Logical Bell inequalities

Samson Abramsky*
Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford OX1 3QD, United Kingdom

Lucien Hardy†
Perimeter Institute, 31 Caroline Street, North Waterloo, ON N2L 2Y5, Canada
(Received 9 March 2012; published 20 June 2012)

Bell inequalities play a central role in the study of quantum nonlocality and entanglement, with many applications in quantum information. Despite the huge literature on Bell inequalities, it is not easy to find a clear conceptual answer to what a Bell inequality is, or a clear guiding principle as to how they may be derived. In this paper, we introduce a notion of logical Bell inequality which can be used to systematically derive testable inequalities for a very wide variety of situations. There is a single clear conceptual principle, based on purely logical consistency conditions, which underlies our notion of logical Bell inequalities. We show that in a precise sense, all Bell inequalities can be taken to be of this form. Our approach is very general. It applies directly to any family of sets of commuting observables. Thus it covers not only the n-partite scenarios to which Bell inequalities are standardly applied, but also Kochen-Specker configurations, and many other examples. There is much current work on experimental tests for contextuality. Our approach directly yields, in a systematic fashion, testable inequalities for a very general notion of contextuality. There has been much work on obtaining proofs of Bell's theorem "without inequalities" or "without probabilities." These proofs are seen as being in a sense more definitive and logically robust than the inequality-based proofs. On the hand, they lack the fault-tolerant aspect of inequalities. Our approach reconciles these aspects, and in fact shows how the logical robustness can be converted into systematic, general derivations of inequalities with provable violations. Moreover, the kind of strong non-locality or contextuality exhibited by the GHZ argument or by Kochen-Specker configurations can be shown to lead to maximal violations of the corresponding logical Bell inequalities. Thus the qualitative and the quantitative aspects are combined harmoniously.

DOI: 10.1103/PhysRevA.85.062114 PACS number(s): 03.65.Ud

I. INTRODUCTION

There is a huge literature on Bell inequalities [1,2], with many ingenious derivations of families of inequalities. However, a unifying principle with a clear conceptual basis has proved elusive. In this paper, we introduce a form of Bell inequality based on logical consistency conditions, which we call *logical Bell inequalities*. This approach is both conceptually illuminating and technically powerful.

To get some feeling for the results, we shall first discuss how Bell inequalities are used. Their main application, of course, is to show the nonlocality of quantum mechanics, as famously first demonstrated in Bell's theorem [1]. More broadly, Bell inequalities are used to delineate those situations which can be accounted for by classical physical concepts from those which are inherently nonclassical; the content of Bell's theorem is exactly that quantum mechanics produces empirically accessible phenomena which fall into the latter category.

An important feature of the inequalities is that they have a *fault-tolerant aspect* which makes them very suitable for experimental verification. Violation of a Bell inequality is quantitative, and allows nonclassicality to be demonstrated without relying on idealized perfect measurements or state preparations.

There are also many applications of Bell inequalities in quantum information, for example, in quantum key distribution [4–6], quantum communication complexity [7], and detection of quantum entanglement [8], so that they also play a leading role in more applied work.

Although a huge literature on Bell inequalities has appeared over the past few decades, it is not easy to distill from this literature a clear conceptual answer to what a Bell inequality *is*, or a clear guiding principle as to how they may be derived.

The present paper addresses this point, and introduces a notion of *logical Bell inequality* which can be used to systematically derive testable inequalities for a very wide variety of situations. The following points in particular are worth emphasizing:

(1) There is a single clear conceptual principle, based on purely logical consistency conditions, which underlies our notion of logical Bell inequalities. We show that in a precise sense, *all* Bell inequalities can be taken to be of this form.

(2) Our approach is very general—much more so than the great majority of the literature on Bell inequalities. It applies directly to *any* family of sets of commuting observables. Thus it covers not only the n-partite scenarios to which Bell inequalities are standardly applied, but also Kochen-Specker configurations, and many other examples. This is important since there is much current work on experimental tests for *contextuality*, a broader phenomenon than nonlocality; see e.g., [9,10]. Our approach directly yields, in a systematic fashion, testable inequalities for a very general notion of contextuality.

*samson@cs.ox.ac.uk
†lhardy@perimeterinstitute.ca

(3) There has been much work on obtaining proofs of Bell's theorem "without inequalities" or 'without probabilities" [11–13]. These proofs are seen as being in a sense more definitive and logically robust than the inequality-based proofs. On the hand, they lack the fault-tolerant aspect of inequalities. Our approach fully reconciles these aspects, and in fact shows how the logical robustness can be converted into systematic, general derivations of inequalities with provable violations. Moreover, the kind of strong nonlocality or contextuality exhibited by the GHZ argument or by Kochen-Specker configurations can be shown to lead to *maximal violations* of the corresponding logical Bell inequalities. Thus the qualitative and the quantitative aspects are combined harmoniously.

We now turn to a more precise, technical summary of our results.

We show that a rational inequality is satisfied by all noncontextual models if and only if it is equivalent to a logical Bell inequality. Thus quantitative tests for contextuality or nonlocality always hinge on purely logical consistency conditions. We obtain explicit descriptions of complete sets of inequalities for the convex polytope of noncontextual probability models, and the derived polytope of expectation values for these models. Moreover, these results are obtained at a high level of generality; they apply not only to the familiar cases of Bell-type scenarios, for any number of parties, but to all Kochen-Specker configurations, and in fact to *any* family of sets of compatible measurements. This generality is achieved by working with *measurement covers*, following the sheaf-theoretic approach to nonlocality and contextuality introduced by the first author and Adam Brandenburger in Ref. [3].

We also obtain results for a number of special cases. We show that a model achieves maximal violation of a logical Bell inequality if and only if it is strongly (or maximally) contextual. We show that all Kochen-Specker configurations lead to maximal violations of logical Bell inequalities in a state-independent fashion. We also derive specific violations of logical Bell inequalities for models which are possibilistically contextual, meaning that they admit logical proofs of contextuality. Well-known examples of such models are those arising from a construction given by one of us 20 years ago [12,14].

Inspiration for the present work was drawn from [15], which derives some particular cases of logical Bell inequalities. Developing these ideas in the general setting provided by [3] proves to be fruitful, and indicates the potential for a structural approach to quantum foundations.

A. A simple observation

We begin with a simple and very general scenario.

Suppose we have propositional formulas $\varphi_1, \ldots, \varphi_N$. We suppose further that we can assign a probability p_i to each φ_i.

In particular, we have in mind the situation where the Boolean variables appearing in φ_i correspond to empirically testable quantities; φ_i then expresses a condition on the outcomes of an experiment involving these quantities. The probabilities p_i are obtained from the statistics of these experiments.

Now let P be the probability of $\Phi := \bigwedge_i \varphi_i$. Using elementary probability theory, we can calculate as follows:

$$1 - P = \text{Prob}(\neg\Phi) = \text{Prob}\left(\bigvee_i \neg\varphi_i\right) \leqslant \sum_i \text{Prob}(\neg\varphi_i)$$

$$= \sum_i (1 - p_i) = N - \sum_i p_i.$$

Tidying this up yields $\sum_i p_i \leqslant N - 1 + P$.

Now suppose that the formulas φ_i are *jointly contradictory*; (i.e., Φ is unsatisfiable). This implies that $P = 0$. Hence we obtain the inequality,

$$\sum_i p_i \leqslant N - 1.$$

This inequality was obtained in Ref. [15], where it was used to derive chained Bell inequalities (as originally obtained in Ref. [16]). It is an example of a *logical Bell inequality*. In Sec. V we shall give a general form for logical Bell inequalities.

B. A curious observation

Quantum mechanics tells us that we can find propositions φ_i describing outcomes of certain measurements, which not only *can* but *have* been performed. From the observed statistics of these experiments, we have *very* highly confirmed probabilities p_i. These propositions are easily seen to be jointly contradictory. Nevertheless, the inequality,

$$\sum_i p_i \leqslant N - 1,$$

is observed to be *strongly violated*. In fact, the maximum violation of 1 can be achieved [17].

How can this be?

The best resolution to this puzzle on offer is that each formula φ_i involves a proper subset X_i of the total set X of Boolean variables which appear in the family, and hence in the conjunction Φ. There is no global assignment of probabilities to all the variables X simultaneously which yields the empirically observed probabilities. Hence the ascription of a probability to Φ is the invalid step. This is given general mathematical meaning in terms of an obstruction to the existence of a global section in Refs. [3,18], extending [19].

This does seem an uncomfortably slender basis on which to defend logical consistency, since it seems hard to avoid the conclusion that the null event should be assigned probability 0.

This argument can be seen as a theory-independent derivation of the impossibility of measuring all the variables in X simultaneously, even in principle, on pain of a direct clash between logical consistency and empirical evidence. We simply cannot regard the variables as each representing a global, context-independent quantity.

C. Logical Bell and CHSH inequalities

We shall call the inequality,

$$\sum_i p_i \leqslant N - 1,$$

a logical Bell inequality. We can also derive an associated inequality for expectations.

We shall associate truth of a formula with the value $+1$, and falsity with -1. We then have the expected value E_i of the formula φ_i given by

$$E_i = (+1)p_i + (-1)(1 - p_i) = 2p_i - 1.$$

From the Bell inequality, we obtain

$$\begin{aligned}\sum_i E_i &= \sum_i (2p_i - 1)\\ &= 2\sum_i p_i - N \leqslant 2(N-1) - N = N - 2.\end{aligned}$$

Moreover, if K is an upper bound as the expectations range over probability assignments, $-K$ must be a lower bound, as we can substitute $1 - p_i$ for p_i to get the expected value $-E_i$. Thus this is a bound on the absolute value of the expectations, so we obtain the logical CHSH inequality:

$$\left|\sum_i E_i\right| \leqslant N - 2. \tag{1}$$

Note that these inequalities are very general, and independent of any particular setting. We shall now show how they apply to familiar scenarios arising from quantum mechanics and the study of nonlocality.

II. PROBABILISTIC MODELS OF EXPERIMENTS

Our general setting will be the probability models commonly studied in quantum information and quantum foundations [20]. In these models, a number of agents each has the choice of one of several measurement settings; and each measurement has a number of distinct outcomes. For most of this paper, we shall focus on measurements with two possible outcomes; however, we will show how our results can be extended to measurements with multiple outcomes in Sec. VIII. For each choice of a measurement setting by each of the agents, we have a probability distribution on the joint outcomes of the measurements.

For example, consider the following tabulation of such a model.

	(0,0)	(1,0)	(0,1)	(1,1)
(a,b)	1/2	0	0	1/2
(a,b')	3/8	1/8	1/8	3/8
(a',b)	3/8	1/8	1/8	3/8
(a',b')	1/8	3/8	3/8	1/8

Here we have two agents, Alice and Bob. Alice can choose from the settings a or a', and Bob can choose from b or b'. These choices correspond to the rows of the table. The columns correspond to the joint outcomes for a given choice of settings by Alice and Bob, the two possible outcomes for each individual measurement being represented by 0 and 1. The numbers along each row specify a probability distribution on these joint outcomes.

A. The Bell model

A standard version of Bell's theorem uses the probability table given above. This table can be realized in quantum mechanics (e.g., by a Bell state), written in the Z basis as

$$\frac{|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle}{\sqrt{2}},$$

subjected to spin measurements in the XY plane of the Bloch sphere, at a relative angle of $\pi/3$.

Logical analysis of the Bell table

We now pick out a subset of the elements of each row of the table, as indicated in the following table.

	(0, 0)	(1, 0)	(0, 1)	(1, 1)
(a,b)	$\boxed{1/2}$	0	0	$\boxed{1/2}$
(a,b')	$\boxed{3/8}$	1/8	1/8	$\boxed{3/8}$
(a',b)	$\boxed{3/8}$	1/8	1/8	$\boxed{3/8}$
(a',b')	1/8	$\boxed{3/8}$	$\boxed{3/8}$	1/8

If we read 0 as true and 1 as false, the boxed positions in the table are represented by the following propositions:

$$\begin{aligned}\varphi_1 &= (a \wedge b) \quad\ \vee (\neg a \wedge \neg b) \ = a \leftrightarrow b,\\ \varphi_2 &= (a \wedge b') \quad\ \vee (\neg a \wedge \neg b') = a \leftrightarrow b',\\ \varphi_3 &= (a' \wedge b) \quad\ \vee (\neg a' \wedge \neg b) = a' \leftrightarrow b,\\ \varphi_4 &= (\neg a' \wedge b') \vee (a' \wedge \neg b') \ = a' \oplus b'.\end{aligned}$$

The first three rows are the correlated outcomes; the fourth is anticorrelated. These propositions are easily seen to be contradictory. Indeed, starting with φ_4, we can replace a' with b using φ_3, b with a using φ_1, and a with b' using φ_2, to obtain $b' \oplus b'$, which is obviously unsatisfiable.

We see from the table that $p_1 = 1$, $p_i = 6/8$ for $i = 2,3,4$. Hence the violation of the Bell inequality is $1/4$, and of the CHSH inequality $1/2$.

We may note that the logical pattern shown by this jointly contradictory family of propositions underlies the familiar CHSH correlation function.

B. Some notation

Later we will develop some notation for the general case. To prepare the way we will indicate how this notation will work in the Bell model. First, in this case, we put

$$X = \{a,b,a',b'\},$$

as this is the set of Boolean variables we are interested in. Next, we consider subsets $U \subseteq X$, corresponding to the different combinations of measurements we might perform—the *measurement contexts*. One such subset is $U = \{a,b\}$. We denote the set of all such subsets by $\mathcal{U}$. Thus, in this case, we have

$$\mathcal{U} = \{\{a,b\},\{a,b'\},\{a',b\},\{a',b'\}\}.$$

A basic measurement such as a has possible outcomes 0 or 1. We shall write $\mathbf{2} := \{0,1\}$ for the set of possible outcomes. A joint outcome for a set of measurements U can be specified by a function $s : U \to \mathbf{2}$. For example, if we perform the measurements in $U = \{a,b\}$, and a has outcome 0 and b has outcome 1, this is described by the function,

$$\{a \mapsto 0,\ b \mapsto 1\},$$

which maps a to 0 and b to 1. This function corresponds to the cell in the first row and third column of the Bell table. The set of all such functions is denoted by $\mathbf{2}^U$. Thus for $U = \{a,b\}$,

$$\mathbf{2}^U = \{f_{ij} \mid i,j = 0,1\},$$

where $f_{ij} = \{a \mapsto i, b \mapsto j\}$. This corresponds to the set of cells in the first row of the table.

A probability model such as the Bell table shown above is given by specifying a probability distribution d_U on $\mathbf{2}^U$ for each $U \in \mathcal{U}$. Thus d_U is a function $d_U : \mathbf{2}^U \to [0,1]$ such that $\sum_{s \in \mathbf{2}^U} d_U(s) = 1$. These distributions correspond to the rows of the Bell table.

The proposition φ_1 pertains to the context $U = \{a,b\}$; note that it only uses the variables in U. We can think of $\mathbf{2}^U$ as the set of truth-value assignments to the Boolean variables in U, where we interpret 0 as true and 1 as false. The set of satisfying assignments for the formula φ_1—the subset of $\mathbf{2}^U$ for which this proposition is true—is

$$S(U) = \{\{a \mapsto 0,\ b \mapsto 0\}, \{a \mapsto 1,\ b \mapsto 1\}\}.$$

We have given such a proposition φ_i for each element of $\mathcal{U}$. The highlighted items in the ith row of the table form the set $S(U_i)$ of satisfying assignments for φ_i, where U_i is the corresponding measurement context.

C. A bipartite logical model

We know turn to the model introduced by one of us in 1992 [12,14]. The original purpose of this construction was to show a "logical" proof of Bell's theorem in the bipartite case, following the GHZ tripartite construction. Reflecting this, we shall only need to consider the *support table* of the model to demonstrate a violation of the inequalities.

Consider, for example, the following table, which has a quantum realization as described in Ref. [14].

	(0,0)	(1,0)	(0,1)	(1,1)
(a,b)	1	1	1	1
(a',b)	0	1	1	1
(a,b')	0	1	1	1
(a',b')	1	1	1	0

This table has a 1 for every entry in the model with a positive probability.

If we interpret outcome 0 as true and 1 as false, then the following formulas all have positive probability:

$$a \wedge b, \quad \neg(a \wedge b'), \quad \neg(a' \wedge b), \quad a' \vee b'.$$

However, these formulas are not simultaneously satisfiable.

Note that the formulas φ_i for $i = 2,3,4$ describe the full support of this model for the corresponding rows; hence $p_2 = p_3 = p_4 = 1$. It follows that the model achieves a violation of $p_1 = \text{Prob}(a \wedge b)$ for the Bell inequality, and a violation of $2p_1$ for the CHSH inequality.

Note that this calculation can be made purely on the basis of the support table.

III. THE GENERAL CASE: STRUCTURE OF SUPPORTS

We now turn to a general analysis. The setting will be that of [3], but we shall develop what we need in a self-contained fashion.

We shall begin by looking just at the *supports* of probability models, which suffice to describe many forms of contextual and nonlocal behavior, as we have already illustrated with the model described in Sec. II C. We shall then go on to look at generalized probability models themselves.

A. Notation

We shall use the notation introduced in the previous section: We define $\mathbf{2} := \{0,1\}$, and write $\mathbf{2}^U$ for the set of all functions from a set U into $\mathbf{2}$. We shall also use the following notation for function restriction. If $s : X \to \mathbf{2}$ is a function, and $U \subseteq X$, then we write $s|U : U \to \mathbf{2}$ for the restriction of s to U. For example, if $X = \{a,b,a',b'\}$, $U = \{a,b\}$, and $s : X \to \mathbf{2}$ is the function,

$$\{a \mapsto 0,\ b \mapsto 1,\ a' \mapsto 1,\ b' \mapsto 0\},$$

then $s|U$ is the function,

$$\{a \mapsto 0,\ b \mapsto 1\}.$$

B. Structure of support tables

We fix a set of Boolean variables X, and a *cover* $\mathcal{U}$ (i.e., a family of subsets of X such that $\bigcup \mathcal{U} = X$).

A *probability model* on a cover $(X,\mathcal{U})$ is a family $\{d_U\}_{U \in \mathcal{U}}$, where d_U is a probability distribution on $\mathbf{2}^U$.

We think of the sets $U \in \mathcal{U}$ as the *compatible sets of measurements*, which index the "rows" of the probability table. Given such a row U, $\mathbf{2}^U$ is the set of possible joint outcomes of these measurements. The distribution d_U gives the probability for each such joint outcome.

The *support* of the model at $U \in \mathcal{U}$ is the set $S(U) \subseteq \mathbf{2}^U$ of those $s \in \mathbf{2}^U$ such that $d_U(s) > 0$.

A *global section* for the support of the model is an assignment,

$$s : X \to \mathbf{2},$$

such that, for all $U \in \mathcal{U}$, $s|U \in S(U)$ [21].

We can think of global sections in geometric terms, as coherently gluing together a family of local sections $s_U \in S(U)$, indexed by $U \in \mathcal{U}$. This geometrical idea of a global section can be related to logical notions. A formula φ_U over a set of variables $U \in \mathcal{U}$ has a set of satisfying assignments which is a subset of $\mathbf{2}^U$. Note that, if U is finite, *any* subset of $\mathbf{2}^U$ can be defined in this way by a propositional formula. For each $U \in \mathcal{U}$, let φ_U be a formula whose set of satisfying assignments is $S(U)$. Global sections correspond precisely to satisfying assignments for the formula,

$$\varphi = \bigwedge_{U \in \mathcal{U}} \varphi_U.$$

As shown in detail in Ref. [3], the existence of global sections provides a canonical form for *noncontextual hidden-variable theories*.

We can define a probabilistic model to be *possibilistically noncontextual* [3] if for every element $s \in S(U)$ of its support, there is a global section s' such that $s'|U = s$. If this does not hold, the model is *contextual*, or in particular *nonlocal*. In fact, as shown in Ref. [3], this form of contextuality or nonlocality is strictly stronger than the usual probabilistic notions. For example, the Bell model studied in the previous section is nonlocal, but is in fact possibilistically noncontextual. The possibilistically contextual models are those which admit logical proofs of Bell's theorem: "Bell's theorem without inequalities" [11].

We can now give a completely general argument that for any model which is contextual in this strong possibilistic sense, we can obtain a violation of instances of the generalized Bell and CHSH inequalities.

Proposition 1. Any possibilistically contextual model violates a logical Bell or CHSH inequality.

Proof. Suppose that a model is possibilistically contextual, with $s \in S(U)$ such that there is no global section for S restricting to s. We define a formula φ_s, describing s. This formula can be written explicitly as

$$\varphi_s := \bigwedge_{s(x)=0} x \wedge \bigwedge_{s(x)=1} \neg x. \tag{2}$$

The only satisfying assignment for φ_s in $\mathbf{2}^U$ is s.

For all $U' \in \mathcal{U}$ with $U' \neq U$, we define $\varphi_{U'}$ to be a formula which defines the support of the model on the "row" U'. Explicitly, we can define

$$\varphi_{U'} := \bigvee_{s' \in S(U')} \varphi_{s'}.$$

The fact that there is no global section on the support which restricts to s says exactly that the formula $\varphi_s \wedge \bigwedge_{U' \neq U} \varphi_{U'}$ is not satisfiable. Since $p_{U'} = 1$ for $U \neq U' \in \mathcal{U}$, the Bell inequality with respect to these formulas is violated by $p_{\varphi_s} = p(s) > 0$, while violation of the CHSH inequality is by $2p(s)$. ■

C. Strong contextuality

A still stronger form of contextuality is identified in Ref. [3]. A model is defined to be *strongly contextual* if its support has no global section; equivalently, the propositional formulas defining its support are not simultaneously satisfiable.

It is shown in Ref. [3] that all n-partite states GHZ(n), for $n \geqslant 3$, are strongly contextual in this sense. It is also shown that strong contextuality is equivalent to the model being maximally contextual, in the sense of having no nontrivial convex decomposition into a noncontextual model and a no-signaling model.

We now have the following result.

Proposition 2. A model achieves maximal violation of a logical Bell inequality if and only if it is strongly contextual.

Proof. Suppose that the model is strongly contextual. For each row U, we can define the formula φ_U corresponding to the support of the model on that row, as in the proof of the previous proposition. Since the probability of each φ_U is 1, we obtain the maximum violation of 1.

For the converse, if maximal violation is achieved, there are a family of rows $U_1, \dots, U_N$, and propositions φ_i defining subsets $S(U_i) \subseteq \mathbf{2}^{U_i}$, such that $\bigwedge_i \varphi_i$ is unsatisfiable, and $\sum_i p_i = N$. This implies that $p_i = 1$ for all i, and hence that $S(U_i)$ contains the support of the model on U_i. The unsatisfiability of $\bigwedge_i \varphi_i$ means that there is no global section which restricts to each $S(U_i)$, which means *a fortiori* that the model is strongly contextual. ■

1. Example: the GHZ state

We consider the tripartite GHZ state [11,22], which we write in the Z basis as

$$\frac{|\uparrow\uparrow\uparrow\rangle + |\downarrow\downarrow\downarrow\rangle}{\sqrt{2}},$$

with X and Y measurements in each component. The relevant part of the support table for the resulting probability model can be specified as follows:

	000	001	010	011	100	101	110	111
abc	1	0	0	1	0	1	1	0
$ab'c'$	0	1	1	0	1	0	0	1
$a'bc'$	0	1	1	0	1	0	0	1
$a'b'c$	0	1	1	0	1	0	0	1

Given Boolean variables x, y, z, we define

$$\Psi_{xyz} := \neg x \oplus \neg y \oplus \neg z. \tag{3}$$

The support for each row can be specified by the following formulas:

$$\varphi_1 := \neg\Psi_{abc}, \quad \varphi_2 := \Psi_{ab'c'}, \quad \varphi_3 := \Psi_{a'bc'}, \quad \varphi_4 := \Psi_{a'b'c}.$$

It can be verified that these formulas are not simultaneously satisfiable; in fact, such a verification is what the well-known argument by Mermin in terms of "instruction sets" [23] amounts to.

Thus the tripartite GHZ state maximally violates a logical Bell inequality. Similar arguments apply to n-partite GHZ states for all $n > 3$; see Ref. [3].

2. Example: the PR box

We consider the Popescu-Rohrlich box [24], which achieves superquantum correlations while respecting no-signaling.

	(0,0)	(0,1)	(1,0)	(1,1)
(a,b)	1	0	0	1
(a,b')	1	0	0	1
(a',b)	1	0	0	1
(a',b')	0	1	1	0

The supports of the rows of this table are specified by the following formulas:

$$a \leftrightarrow b, \quad a \leftrightarrow b', \quad a' \leftrightarrow b, \quad a' \oplus b',$$

which are not simultaneously satisfiable. Thus this model maximally violates a logical Bell inequality.

Note that these formulas are the same as those we used for the Bell model in Sec. II A. In this case, however, they cover the whole support of the model, corresponding to the fact that the PR box attains the algebraic maximum of the CHSH correlation function.

D. Kochen-Specker configurations

The notion of model we are considering, following [3], is much more general than the usual "Bell scenarios." For example, any set $\mathcal{X}$ of quantum observables gives rise to a cover in our sense, where the sets in the cover correspond to the maximal compatible subsets of $\mathcal{X}$. Since we are currently restricting our attention to two-outcome measurements, we shall only consider dichotomic observables. If we fix a state, then for each maximal set of compatible observables, (i.e., each row of the table), we get a probability distribution on joint outcomes of the observables in the family, following the usual quantum mechanical recipe. The details are spelled out in Ref. [3].

The usual Bell case arises when the observables are partitioned according to the sites or parties; the sets in the cover correspond to a choice of one observable from each site, represented on a tensor product in the standard fashion.

Equally, however, any Kochen-Specker configuration gives rise to a cover in our sense [25]. Given a family of unit vectors representing distinct rays in $\mathbb{R}^d$, we consider the corresponding dichotomic observables, whose spectral resolutions project onto the ray and its orthogonal. We shall label the outcome corresponding to the ray as 0, and the orthogonal outcome as 1.

These observables are compatible if and only if the corresponding rays are orthogonal. Thus the maximal compatible families correspond to the families of vectors which determine orthonormal bases of $\mathbb{R}^d$. It follows that, for any quantum state, the only possible outcomes for one of these maximal compatible families are those where exactly one of the outcomes is labeled 0. Thus for any state, the support of the probability model it gives rise to satisfies the following formula for each set U in the cover:

$$\Theta(U) := \bigvee_{x\in U}(x \wedge \bigwedge_{x'\in U\setminus\{x\}} \neg x').$$

The essential property of Kochen-Specker configurations is exactly that there is no global section for this family of supports, or equivalently, that the formula,

$$\bigwedge_{U\in\mathcal{U}} \Theta(U),$$

is unsatisfiable. It follows immediately that, given a Kochen-Specker configuration, the probability model generated by *any* quantum state with respect to the corresponding family of observables is strongly contextual. This fully explicates the state-independent nature of the Kochen-Specker theorem.

Hence we obtain the following corollary to Proposition 2.

Proposition 3. For any Kochen-Specker configuration, and for any quantum state, the corresponding probability model maximally violates a logical Bell inequality.

Thus we have a perfectly general way of obtaining experimentally testable inequalities, with maximal violations, from any Kochen-Specker configuration.

1. Example: the 18-vector configuration in $\mathbb{R}^4$

We look at the 18-vector construction in $\mathbb{R}^4$ from [26]. This uses the following measurement cover $\mathcal{U} = \{U_1, \ldots, U_9\}$, where the columns U_i are the sets in the cover.

U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9
A	A	H	H	B	I	P	P	Q
B	E	I	K	E	K	Q	R	R
C	F	C	G	M	N	D	F	M
D	G	J	L	N	O	J	L	O

The standard argument that this is a Kochen-Specker configuration [3,26] amounts to verifying that the formula,

$$\bigwedge_{i=1}^{9} \Theta(U_i),$$

is unsatisfiable. Thus for any quantum state, the resulting probability model will maximally violate a logical Bell inequality.

2. Example: the Peres-Mermin square

We look at an important example, the Peres-Mermin square [27,28], which can be realized in quantum mechanics using two-qubit observables.

The structure of the square is as follows:

A	B	C
D	E	F
G	H	I

The compatible families of measurements are the rows and columns of this table. The key property differs from the usual Kochen-Specker situation in that we don't ask for exactly one 1 at each maximal context. Instead, we ask that each "row context" has an odd number of 1's whereas each "column context" has an even number of 1's. Hence the support table is the following.

	000	001	010	011	100	101	110	111
ABC	0	1	1	0	1	0	0	1
DEF	0	1	1	0	1	0	0	1
GHI	0	1	1	0	1	0	0	1
ADG	1	0	0	1	0	1	1	0
BEH	1	0	0	1	0	1	1	0
CFI	1	0	0	1	0	1	1	0

Note that the first three lines correspond to the row contexts and the remaining three to the column contexts from the square.

The following formulas characterize the supports for each line of the table:

$$\varphi_1 := \Psi_{ABC}, \quad \varphi_2 := \Psi_{DEF}, \quad \varphi_3 := \Psi_{GHI},$$
$$\varphi_4 := \neg\Psi_{ADG}, \quad \varphi_5 := \neg\Psi_{BEH}, \quad \varphi_6 := \neg\Psi_{CFI}.$$

Here we use Ψ_{xyz} as defined in Eq. (3).

It can be verified that these formulas are not simultaneously satisfiable. Thus the Peres-Mermin square maximally violates a logical Bell inequality.

IV. GENERAL PROBABILISTIC MODELS

Suppose we are given a cover $\mathcal{U}$ on a set X. A general probability model over $\mathcal{U}$ assigns a probability distribution d_U on the set $\mathbf{2}^U$ for each $U \in \mathcal{U}$ [29].

Each global assignment $t \in \mathbf{2}^X$ induces a deterministic probability model δ^t:

$$\delta^t_U(s) = \begin{cases} 1, & t|U = s \\ 0 & \text{otherwise.} \end{cases}$$

We have the following result from [3, Theorem 8.1].

Theorem 1. A probability model $\{d_U\}$ is noncontextual if and only if it can be written as a convex combination $\sum_{j\in J} \mu_j \delta^{t_j}$ where $t_j \in \mathbf{2}^X$ for each $j \in J$. This means that for each $U \in \mathcal{U}$,

$$d_U = \sum_j \mu_j \delta^{t_j}_U.$$

In fact, this gives a canonical form for such models, subsuming the usual notions of local or noncontextual hidden-variable models.

V. THE GENERAL FORM OF LOGICAL BELL INEQUALITIES

It will be useful to establish some notation for expressing logical Bell inequalities. Suppose we are given a cover $\mathcal{U}$ on a set X. As illustrated in the examples we have looked at previously, we will regard X as a set of Boolean variables. We shall consider expressions of the form,

$$\sum_{i=1}^{N} k_i \varphi_i,$$

where for each i, k_i is a non-negative integer, and φ_i is a formula whose variables are drawn from $U_i \in \mathcal{U}$.

We think of such expressions as multisets of formulas, where φ_i appears with multiplicity k_i. A sub-multiset of $\sum_{i\in I} k_i\varphi_i$ is an expression of the form $\sum_{i\in I} k'_i\varphi_i$, where for each i, $0 \leqslant k'_i \leqslant k_i$. The *cardinality* of $\sum_{i\in I} k_i\varphi_i$ is $\sum_{i\in I} k_i$. We say that $\sum_{i\in I} k_i\varphi_i$ is K-consistent if for every sub-multiset of cardinality $>K$, the underlying set of formulas with positive support has no satisfying assignment.

Given a positive integer K, we consider the expression,

$$\sum_{i=1}^{N} k_i\, p(\varphi_i) \leqslant K. \tag{4}$$

If we are given a probability model $\{d_U\}_{U\in\mathcal{U}}$, we can evaluate the formal expression $p(\varphi_i)$ as $p_i := d_{U_i}(S_i)$, where S_i is the set of satisfying assignments in $\mathbf{2}^{U_i}$ for φ_i (i.e., the event defined by φ_i).

The model *satisfies* the expression (4) if

$$\sum_{i=1}^{N} k_i p_i \leqslant K.$$

Proposition 4. The inequality (4) is satisfied by all noncontextual models if and only if the multiset $\sum_{i\in I} k_i\varphi_i$ is K-consistent.

Proof. By Theorem 1, any noncontextual model can be written as a convex combination $\sum_j \mu_j \delta^{t_j}$, where $t_j \in \mathbf{2}^X$.

It suffices to verify (4) for the deterministic models δ^t, since if for each j we have $\sum_i k_i p_{i,j} \leqslant K$, where $p_{i,j} = \delta^{t_j}_{U_i}(S_i)$, then,

$$\sum_i k_i \left(\sum_j \mu_j p_{i,j}\right) = \sum_j \mu_j \left(\sum_i k_i p_{i,j}\right) \leqslant \sum_j \mu_j K = K.$$

Now if the multiset $\sum_{i\in I} k_i\varphi_i$ is K-consistent, any $t \in \mathbf{2}^X$, viewed as a Boolean assignment on X, can satisfy a sub-multiset of cardinality at most K, and hence δ^t will satisfy the inequality (4).

Conversely, if t satisfies $K + 1$ formulas in the multiset, each corresponding term in Eq. (4) will be assigned probability 1 in δ^t, and hence we will have $\sum_i k_i p_i \geqslant K + 1$. ■

Note that the form of logical inequality which we have considered previously is a special case, where we have a set of N formulas which is $(N-1)$ consistent. Allowing for the more general notion of K-consistency leads to sharper inequalities, which are needed to obtain completeness.

It is important to note that there is no requirement for the sets U_i to be distinct. Thus different formulas occurring in the multiset may define overlapping subsets of the same row.

We define the general notion of logical Bell inequality over a cover $\mathcal{U}$ to be given by expressions of the form (4), where the multiset of formulas is K-consistent. Note that this class of inequalities is defined solely in terms of the cover $\mathcal{U}$, and a purely logical condition on the formulas. Thus we may indeed regard this as a logical class; the interesting point is that we can obtain quantitative information about contextuality from conditions which are derived in a purely logical fashion.

VI. COMPLETENESS OF LOGICAL BELL INEQUALITIES

We shall now show that logical Bell inequalities completely characterize contextuality.

We begin by recalling the definition of the *incidence matrix* from [3]. Given a cover $\mathcal{U}$ on a set X, we define a matrix $\mathbf{M}$ whose rows are indexed by pairs (U,s), where $U \in \mathcal{U}$, and $s \in \mathbf{2}^U$, and whose columns are indexed by global assignments $t \in \mathbf{2}^X$. The matrix entries are defined by

$$\mathbf{M}[(U,s),t] = \begin{cases} 1, & t|U = s \\ 0 & \text{otherwise.} \end{cases}$$

Note that the column $\mathbf{M}[\_,t]$ of the matrix corresponds to the deterministic model δ^t. We can regard a probabilistic model $\{d_U\}_{U\in\mathcal{U}}$ as a real vector $\mathbf{v}$ in the row space of $\mathbf{M}$, where $\mathbf{v}[U,s] = d_U(s)$.

Proposition 5. The noncontextuality of the probabilistic model represented by the vector $\mathbf{v}$ is equivalent to the existence of a non-negative solution $\mathbf{x} \geqslant \mathbf{0}$ for the linear system

$$\mathbf{M}\mathbf{x} = \mathbf{v}.$$

Proof. For each $U \in \mathcal{U}$, the subvector $\mathbf{v}_U$ of $\mathbf{v}$ forms a probability distribution on $\mathbf{2}^U$, and hence sums to 1. Since the restriction map $\mathbf{2}^X \to \mathbf{2}^U$ is surjective, and $\mathbf{M}\mathbf{x} = \mathbf{v}$ implies $(\mathbf{M}\mathbf{x})_U = \mathbf{v}_U$, it follows that the entries of $\mathbf{x}$ sum to 1. Thus $\mathbf{x}$ defines a probability distribution μ on $\mathbf{2}^X$. Moreover, the equation $(\mathbf{Mx})_U = \mathbf{v}_U$ is equivalent to

$$d_U = \sum_{t \in \mathbf{2}^X} \mu(t)\delta^t_U.$$

Thus a solution $\mathbf{x}$ exists if and only if the model can be written as a convex combination as in Theorem 1. ■

Thus the set $\mathcal{N}$ of noncontextual probability models is the convex hull of the set of deterministic models δ^t, $t \in \mathbf{2}^X$. By the fundamental properties of convex polytopes [30–32], $\mathcal{N}$ is equivalently specified by a finite set of linear inequalities.

To state this more explicitly, we first recall the well-known Fourier-Motzkin elimination procedure [30–32].

Proposition 6. (Fourier-Motzkin elimination). If we are given a finite system $I(\mathbf{x},\mathbf{y})$ of linear inequalities in the variables $\mathbf{x}$, $\mathbf{y}$, we can effectively obtain a finite system $J(\mathbf{y})$ of inequalities in the variables $\mathbf{y}$, such that $\mathbf{v}$ satisfies J if and only if for some $\mathbf{w}$, $(\mathbf{w},\mathbf{v})$ satisfies I. Moreover, J is constructed from I using only the field operations, so if I is rational, so is J.

The size of J is, in the worst case, doubly exponential in the size of I. Nevertheless, Fourier-Motzkin elimination is widely used in computer-assisted verification and polyhedral computation [33,34].

In our case, we begin with the "symbolic" system,

$$\mathbf{Mx} = \mathbf{y}, \quad \mathbf{x} \geqslant \mathbf{0}, \quad \mathbf{1} \cdot \mathbf{x} = 1,$$

in variables $\mathbf{x}$, $\mathbf{y}$. This can be written as

$$\begin{aligned} a_{1,j}x_1 + \cdots + a_{N,j}x_N - y_j &\geqslant 0, \quad j = 1,\ldots,D,\\ -a_{1,j}x_1 + \cdots + -a_{N,j}x_N + y_j &\geqslant 0, \quad j = 1,\ldots,D,\\ x_i &\geqslant 0, \quad i = 1,\ldots,N,\\ x_1 + \cdots + x_N &\geqslant 1\\ -x_1 + \cdots + -x_N &\geqslant -1, \end{aligned}$$

where the coefficients $a_{i,j}$ come from the incidence matrix $\mathbf{M}$, and

$$N := 2^{|X|}, \quad D := \sum_{U \in \mathcal{U}} 2^{|U|}$$

are the dimensions of $\mathbf{M}$. Note that, since the system is symbolic, we have to add the constraint that $\mathbf{x}$ sums to 1 explicitly.

Writing this system as $I(\mathbf{x},\mathbf{y})$, by Proposition 5, we have

$$\mathcal{N} = \{\mathbf{v} \mid \exists\, \mathbf{w},\, I(\mathbf{w},\mathbf{v})\}.$$

By Proposition 6, we can eliminate the variables $\mathbf{x}$ from this system, producing a system J of inequalities in the variables $\mathbf{y}$, such that $\mathbf{v}$ satisfies J if and only if for some $\mathbf{w}$, $(\mathbf{w},\mathbf{v})$ satisfies I. Thus $\mathbf{v}$ is in $\mathcal{N}$ if and only if $\mathbf{v}$ satisfies J.

Thus we obtain the following result.

Proposition 7. There is a finite set of rational vectors $\mathbf{r}_1,\ldots,\mathbf{r}_p$, and rational numbers $r_1,\ldots,r_p$, such that, for all probability models $\mathbf{v}$,

$$\mathbf{v} \in \mathcal{N} \iff \forall\, i = 1,\ldots,p, \quad \mathbf{r}_i \cdot \mathbf{v} \leqslant r_i.$$

A. Completeness of logical Bell inequalities

Suppose we are given a cover $\mathcal{U}$. A rational inequality over $\mathcal{U}$ is given by a rational vector $\mathbf{r}$ and a rational number r. A probability model $\mathbf{v}$ satisfies this inequality if $\mathbf{r} \cdot \mathbf{v} \leqslant r$. Two inequalities are equivalent if they are satisfied by the same probability models.

Theorem 2. A rational inequality is satisfied by all noncontextual models over $\mathcal{U}$ if and only if it is equivalent to a logical Bell inequality.

Proof. A rational inequality determines an equivalent integer inequality given by an integer vector $\mathbf{k}$ and an integer M, obtained by clearing denominators.

Suppose that we are given an integer vector $\mathbf{k}$ indexed by (U,s), where $U \in \mathcal{U}$ and $s \in \mathbf{2}^U$. For each (U,s), we define non-negative integers k^U_s, and formulas θ^U_s in the variables U:

$$\begin{aligned} k^U_s &= |\mathbf{k}[U,s]|,\\ \theta^U_s &= \begin{cases} \varphi_s, & \mathbf{k}[U,s] \geqslant 0\\ \neg\varphi_s, & \mathbf{k}[U,s] < 0. \end{cases} \end{aligned}$$

Here we use φ_s as defined in Eq. (2).

Now suppose we are given a probability model $\mathbf{v}$. For each (U,s), we define p^U_s to be the probability assigned by $\mathbf{v}$ to the subset of $\mathbf{2}^U$ defined by θ^U_s.

We claim that

$$\mathbf{k} \cdot \mathbf{v} = \sum_{U,s} k^U_s p^U_s - \sum_{\mathbf{k}[U,s]<0} k^U_s. \tag{5}$$

To see this, for each (U,s) we compare $\mathbf{k}[U,s] \cdot \mathbf{v}[U,s]$ with $k^U_s p^U_s$:

If $\mathbf{k}[U,s] \geqslant 0$, then $\mathbf{k}[U,s] \cdot \mathbf{v}[U,s] = k^U_s p^U_s$.

If $\mathbf{k}[U,s] < 0$, we have the following:

$$\mathbf{k}[U,s] \cdot \mathbf{v}[U,s] = k^U_s\{[1 - p(\varphi_s)] - 1\} = k^U_s\left(p^U_s - 1\right).$$

Collecting terms, we obtain (5).

We now consider the expression,

$$\sum_{U,s} k^U_s\, p\left(\theta^U_s\right) \leqslant K, \tag{6}$$

where $K = M + \sum_{\mathbf{k}[U,s]<\mathbf{0}} k^U_s$.

By (5), a probability model $\mathbf{v}$ will satisfy this inequality if and only if $\mathbf{k} \cdot \mathbf{v} \leqslant M$. Thus this inequality is equivalent to the rational inequality we began with.

Now suppose that this inequality is satisfied by all noncontextual models. Since the coefficients k^U_s in Eq. (6) are non-negative, K must be non-negative. By Proposition 4, the multiset of formulas $\sum_{U,s} k^U_s \theta^U_s$ is K-consistent, and thus (6) is a logical Bell inequality.

Thus every rational inequality satisfied by all noncontextual models is equivalent to a logical Bell inequality. From Proposition 4, we also have the converse: Every logical Bell inequality is satisfied by all noncontextual models.

Combining Proposition 7 and Theorem 2, we obtain the following completeness result.

Theorem 3. The polytope of noncontextual probability models over any cover $\mathcal{U}$ is determined by a finite set of logical Bell inequalities. Moreover, these inequalities can be obtained effectively from $\mathcal{U}$. Thus a probabilistic model over any cover is contextual if and only if it violates one of finitely many logical Bell inequalities.

Proof. By Proposition 7, given $\mathcal{U}$ we can effectively obtain a finite set of rational inequalities defining the noncontextual polytope. Using the construction given in the proof of

Theorem 2, we can effectively transform these rational inequalities into equivalent logical Bell inequalities. ■

VII. LOGICAL DESCRIPTION OF CORRELATION INEQUALITIES

We shall now show that correlation inequalities can also be analyzed logically; in fact, they form a special case of the logical inequalities we have already described.

A probability model $\mathbf{v}$ determines a vector $\boldsymbol{\eta}^{\mathbf{v}} = (E_U)_{U \in \mathcal{U}}$ of expectation values. Here,

$$E_U := (+1)p(\psi_U) + (-1)p(\neg\psi_U),$$

where ψ_U is a formula whose satisfying assignments are those with an even number of 1 outcomes. Thus we can define

$$\psi_U := \neg \bigoplus_{x \in U} \neg x. \tag{7}$$

The set of expectation vectors of noncontextual models is the image under a linear map of the convex polytope of noncontextual probability models, and hence forms a convex polytope $\mathcal{E}$, with vertices given by the vectors $\boldsymbol{\eta}^t$, $t \in \mathbf{2}^X$.

Clearly, any probability model $\mathbf{v}$ whose expectation vector $\boldsymbol{\eta}^{\mathbf{v}}$ is not in $\mathcal{E}$ must be contextual. However, the converse is *not* the case. We shall return to this point in Sec. VII B. Nevertheless, the correlation inequalities have received a great deal of attention in the literature on nonlocality, and it is clearly of considerable interest to give a complete characterization.

We shall now give a logical characterization of a complete set of inequalities for the polytope $\mathcal{E}$ on an arbitrary cover $\mathcal{U}$.

Theorem 4. For any probability model $\mathbf{v}$ such that $\boldsymbol{\eta}^{\mathbf{v}} \notin \mathcal{E}$, there is a logical Bell inequality,

$$\sum_{U \in \mathcal{U}} k_U\, p(\theta_U) \leqslant K, \tag{8}$$

which is violated by $\mathbf{v}$, where for each U, θ_U is either ψ_U or $\neg\psi_U$.

Proof. By similar reasoning to that used in the proof of Theorem 3, there is an integer vector $\mathbf{k}$ and an integer M such that $\mathbf{k} \cdot \boldsymbol{\eta}^{\mathbf{w}} \leqslant M$ for all noncontextual models $\mathbf{w}$, and $\mathbf{k} \cdot \boldsymbol{\eta}^{\mathbf{v}} > M$.

For each U, and any probability model $\mathbf{w}$, we consider two cases:

If $\mathbf{k}[U] = k_U$ is positive, then we can write $\mathbf{k}[U] \cdot \boldsymbol{\eta}^{\mathbf{w}}[U] = k_U[2p(\psi_U) - 1]$.

If $\mathbf{k}[U] = -k_U$ is negative, we can write

$$\begin{aligned}\mathbf{k}[U] \cdot \boldsymbol{\eta}^{\mathbf{w}}[U] &= -k_U[2p(\psi_U) - 1] = k_U[1 - 2p(\psi_U)] \\ &= k_U\{2[1-p(\psi_U)]-1\} = k_U[2p(\neg\psi_U) - 1].\end{aligned}$$

Rearranging terms, we have

$$\mathbf{k} \cdot \boldsymbol{\eta}^{\mathbf{w}} = \sum_{U \in \mathcal{U}} 2k_U\, p(\theta_U) - P,$$

where each θ_U is either ψ_U or $\neg\psi_U$, and P is a positive integer. Hence the inequality $\mathbf{k} \cdot \boldsymbol{\eta}^{\mathbf{w}} \leqslant M$ is equivalent to $\mathbf{w}$ satisfying the inequality,

$$\sum_{U \in \mathcal{U}} 2k_U\, p(\theta_U) \leqslant K, \tag{9}$$

where $K = M + P$. By Proposition 5, the fact that all noncontextual models $\mathbf{w}$ satisfy $\mathbf{k} \cdot \boldsymbol{\eta}^{\mathbf{w}} \leqslant M$ implies that (9) is a logical Bell inequality. Since $\mathbf{k} \cdot \boldsymbol{\eta}^{\mathbf{v}} > M$, $\mathbf{v}$ violates this inequality.

Note that, for any vector $\boldsymbol{\eta} \in \mathcal{E}$, $\boldsymbol{\eta} = \boldsymbol{\eta}^{\mathbf{w}}$ for some noncontextual model $\mathbf{w}$, and $\mathbf{w}$ satisfies all the logical Bell inequalities.

It is also possible to reverse the procedure described in Theorem 4, to obtain a complete set of inequalities directly applicable to expectation vectors.

Given a logical Bell inequality of the form,

$$\sum_{U \in \mathcal{U}} 2k_U\, p(\theta_U) \leqslant K,$$

where for each U, θ_U is either ψ_U or $\neg\psi_U$, we can form the inequality,

$$\sum_{U \in \mathcal{U}} l_U E_U \leqslant M,$$

where $M = K - \sum_{U \in \mathcal{U}} k_U$, and

$$l_U = \begin{cases} k_U, & \theta_U = \psi_U \\ -k_U, & \theta_U = \neg\psi_U. \end{cases}$$

We call this class of inequalities on expectation vectors the *logical correlation inequalities*.

As an immediate consequence of Theorem 4, we have the following.

Theorem 5. An expectation vector $\boldsymbol{\eta}$ is in $\mathcal{E}$ if and only if it satisfies all the logical correlation inequalities.

A. Example

We consider the following correlation inequality for the (3,2,2) case given by Werner and Wolf in Ref. [35]:

$$\frac{1}{4}\sum_{i=1}^{8} E_i \;-\; E_8 \;\leqslant\; 1. \quad (A2)$$

Here E_i, for $i = 1, \ldots, 8$, is the expectation value for the combination of measurements whose value, written as a binary string, is $i - 1$.

If we write this more explicitly, and clear the denominator of the scaling factor $1/4$, we obtain

$$\sum_{i=1}^{7} E_i \;-\; 3E_8 \;\leqslant\; 4.$$

If we now convert this to the form (4), following the procedure given in the proof of Theorem 5, we obtain the following inequality:

$$\sum_{i=1}^{7} p(\psi_i) \;+\; 3p(\neg\psi_8) \leqslant 7.$$

We can see that the multiset,

$$\sum_{i=1}^{7} 1\psi_i \;+\; 3(\neg\psi_8),$$

is 7-consistent. In fact, $\neg\psi_8$, together with any five of the formulas $\psi_1, \ldots, \psi_7$, is inconsistent.

B. Example

There are a number of extremal vertices of the no-signaling polytope in the (3,2,2) case, as listed in Ref. [36], which satisfy all the correlation inequalities from [35,37].

We shall examine one of these in detail. This is the vertex numbered 4 in the listing in Ref. [36].

We shall label the measurements as a, a' for site 1; b, b' for site 2; and c, c' for site 3. The support of the model for each measurement combination m can be represented by a formula φ_m; since the distribution on each row is uniform on the support, this completely specifies the model.

We recall the definition of ψ_U from (7). The formulas for the support of the model are defined as follows:

$$\varphi_{abc} = \varphi_{abc'} = \psi_{ab}; \quad \varphi_{ab'c'} = \varphi_{a'b'c'} = \psi_{b'c'},$$

$$\varphi_{a'bc} = \varphi_{a'b'c} = \psi_{a'c}; \quad \varphi_{ab'c} = \psi_{ab'c}; \quad \varphi_{a'bc'} = \neg\psi_{a'bc'}.$$

Combining these, we obtain the following multiset of formulas:

$$2\psi_{ab} + 2\psi_{b'c'} + 2\psi_{a'c} + \psi_{ab'c} + \neg\psi_{a'bc'}.$$

Since ψ_{ab} is equivalent to $a \leftrightarrow b$, in the presence of the first three formulas $\psi_{ab'c}$ is equivalent to $\psi_{bc'c}$, and $\neg\psi_{a'bc'}$ is equivalent to $\neg\psi_{cbc'}$. Since ψ_U is independent of the order in which the elements of U are listed, we obtain a contradiction. In fact, this multiset of formulas is 7-consistent, so the model achieves a maximal violation of the logical Bell inequality,

$$2p(\psi_{ab}) + 2p(\psi_{b'c'}) + 2p(\psi_{a'c}) + p(\psi_{ab'c}) + p(\neg\psi_{a'bc'}) \leqslant 7.$$

This yields a concrete example of a no-signaling model which satisfies all the correlation inequalities, while maximally violating the canonical logical Bell inequality arising from its support.

VIII. MULTIPLE OUTCOMES

Thus far we have focused exclusively on dichotomic measurements, which are particularly convenient for connecting to logic. However, the general format of measurement covers easily allows the results to be extended to measurements with multiple outcomes.

For example, we consider the case of $(n,k,2^p)$ Bell scenarios: n sites, k measurements per site, and 2^p outcomes per measurement. This corresponds to the following situation in our setting. We have a set X with nkp elements $\{m^i_{j,l}\}$, where $i = 1,\ldots,n$, $j = 1,\ldots,k$, and $l = 1,\ldots,p$. We write

$$X^i_j := \{m^i_{j,l} | l = 1,\ldots,p\}, \quad X_i := \bigcup_{j=1}^{k} X^i_j.$$

The cover $\mathcal{U}$ comprises all those subsets U of X such that, for all $i = 1,\ldots,n$, for some j, $U \cap X_i = X^i_j$. The idea is that X_i is the set of measurements which can be performed at site i. There are k choices available at each site between sets X^i_j of p dichotomic measurements each. Because these measurements are compatible, they can be performed together, resulting in a measurement with 2^p possible outcomes. An overall choice of measurements consists of selecting one such compatible family for each site.

All our results apply directly to this situation, which is itself a very special case of the general notion of cover. Thus from Theorem 5, we have an explicit description of a complete set of correlation inequalities characterizing the $(n,k,2^p)$ Bell scenarios.

IX. FINAL REMARKS

For further directions, it would be of particular interest to see how much of the present approach could be lifted to the quantum set, and the Tsirelson inequality [38].

As regards related work, the form of expressions we have used for the logical inequalities correspond to basic weight formulas in the logic for reasoning about probabilities studied in Ref. [39], following [40], which was motivated by applications in artificial intelligence.

The correlation polytopes of Pitowsky [41], which have a lineage going back to Boole's "conditions of possible experience" [42], should also be mentioned. Although this line of thought is certainly in a kindred spirit, Boole's conditions are arithmetical in nature; while the central theme of the present paper is that complete sets of Bell inequalities can be defined in terms of purely logical consistency conditions.

The notion of K-consistency is closely related to the well-known maximum satisfiability problem in computational complexity [43]. This asks for the maximum number of clauses in a given set which are satisfiable.

ACKNOWLEDGMENTS

S.A. thanks Harvey Brown, Matty Hoban, Shane Mansfield, and Rui Soares Barbosa for a number of valuable discussions. We gratefully acknowledge financial support from the Perimeter Institute, Engineering and Physical Sciences Research Council Senior Research Fellowship No. EP/E052819/1, and US Office of Naval Research Grant No. N000141010357.

[1] J. Bell, Physics **1**, 195 (1964).
[2] J. Clauser, M. Horne, A. Shimony, and R. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[3] S. Abramsky and A. Brandenburger, New J. Phys. **13**, 113036 (2011).
[4] A. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[5] J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. **95**, 10503 (2005).
[6] A. Acin, N. Gisin, and L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006).
[7] Č. Brukner, M. Żukowski, J. Pan, and A. Zeilinger, Phys. Rev. Lett. **92**, 127901 (2004).
[8] B. Terhal, Theor. Comput. Sci. **287**, 313 (2002).
[9] H. Bartosik, J. Klepp, C. Schmitzer, S. Sponar, A. Cabello, H. Rauch, and Y. Hasegawa, Phys. Rev. Lett. **103**, 040403 (2009).

[10] G. Kirchmair, F. Zähringer, R. Gerritsma, M. Kleinmann, O. Gühne, A. Cabello, R. Blatt, and C. Roos, Nature (London) **460**, 494 (2009).

[11] D. Greenberger, M. Horne, A. Shimony, and A. Zeilinger, Am. J. Phys. **58**, 1131 (1990).

[12] L. Hardy, Phys. Rev. Lett. **68**, 2981 (1992).

[13] J. Zimba and R. Penrose, Studies in History and Philosophy of Science Part A **24**, 697 (1993).

[14] L. Hardy, Phys. Rev. Lett. **71**, 1665 (1993).

[15] L. Hardy, Phys. Lett. A **161**, 21 (1991).

[16] S. Braunstein and C. Caves, Ann. Phys. **202**, 22 (1990).

[17] Since each p_i is a probability, its maximum value is 1, so the "algebraic maximum" of the sum $\sum_i p_i$ is N.

[18] S. Abramsky, S. Mansfield, and R. Barbosa, arXiv:1111.3620v1.

[19] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).

[20] We shall eventually consider a much more general form of such models introduced in Ref. [3], which allow a uniform treatment of contextuality, including Kochen-Specker configurations, etc.

[21] The support of a probability model can be given the structure of a presheaf, in such a way that the above definition corresponds to the usual notion of a global section. See Ref. [3] for details.

[22] D. Greenberger, M. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer, Norwell, 1989), pp. 69–72.

[23] N. Mermin, Am. J. Phys. **58**, 731 (1990).

[24] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).

[25] For further details on this perspective on the Kochen-Specker theorem, see Ref. [3].

[26] A. Cabello, J. Estebaranz, and G. García-Alcaine, Phys. Lett. A **212**, 183 (1996).

[27] A. Peres, Phys. Lett. A **151**, 107 (1990).

[28] N. D. Mermin, Phys. Rev. Lett. **65**, 3373 (1990).

[29] In this note, following [3], we only consider finite sets of measurements, and thus discrete probability distributions suffice.

[30] J. Matoušek and B. Gärtner, *Understanding and Using Linear Programming* (Springer Verlag, Berlin, 2007).

[31] A. Schrijver, *Theory of Linear and Integer Programming* (John Wiley & Sons, New York, 1998).

[32] G. Ziegler, *Lectures on Polytopes*, Vol. 152 (Springer, New York, 1995).

[33] O. Strichman, in *Formal Methods in Computer-Aided Design* (Springer, New York, 2002), pp. 160–170.

[34] T. Christof, A. Löbel, and M. Stoer, ftp://ftp.zib.de/pub/ Packages/mathprog/polyth/porta (1997).

[35] R. F. Werner and M. M. Wolf, Phys. Rev. A **64**, 032112 (2001).

[36] S. Pironio, J. Bancal, and V. Scarani, J. Phys. A: Math. Theor. **44**, 065303 (2011).

[37] The first author is grateful to Matty Hoban for bringing these examples to his attention.

[38] B. Tsirelson, Lett. Math. Phys. **4**, 93 (1980).

[39] R. Fagin, J. Halpern, and N. Megiddo, Inf. Comput. **87**, 78 (1990).

[40] N. Nilsson, Artif. Intell. **28**, 71 (1986).

[41] I. Pitowsky, The British Journal for the Philosophy of Science **45**, 95 (1994).

[42] G. Boole, Philos. Trans. R. Soc. London **152**, 225 (1862).

[43] C. Papadimitriou, *Computational Complexity* (John Wiley & Sons, New York, 2003).