Bourennane, and Adán Cabello, “State-independent quantum contextuality with single photons,” *Physical Review Letters* **103**, 160405 (2009).
- [37] Ryuji Takagi and Bartosz Regula, “General resource theories in quantum mechanics and beyond: Operational characterization via discrimination tasks,” *Physical Review X* **9**, 031053 (2019).
- [38] Note that we do not assume a fixed finite outcome set Λ , or a bound on its size. The classical polytope consists of the edge weightings that arise from jointly distributed random variables with outcomes in *some* set Λ . We could fix a single Λ as long as it is countably infinite. But in practice, for a fixed graph G with n vertices, it suffices to consider $\Lambda = \{1, \dots, n\}$.
- [39] George Boole, *An investigation of the laws of thought: On which are founded the mathematical theories of logic and probabilities* (Macmillan, London, 1854).
- [40] Itamar Pitowsky, “George Boole’s ‘conditions of possible experience’ and the quantum puzzle,” *The British Journal for the Philosophy of Science* **45**, 95–125 (1994).
- [41] Samson Abramsky and Lucien Hardy, “Logical Bell inequalities,” *Physical Review A* **85**, 062114 (2012).
- [42] The inequalities found in this work were obtained using the **traf** option from the **PORTA** program, which converts a V-representation of a polytope into an H-representation.
- [43] “Inequality zoo for event graphs,” GitHub repository, <https://github.com/rafunds/CoherenceContextuality> (2022).
- [44] Taira Giordani, Rafael Wagner, Chiara Esposito, Anita Camillini, Francesco Hoch, Gonzalo Carvacho, Ciro Pentangelo, Francesco Ceccarelli, Simone Piacentini, Andrea Crespi, Nicolò Spagnolo, Roberto Osellame, Ernesto F Galvão, and Fabio Sciarrino, “Experimental certification of contextuality, coherence, and dimension in a programmable universal photonic processor,” *Science Advances* **9**, eadj4249 (2023).
- [45] Tillmann Baumgratz, Marcus Cramer, and Martin B. Plenio, “Quantifying coherence,” *Physical Review Letters* **113**, 140401 (2014).
- [46] Johan Åberg, “Quantifying superposition,” arXiv preprint: quant-ph/0612146 (2006).
- [47] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf, “Quantum fingerprinting,” *Physical Review Letters* **87**, 167902 (2001).
- [48] Simon Kochen and Ernst Specker, “The problem of hidden variables in quantum mechanics,” *Journal of Mathematics and Mechanics* **17**, 59–87 (1967).
- [49] John S. Bell, “On the Einstein Podolsky Rosen paradox,” *Physics Physique Fizika* **1**, 195–200 (1964).
- [50] Arthur Fine, “Hidden variables, joint probability, and the Bell inequalities,” *Physical Review Letters* **48**, 291–295 (1982).
- [51] Arthur Fine, “Joint distributions, quantum correlations, and commuting observables,” *Journal of Mathematical Physics* **23**, 1306–1310 (1982).
- [52] Mateus Araújo, *Quantum realism and quantum surrealism*, Master’s thesis, Universidade Federal de Minas Gerais (UFMG) (2012), available as arXiv:1208.6283 [quant-ph].
- [53] Martti Karvonen, “Neither contextuality nor nonlocality admits catalysts,” *Phys. Rev. Lett.* **127**, 160402 (2021).
- [54] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt, “Proposed experiment to test local hidden-variable theories,” *Physical Review Letters* **23**, 880–884 (1969).
- [55] Equal outcomes and different outcomes are described in a dual way. Without loss of generality, we can use either the probability that the two measurements return equal outcomes or the probability that they return different outcomes. The forbidden deterministic edge labellings are essentially the same up to permuting 0 and 1.
- [56] Daniel Collins, Nicolas Gisin, Noah Linden, Serge Massar, and Sandu Popescu, “Bell inequalities for arbitrarily high-dimensional systems,” *Physical Review Letters* **88**, 040404 (2002).
- [57] Samuel L. Braunstein and Carlton M. Caves, “Wringing out better Bell inequalities,” *Annals of Physics* **202**, 22–56 (1990).
- [58] Mateus Araújo, Marco Túlio Quintino, Costantino Budroni, Marcelo Terra Cunha, and Adán Cabello, “All noncontextuality inequalities for the n -cycle scenario,” *Physical Review A* **88**, 022118 (2013).
- [59] Thinh P. Le, Chiara Meroni, Bernd Sturmels, Reinhard F. Werner, and Timo Ziegler, “Quantum Correlations in the Minimal Scenario,” *Quantum* **7**, 947 (2023).
- [60] One may think of a *measurement event* as a pair (m, o) describing that the measurement m is performed and the outcome o is observed.
- [61] Alexander A. Klyachko, M. Ali Can, Sinem Binicioglu, and Alexander S. Shumovsky, “Simple test for hidden variables in spin-1 systems,” *Physical Review Letters* **101**, 020403 (2008).
- [62] Robert W. Spekkens, “Contextuality for preparations, transformations, and unsharp measurements,” *Physical Review A* **71**, 052108 (2005).
- [63] Matteo Lostaglio and Gabriel Senno, “Contextual advantage for state-dependent cloning,” *Quantum* **4**, 258 (2020).
- [64] David Schmid and Robert W. Spekkens, “Contextual advantage for state discrimination,” *Physical Review X* **8**, 011015 (2018).
- [65] Robert W. Spekkens, “Negativity and contextuality are equivalent notions of nonclassicality,” *Physical Review Letters* **101**, 020401 (2008).
- [66] Roberto D. Baldijão, Rafael Wagner, Cristhiano Duarte, Bárbara Amaral, and Marcelo Terra Cunha, “Emergence of noncontextuality under quantum Darwinism,” *PRX Quantum* **2**, 030351 (2021).
- [67] Michał Oszmaniec, Daniel J. Brod, and Ernesto F. Galvão, “Measuring relational information between quantum states, and applications,” arXiv preprint arXiv:2109.10006 [quant-ph] (2021).
- [68] Rafael Wagner, Anita Camillini, and Ernesto F. Galvão, “Coherence and contextuality in a Mach–Zehnder interferometer,” arXiv preprint arXiv:2210.05624 [quant-ph] (2022).
- [69] OEIS Foundation Inc., “The Bell or exponential numbers,” Entry A000110 in The On-line Encyclopedia of Integer Sequences, <https://oeis.org/A000110> (2022).
- [70] Eric T. Bell, “Exponential numbers,” *The American Mathematical Monthly* **41**, 411–419 (1934).
- [71] Nikolai N. Vorob’ev, “Consistent families of measures and their extensions,” *Theory of Probability and its Applications (Teoriya Veroyatnostei i ee Primeneniya)* **7**, 147–163 (English) / 153–159 (Russian) (1962).
- [72] J. Robert Johansson, Paul D. Nation, and Franco Nori, “QuTiP: An open-source Python framework for the dy-

- namics of open quantum systems,” Computer Physics Communications **183**, 1760–1772 (2012).
- [73] Nicolas Gisin, “Entanglement 25 years after quantum teleportation: Testing joint measurements in quantum networks,” Entropy **21**, 325 (2019).
- [74] Roberto D. Baldijão and Marcelo Terra Cunha, “Classical limits and contextuality in a scenario of multiple observers,” Physical Review A **102**, 052226 (2020).
- [75] Lina Vandré and Marcelo Terra Cunha, “Quantum sets of the multicolored-graph approach to contextuality,” Physical Review A **106**, 062210 (2022).
- [76] H-representation is standard terminology referring to the description of a polytope as an intersection of half-spaces, i.e. in terms of (facet-defining) inequalities. The ‘H’ in ‘H-representation’ is not to be confused with the symbol ‘ H ’ that we use to denote an exclusivity graph.
- [77] Giacomo Mauro D’Ariano, Giulio Chiribella, and Paolo Perinotti, *Quantum theory from first principles: An informational approach* (Cambridge University Press, 2017).
- [78] David Schmid, John H. Selby, Matthew F. Pusey, and Robert W. Spekkens, “A structure theorem for generalized-noncontextual ontological models,” arXiv preprint arXiv:2005.07161 [quant-ph] (2020).
- [79] Ravi Kunjwal, “Beyond the Cabello-Severini-Winter framework: Making sense of contextuality without sharpness of measurements,” Quantum **3**, 184 (2019).
- [80] In the continuous case, it is often rendered as $\|\mu_i - \mu_j\|_{\text{TV}} = \int_{\Lambda} |\mu_i(\lambda) - \mu_j(\lambda)| d\lambda$ in terms of a reference measure such as the Lebesgue measure on the real line.
- [81] Taira Giordani, Chiara Esposito, Francesco Hoch, Gonzalo Carvacho, Daniel J. Brod, Ernesto F. Galvão, Nicolò Spagnolo, and Fabio Sciarrino, “Witnesses of coherence and dimension from multiphoton indistinguishability tests,” Physical Review Research **3**, 023031 (2021).