Deriving robust noncontextuality inequalities from algebraic proofs of the Kochen–Specker theorem: the Peres–Mermin square

Anirudh Krishna1,2,3, Robert W Spekkens1 and Elie Wolfe1
1 Perimeter Institute for Theoretical Physics, Waterloo, Ontario, N2L 2Y5, Canada
2 Université de Sherbrooke, Sherbrooke, Québec, J1K 2R1, Canada
3 Author to whom any correspondence should be addressed. E-mail: anirudh.krishna@usherbrooke.ca, rspekkens@perimeterinstitute.ca and ewolfe@perimeterinstitute.ca

Keywords: noncontextuality, quantum foundations, Peres–Mermin square

Abstract

When a measurement is compatible with each of two other measurements that are incompatible with one another, these define distinct contexts for the given measurement. The Kochen–Specker theorem rules out models of quantum theory that satisfy a particular assumption of context-independence: that sharp measurements are assigned outcomes both deterministically and independently of their context. This notion of noncontextuality is not suited to a direct experimental test because realistic measurements always have some degree of unsharpness due to noise. However, a generalized notion of noncontextuality has been proposed that is applicable to any experimental procedure, including unsharp measurements, but also preparations as well, and for which a quantum no-go result still holds. According to this notion, the model need only specify a probability distribution over the outcomes of a measurement in a context-independent way, rather than specifying a particular outcome. It also implies novel constraints of context-independence for the representation of preparations. In this article, we describe a general technique for translating proofs of the Kochen–Specker theorem into inequality constraints on realistic experimental statistics, the violation of which witnesses the impossibility of a noncontextual model. We focus on algebraic state-independent proofs, using the Peres–Mermin square as our illustrative example. Our technique yields the necessary and sufficient conditions for a particular set of correlations (between the preparations and the measurements) to admit a noncontextual model. The inequalities thus derived are demonstrably robust to noise. We specify how experimental data must be processed in order to achieve a test of these inequalities. We also provide a criticism of prior proposals for experimental tests of noncontextuality based on the Peres–Mermin square.

1. Introduction

Ontological models of quantum theory are an attempt to explain the statistical predictions of quantum theory. They take every system to be associated with a space of possible physical states, termed ontic states, every quantum state to be represented by a statistical distribution over these ontic states, and every measurement to be represented by a conditional probability distribution for the outcome given the ontic state [1]. Hidden variable models are examples of ontological models, but so is the physicist’s orthodox conception of quantum theory, wherein the ontic states are simply the pure quantum states, not supplemented by any additional variables.

The principle of noncontextuality is an assumption about ontological models that seeks to capture a notion of classicality. It started its life as an assumption about outcome-deterministic ontological models of quantum theory, that is, ontological models wherein the outcome of every measurement was fixed deterministically by the ontic state (in contrast to the orthodox conception). This assumption was famously demonstrated to be in contradiction with the predictions of quantum theory by Kochen and Specker [2] and Bell [3].

The latter is termed the ψ-complete ontological model in [1].
Kochen–Specker theorem is one of the strongest constraints on the interpretation of quantum theory. Furthermore, failing to admit of a noncontextual model appears to be a resource. For instance, in the context of the state injection model for quantum computation [4, 5], the failure of noncontextuality has been shown in some cases to be necessary for achieving universal quantum computation [6, 7].

In [8], a generalized notion of noncontextuality was proposed. For measurements, it constitutes a relaxation of what was assumed by Kochen–Specker and by Bell. Specifically, it allows the assignment of measurement outcomes by ontic states to be indeterministic. In this way, it redefines the notion of noncontextuality for measurements in a way that excised the notion of determinism. This is desirable from a foundational perspective as it allows one to separate the issue of noncontextuality from that of determinism (recall that Bell’s notion of local causality does not presume that the outcomes of measurements are fixed deterministically). Additionally, it can be shown that the assumption of outcome determinism is unwarranted for any unsharp measurement (i.e., a measurement for which one cannot find a basis of preparations relative to which it is perfectly predictable), and every measurement appearing in a real experiment is of this sort [9]. As such, this generalization is important if one hopes to turn the proven theoretical advantages for computation into practical advantages, because in practice, sharpness is an idealization that is never strictly satisfied.

Although the revised notion of noncontextuality yields a weaker constraint on the representation of measurements in the ontological model than did the traditional notion, it naturally applies not only to measurements but to preparations as well, and thereby implies novel constraints on how quantum states can be represented by distributions over ontic states in the model. It was argued in [8] that whatever motivations can be given for assuming noncontextuality for one type of procedure, such as a measurement, this same motivation can be given for assuming it of any other type of procedure, such as a preparation. Consequently, the only natural assumption to consider in this approach is that the revised notion of noncontextuality applies to all procedures. This assumption is termed universal noncontextuality or simply noncontextuality. We will henceforth refer to the traditional notion of noncontextuality as KS-noncontextuality (for Kochen–Specker) to avoid any confusion.

In [8], it was shown that quantum theory does not admit of a universally noncontextual ontological model. It was also demonstrated that if one replaces the assumption of KS-noncontextuality for measurements with the assumption of universal noncontextuality for all procedures, the relaxation of the constraints on the representation of measurements is compensated by the strengthening of the constraints on the representation of preparations in such a way that any proof that quantum theory fails to admit of a KS-noncontextual model can be translated into a proof that it fails to admit of a universally noncontextual model.

Much of the research on noncontextuality to date has centred on the question of whether quantum theory admits of a noncontextual model. A more general question, which has been the impetus for much recent work, is whether one can devise a direct experimental test of the assumption of a noncontextual ontological model, one that is independent of the validity of quantum theory. Just as a Bell inequality is a constraint on experimental statistics that follows directly from the assumption of a locally causal ontological model, without any reference to the quantum formalism, what one wants of a test of noncontextuality is a constraint on experimental statistics that follows directly from the assumption of a noncontextual ontological model, without any reference to the quantum formalism. Such constraints will here be termed noncontextuality inequalities. If experimental statistics are found to violate these inequalities, then one can conclude that not just quantum theory but any operational theory that can do justice to the experimental statistics—and therefore nature itself—must fail to admit of such a model, thereby constraining the form of all future physical theories.

The generalized notion of noncontextuality proposed in [8] was defined in such a way as to be applicable to any operational theory, not just quantum theory, such that if an experiment yields data supporting an operational theory distinct from quantum theory, the question of whether it admits of a noncontextual model is still meaningful. The definition asserts that an ontological model of an operational theory is noncontextual if two experimental procedures that are statistically indistinguishable at the operational level are statistically indistinguishable at the ontological level. They key point is that the notion of statistical indistinguishability at the operational level can be assessed in any operational theory.

It has been shown that violations of noncontextuality inequalities defined in terms of this notion can imply advantages for information processing which are independent of the validity of quantum theory. For example, they imply an advantage for the cryptographic task of parity-oblivious random access codes [10–12]. Such

---

5 A consequence of this relaxation is that, by its lights, the ψ-complete ontological model is found to represent measurements noncontextually. However, it fails to represent preparations noncontextually, as noted in [8].

6 It also implies novel constraints on how quantum channels can be represented by conditional probability distributions from the space of ontic states to itself.

7 The fact that the notion of statistical indistinguishability is applicable not just pairs of measurements, but to pairs of preparations and transformations respectively as well, is what allows the generalized notion of noncontextuality to be applicable to these other procedures.
inequalities also hold promise for making the results on quantum computational advantages discussed above robust to noise and for expressing the origin of the advantage in a manner that is independent of the validity of quantum theory.

Several recent works have considered the question of how to derive noncontextuality inequalities and how to subject them to experiment test [13–15]. The present work is concerned with a special case of this problem, namely, how to derive noncontextuality inequalities starting from any given proof of the Kochen–Specker theorem, that is, from a proof of the failure of KS-noncontextuality in quantum theory. As noted above, [8] showed how, in general, to convert a proof of the failure of KS-noncontextuality in quantum theory into a proof of the failure of universal noncontextuality in quantum theory, so the outstanding problem is how to convert a proof of the failure of universal noncontextuality in quantum theory into an operational noncontextuality inequality.

Note that any test of noncontextuality that is devised from a particular no-go theorem requires an experimentalist to target a particular set of preparations and a particular set of measurements, each with specified relations holding among their members (we will say more about the nature of these relations in due course). A more general version of the problem, however, is to figure out how to infer from any experimental data—that is, from an experiment that was not designed to target particular preparations or measurements or any particular relations among them—whether or not it admits of a noncontextual model. Because a test of noncontextuality is a test of classicality, having the capability to test the assumption of noncontextuality on any experimental data is clearly of greater utility than merely knowing how to implement a dedicated experiment for testing the hypothesis of noncontextuality. Pusey [15] identified the conditions that are both necessary and sufficient for the existence of a noncontextual model for experimental data derived from the simplest experimental scenario in which such conditions are expected to be nontrivial. Unfortunately, this simplest scenario does not arise within operational quantum theory7. Extending Pusey’s analysis to more general scenarios is an important open problem.

Nonetheless, there are also advantages to building sets of noncontextuality inequalities from specific proofs of the Kochen–Specker theorem, because such proofs have nontrivial structural properties. Different proofs—and there is now a great diversity of these—capture what is surprising about the failure of noncontextuality in different ways, and these intuitions are likely to be helpful in identifying the applications thereof.

We here focus on deriving noncontextuality inequalities from state-independent proofs of the Kochen–Specker theorem.

Reference [14] has already demonstrated how one can derive one such inequality from any state-independent geometric proof of the Kochen–Specker theorem, that is, any proof expressed in terms of an uncolourable set of rays. Here, we extend this work in two important ways: (1) we provide a technique for finding all of the noncontextuality inequalities that apply to a certain set of correlations starting from any state-independent proof of the Kochen–Specker theorem, and (2) we show how to do so for proofs that are expressed algebraically rather than geometrically. We expand on each of these points presently, in reverse order.

The distinction between geometric and algebraic proofs of the failure of KS-noncontextuality in quantum theory is not fundamental because one can convert any algebraic proof into a geometric form and vice-versa. Nonetheless, each proof style has its advantages. The first known proofs were geometric uncolorability proofs. Algebraic proofs arose later, but in many respects they have a logic that is easier to grasp. Indeed, the paradigm example of a proof of the Kochen–Specker theorem is now arguably the algebraic version of the Peres–Mermin square proof [16, 17], which will be the example we focus on here.

Furthermore, the algebraic structure suggests generalizations of these proofs that might not be obvious from the geometric perspective [18, 19]. Although one could derive a noncontextuality inequality for the Peres–Mermin square by first expressing the latter as a geometric proof (as in [16]) and then applying the technique described in [14], it is more useful to have a technique for deriving noncontextuality inequalities that is native to the algebraic approach. We here provide such a technique.

In order to turn a proof of the failure of universal noncontextuality in quantum theory into a noncontextuality inequality, one must operationalize the description of the experiment provided in the no-go theorem, purging it of any reference of the quantum formalism, and one must robustify the constraints on experimental data that are derived from noncontextuality, which means that these constraints must provide quantitative bounds that can be violated in principle even if the experimental operations are noisy. This progression was achieved in [14], but the resulting inequality provided an upper bound on just a single operational quantity (an average, over certain preparation–measurement pairs, of the degree of correlation between them). The technique described in the present article goes much further towards providing a means of

8 Recall that a set of measurements is said to be tomographically complete for a system if the statistics for any measurement on the system can be computed from the statistics of the measurements in this set. Pusey’s simplest scenario is one wherein a tomographically complete set of measurements consists of just two binary-outcome measurements. This scenario does not arise in operational quantum theory, because the simplest quantum system, a qubit, requires three binary-outcome measurements for tomographic completeness.
deriving all of the noncontextuality inequalities that hold for a given set of preparations and measurements. Although we focus on a subset of the correlations between preparations and measurements that arise in the construction, for this restricted set of experimental data, satisfaction of the inequalities that we derive is both necessary and sufficient for the existence of a noncontextual model.

Finally, we note a difference in the way experiments are described in this article relative to previous treatments of inequalities for universal noncontextuality [10, 13, 14]. We here use the notion of a source, that is, a process which samples a classical variable from a distribution, chooses which preparation procedure to implement on the system based on the value sampled and outputs both the system and the variable. This choice ensures that our derived noncontextuality inequalities are easier to compare with Bell inequalities.

The remainder of the paper is structured as follows.

In section 2, we provide an overview of operational theories (2.1) and ontological models (2.2). In particular, we discuss the concepts of operational equivalence and of compatibility (applied to measurements and sources) and illustrate the concepts with quantum examples. We provide formal definitions of measurement noncontextuality and preparation noncontextuality, in particular, a characterization of these assumptions in terms of expectation values for the outcomes of measurements and sources given the ontic state.

In section 3, we review the well-known proof of the failure of KS-noncontextuality in quantum theory based on the Peres–Mermin square (3.1), and we show how to translate this no-go theorem into one that demonstrates the failure of universal noncontextuality in quantum theory (3.2).

Section 4 is the heart of the article, describing our technique for turning quantum no-go theorems into operational noncontextuality inequalities. In the first section (4.1), we operationalize the description of the quantum measurements and sources that appear in the Peres–Mermin-inspired proof of the failure of universal noncontextuality, thereby obtaining a notion of a Peres–Mermin experimental scenario that is purged of any reference to quantum theory. This provides a template for how to achieve this operationalization for any such construction. The following five sections (4.2–4.6) describe how to derive noncontextuality inequalities from such an operational construction, using Peres–Mermin as the illustrative example. We also show how the ideal quantum realization of the measurements and sources in the Peres–Mermin scenario violate these inequalities (4.6.1), and we demonstrate the robustness of these inequalities to noise (4.6.2), by showing how they can be violated by partially depolarized versions of the ideal quantum realizations of the measurements and sources.

In section 5, we clarify what must be done experimentally in order to test the noncontextuality inequalities we have derived, and in section 6 we provide our concluding remarks.

Appendix A discusses the problem of computationally converting between the vertex and halfspace representations of a polytope, appendix B discusses the symmetries of our noncontextuality inequalities under deterministic processings of the experimental procedures, and appendix C demonstrates that a certain class of inequalities on experimental statistics are trivial. Finally, appendix D reviews a previous proposal for how to implement an experimental test of noncontextuality based on the Peres–Mermin square, and argues against its adequacy.

2. Preliminaries

2.1. Operational concepts

2.1.1. Operational theories

The primitive elements of an operational theory are preparations and measurements, each specified as lists of instructions to be performed in the laboratory.

A source is a device that implements one of a set of preparation procedures on a system, sampled from some probability distribution, and has a classical outcome that heralds which preparation has in fact been implemented. (The use of the term ‘source’ to refer to such a device is conventional in both classical and quantum Shannon theory, where it is the standard way of modelling the input to a communication channel [20].) We will denote a source by \( S \) and the variable describing its classical outcome by \( s \).

A measurement, denoted \( M \), accepts as input a system and returns a classical outcome, denoted by the variable \( m \).

An operational theory provides an algorithm for computing the probability distribution for the outcome of any measurement acting on any preparation, and consequently it allows the computation of the joint probability distribution over the outcome of any measurement \( M \) and the outcome of any source \( S \), \( \text{pr}\{ m, s | M, S \} \). We refer to this as simply the joint distribution on the measurement-source pair \( (M, S) \).

2.1.2. Operational equivalence

Consider two measurement procedures, \( M_i \) and \( M_j \), whose outcomes are random variables, denoted \( m_i \) and \( m_j \) respectively. \( M_i \) and \( M_j \) are said to be operationally equivalent if they define the same joint distribution for all
possible sources:

$$\forall S : \text{pr}\{m_1, s|M_1, S\} = \text{pr}\{m_2, s|M_2, S\}. \quad (1)$$

Letting $\mathcal{M}$ denote an operational equivalence class of measurements, we can express the operational equivalence of two measurements $M_1$ and $M_2$ by specifying that they belong to the same class,

$$M_1, M_2 \in \mathcal{M}. \quad (2)$$

Similarly, two sources, $S_1$ and $S_2$, whose outcomes are random variables $s_1$ and $s_2$ respectively, are said to be operationally equivalent if they define the same joint distribution for all possible measurements:

$$\forall M : \text{pr}\{m, s_1|M_1, S_1\} = \text{pr}\{m, s_2|M_2, S_2\}. \quad (3)$$

Letting $\mathcal{S}$ denote an operational equivalence class of sources, we can express the operational equivalence of two sources, $S_1$ and $S_2$, by stipulating that they are in the same class,

$$S_1, S_2 \in \mathcal{S}. \quad (4)$$

Because the joint distribution $\text{pr}\{s, m|M_1, S\}$ is computable from the operational theory, so too is the correlation between the outcome $s$ of the source $S$ and the outcome $m$ of the measurement $M$,

$$\langle s m \rangle_{M, S} = \sum_{s,m} s \text{pr}\{s, m|M, S\}. \quad (5)$$

Such correlations will be the quantities appearing in our noncontextuality inequalities.

### 2.1.3. Compatibility

In this section, we briefly review the notion of compatibility. The interested reader is pointed to [21] for a detailed overview.

Informally, two or more devices are said to be compatible if their output can be obtained by classical post-processing of the output of a single device. More precisely, one can define the notion of compatibility in terms of a notion of simulatability.

Consider two measurements, $M$ and $M'$, which accept an input system and output random variables $m$ and $m'$ respectively. We say that $M$ can simulate $M'$ if there exists a conditional distribution $\text{pr}\{m'|m\}$ such that

$$\forall S : \text{pr}\{m', s|M', S\} = \sum_m \text{pr}\{m'|m\} \text{pr}\{m, s|M, S\}. \quad (6)$$

Two measurements $M_1$ and $M_2$ are said to be compatible if both of them can be simulated by some third measurement $M$, that is, if there exists $\text{pr}\{m_1|m\}$ and $\text{pr}\{m_2|m\}$ such that

$$\forall S : \text{pr}\{m_1, s|M_1, S\} = \sum_m \text{pr}\{m_1|m\} \text{pr}\{m, s|M, S\},$$

$$\forall S : \text{pr}\{m_2, s|M_2, S\} = \sum_m \text{pr}\{m_2|m\} \text{pr}\{m, s|M, S\}. \quad (7)$$

Similar definitions hold for sources. We say that a source $S$ with classical outcome $s$ simulates source $S'$ with classical outcome $s'$ if there exists a conditional distribution $\text{pr}\{s'|s\}$ such that

$$\forall M : \text{pr}\{m, s'|M', S'\} = \sum_s \text{pr}\{s'|s\} \text{pr}\{m, s|M, S\}. \quad (8)$$

When it comes to defining a notion of compatibility of sources, there is a nuance relative to the case of measurements. The definition we adopt will apply only to those pairs of sources, $S_1$ and $S_2$, that are operationally equivalent when one marginalizes over their outcomes, that is, it presumes that $S_1$ and $S_2$ are such that

$$\forall M : \sum_s \text{pr}\{m, s|s_1, S_1\} = \sum_{s_2} \text{pr}\{m, s_2|M_1, S_2\}. \quad (9)$$

Every set of sources we consider in this article will have this property. Two such sources, $S_1$ and $S_2$, are said to be compatible if there exists a third source $S$ that simulates them both, that is, if there exists $\text{pr}\{s_i|s\}$ and $\text{pr}\{s_2|s\}$ such that

$$\forall M : \text{pr}\{m, s_1|M, S_1\} = \sum_s \text{pr}\{s_i|s\} \text{pr}\{m, s|M, S\}, \forall M : \text{pr}\{m, s_2|M, S_2\} = \sum_s \text{pr}\{s_2|s\} \text{pr}\{m, s|M, S\}. \quad (10)$$

It is instructive to consider what these notions of compatibility correspond to in quantum theory. Recall that every measurement in quantum theory is represented by a positive operator-valued measure (POVM) whose elements are labelled by the outcome $m$, that is, by $\{E_m\}_m$ where $\forall m : E_m \geq 0$ and $\sum_mE_m = 1$. 

---

New J. Phys. 19 (2017) 123031

A Krishna et al
Every source in quantum theory is represented by an ensemble of subnormalized quantum states \( \{p_s \rho_s\} \), where \( \forall s : \rho_s \geq 0 \), \( \text{Tr}(\rho) = 1 \) and \( p_s \) is a probability distribution over \( s \). In other words, the ensemble \( \{p_s \rho_s\} \), defines the source that samples \( s \) from the probability distribution \( p_s \), then prepares the quantum system in the normalized state \( \rho_s \) and outputs \( s \) as its outcome.

If a measurement \( M \) is represented by the POVM \( \{E_m\}_m \) and a source \( S \) is represented by the ensemble \( \{p_s \rho_s\}_s \), then the probability of the source generating outcome \( s \) and the measurement generating outcome \( m \) is given by the Born rule as \( \text{pr}(m,s|M,S) = \text{Tr}(E_m p_s \rho_s) \).

Substituting this Born rule expression into equation (7), we infer that if two measurements, associated to POVMs \( \{E_m\}_m \) and \( \{E_m\}_m \), are compatible, then there exists a third POVM \( \{E_m\}_m \) and conditional distributions \( \text{pr}(m|M) \) and \( \text{pr}(m|m) \) such that
\[
E_m = \sum_m \text{pr}(m|M) E_m,
\]
\[
E_m = \sum_m \text{pr}(m|m) E_m.
\]

For two quantum sources, associated to ensembles \( \{p_s \rho_s\}_s \) and \( \{p_s \rho_s\}_s \), the property that ensures the applicability of our definition of compatibility, equation (9), is that the two ensembles define the same average state \( \sum_s p_s \rho_s = \sum_s p_s \rho_s \). Substituting the Born rule expression into equation (8), we infer that if the two quantum sources are compatible, then there exists a third ensemble \( \{p_s \rho_s\} \), and conditional distributions \( \text{pr}(s|M) \) and \( \text{pr}(s|m) \) such that
\[
p_s \rho_s = \sum_m \text{pr}(s|M) p_s \rho_s \quad p_s \rho_s = \sum_m \text{pr}(s|m) p_s \rho_s.\]  

Note that the notion of compatibility for quantum measurements that we have articulated above concerns only their retrodictive aspect and makes no reference to how the quantum state of a system evolves as a result of the measurement. In other words, it is sufficient to know which POVM is associated to the measurement, while the instrument that is associated to it, i.e., the set of update maps for each outcome, is irrelevant (indeed, it is not even required that there be an update map—the quantum system could be destroyed in the measurement process). This contrasts with the notion of compatibility that is the focus of many other works seeking to devise experimental tests of noncontextuality, where two measurements on a system are deemed compatible if implementing them in one temporal order gives the same statistics as implementing them in the opposite temporal order. The joint-simulatability notion of compatibility articulated above pertains not just to sharp measurements (represented by projector-valued measures) but to all unsharp measurements as well (represented by POVMs). In particular, it allows nontrivial compatibility relations among unsharp measurements that are associated to POVMs wherein the different elements of the POVM do not commute with one another, whereas such POVMs need not even compatible with themselves according to the temporal-reordering notion of compatibility. The wide scope of applicability of the joint-simulatability notion of compatibility makes it particularly well-equipped to contend with experimental noise and imperfections in tests of noncontextuality.

A given POVM defines an operational equivalence class of measurements insofar as it can be implemented in many different ways. This is because any given POVM is generally compatible with many other POVMs which are not compatible with one another, and for each such compatible set of POVMs, there is a different experimental procedure, and hence a different concrete realization of the given POVM. The compatible set of which the POVM is considered a part is therefore an example of a measurement context. Note that if one particularizes the definition of compatibility to projector-valued measures, then the condition for compatibility becomes commutativity of the associated observables, and we recover the standard notion of a measurement context of an observable as the commuting set of observables of which it is considered a part. The important point, however, is that in addition to recovering the standard notion of measurement context for sharp (i.e., projective) measurements, one has a notion of measurement context also for unsharp measurements.

In a similar fashion, a given ensemble of quantum states defines an operational equivalence class of sources because it too can be implemented in many different ways, depending on which compatible set of ensembles it is considered to be a member of. The compatible set of ensembles of which it is a member constitutes the source context. It will later be useful to distinguish between sharp and unsharp quantum sources, where sharp sources are those consisting entirely of states that are normalizable projectors. Clearly, the notion of compatibility of

\footnotetext{The specialization of our notion of compatibility to the quantum case allows us to clarify our motivation for restricting the scope of the notion to pairs of sources satisfying equation (9): if we did not restrict the notion in this manner, then two sources could be incompatible simply by virtue of averaging to different states, while the component states in the two sources were all diagonal in the same basis. For the purpose of evaluating the possibility of a noncontextual model, one prefers to have a notion of compatibility wherein sources being incompatible guarantees that they are not jointly diagonalizable.}
sources that we have introduced applies equally well to sharp and unsharp quantum sources, just as the notion for measurements applies to the sharp and unsharp cases alike.

2.2. Ontological concepts

2.2.1. Ontological models

As proposed in [8], generalized noncontextuality is a constraint on an ontological model of an operational theory. An ontological model is an attempt to reproduce the predictions of the operational theory by imagining that the correlations between the outcome of the source and that of the measurement are explained by the physical system that acts as a causal mediary between them. All of the physical attributes of the system at any given point in time is termed the \textit{ontic state} of the system at that time. We shall denote this by \( \lambda \), and the space of all possible ontic states of the system will be denoted by \( \Lambda \).

Consider the most general way of representing a measurement procedure \( M \) in an ontological model. The output \( m \) might not be completely determined by the ontic state \( \lambda \) of the system. Instead, specifying \( M \) might only specify the conditional probability of obtaining output \( m \) if the system is in the ontic state \( \lambda \). This could arise because of objective indeterminism or because the outcome of the measurement depends not only on the input system but also on degrees of freedom of the measurement apparatus. We denote this conditional probability by \( \xi(m|\lambda, M) \) and refer to it as the \textit{response function} associated to \( M \).

Similarly, the most general way of representing a preparation procedure in an ontological model is to allow that the preparation does not uniquely fix the ontic state of the system, but rather that the ontic state might only be sampled probabilistically from a distribution that is specified by the preparation. This implies that the most general way of representing a source \( S \) in an ontological model is as a joint distribution over its outcome \( s \) and the ontic state \( \lambda \) that it outputs, \( \mu(s, \lambda|S) \).

The purpose of an ontological model for an operational theory is to reproduce the statistics of that theory. This occurs if the ontological model is such that for all sources \( S \) and measurements \( M \),

\[
\text{pr}(m, s|M, S) = \sum_{\lambda \in \Lambda} \xi(m|\lambda, M) \mu(\lambda, s|S). \tag{12}
\]

It is useful to express the connection between the ontological model and the operational theory in terms of expectation values as well. The expectation value of outcome \( m \) of measurement \( M \) for the ontic state \( \lambda \) is

\[
\langle m \rangle_{\lambda, M} \equiv \sum_m m \xi(m|\lambda, M). \tag{13}
\]

The expectation value of outcome \( s \) of measurement \( S \) for the ontic state \( \lambda \) is a \textit{retrodictive} expectation value and so is a bit more subtle to express. Consider a source \( S \) with classical outcome \( s \) that is associated with the joint distribution \( \mu(s, \lambda|S) \). What probability ought one to assign to the outcome variable \( s \) having taken a particular value if one knows that the ontic state emitted by the source was \( \lambda \)? The answer is given by a simple Bayesian inversion:

\[
\mu(s|\lambda, S) = \frac{\mu(s, \lambda|S)}{\mu(\lambda|S)}, \tag{14}
\]

where

\[
\mu(\lambda|S) \equiv \sum_{s'} \mu(s', \lambda|S). \tag{15}
\]

We can then use this conditional probability to define an expectation value for an outcome \( s \) of a source \( S \) given knowledge of the ontic state \( \lambda \), as:

\[
\langle s \rangle_{\lambda, S} \equiv \sum_s s \mu(s|\lambda, S). \tag{16}
\]

It follows that the correlation between \( m \) and \( s \) can be expressed as

\[
\langle ms \rangle_{M, S} = \sum_{m,s} ms \text{pr}(m, s|M, S) = \sum_{\lambda \in \Lambda} \langle m \rangle_{\lambda, M} \langle s \rangle_{\lambda, S} \mu(\lambda|S), \tag{17}
\]

where we have used equations (12)–(14) and (16). We will use the latter expression when deriving our quantum no-go theorems and noncontextuality inequalities.

2.2.2. Measurement noncontextuality

The assumption of measurement noncontextuality stipulates that if two measurements \( M_1 \) and \( M_2 \) are operationally equivalent, then the response functions associated to these measurements in the ontological model
are equal. Equivalently, we can express measurement noncontextuality as the assumption that the response function associated to a measurement depends only on its operational equivalence class and not on the measurement context (which explains the appropriateness of the term ‘noncontextual’). Denoting the operational equivalence class of $M_1$ and $M_2$ by $\mathcal{M}$ (with outcome denoted by $m$), and denoting the response function for $M_1$ and $M_2$ and $\mathcal{M}$ by $\xi(m|\lambda, M_1)$, $\xi(m|\lambda, M_2)$ and $\xi(m|\lambda, \mathcal{M})$ respectively, the assumption of measurement noncontextuality can be formalized as follows:

$$\text{If } M_1, M_2 \in \mathcal{M}, \text{ then } \forall \lambda \in \Lambda : \xi(m|\lambda, M_1) = \xi(m|\lambda, M_2) \equiv \xi(m|\lambda, \mathcal{M}). \quad (18)$$

We can also express measurement noncontextuality in terms of expectation values, using equation (13), as

$$\text{If } M_1, M_2 \in \mathcal{M}, \text{ then } \forall \lambda \in \Lambda : \langle m|\lambda, M_1 \rangle = \langle m|\lambda, M_2 \rangle \equiv \langle m|\lambda, \mathcal{M} \rangle. \quad (19)$$

An ontic state $\lambda$ will be said to be make a deterministic assignment to an equivalence class of measurements $\mathcal{M}$ if

$$\xi(m|\lambda, \mathcal{M}) \in \{0, 1\}, \quad (20)$$

or equivalently, if $\langle m|\lambda, \mathcal{M} \rangle \in \text{Range}(m)$.

We pause to note the how the standard notion of measurement context that appears in proofs of the Kochen–Specker theorem is understood in our framework. Consider a measurement associated to an observable $O$. If $O$ is compatible with $O_1$ and $O$ is compatible with $O_2$, but $O_1$ and $O_2$ are not compatible with one another, then there are two operationally equivalent ways of implementing a measurement of $O$, namely, by measuring it jointly with $O_1$ and by measuring it jointly with $O_2$. If $\langle m|\lambda, O \rangle_2$ denotes the expectation value for the outcome of the $O$ measurement when it is measured jointly with $O_1$ and $\langle m|\lambda, O \rangle_2$ denotes the expectation value for the outcome of the $O$ measurement when it is measured jointly with $O_2$, then the assumption of measurement noncontextuality implies that $\forall \lambda \in \Lambda : \langle m|\lambda, O \rangle_{21} = \langle m|\lambda, O \rangle_{22} \equiv \langle m|\lambda, O \rangle$. The fact that the notion of measurement noncontextuality does not include the assumption of outcome determinism translates, in the case of quantum observables, to the fact that these expectation values are not assumed, a priori, to lie in the eigenspectrum of $O$.

The generalized notion of noncontextuality allows one to extend this analysis to the case of a quantum measurement that is associated to a POVM that is nonprojective. If it can be measured jointly with either one of two other POVMs which are not compatible with one another, then the assumption of measurement noncontextuality implies that the expectation value of its outcome given the ontic state $\lambda$ should be independent of which of the two other POVMs it is measured jointly with.

2.2.3. Preparation noncontextuality

The assumption of preparation noncontextuality has previously been expressed in terms of individual preparations. However, in this article we will be describing experiments in terms of sources, and so we will here express it in the language of sources.

The assumption of preparation noncontextuality stipulates that if two sources $S_1$ and $S_2$ are operationally equivalent, then the joint distributions over ontic states and outcomes that represent these sources in the ontological model are equal. Equivalently, the joint distribution over ontic states and outcomes representing a source depends only on the operational equivalence class of that source. Denoting the operational equivalence class of $S_1$ and $S_2$ by $\mathcal{S}$ (with outcome denoted by $s$), and denoting the joint distribution over ontic states and outcomes for $S_1$ and $S_2$ and $\mathcal{S}$ by $\mu(s_1, \lambda|S_1)$, $\mu(s_2, \lambda|S_2)$, and $\mu(s, \lambda|\mathcal{S})$ respectively, the assumption of preparation noncontextuality can be formalized as follows:

$$\text{If } S_1, S_2 \in \mathcal{S}, \text{ then } \forall \lambda \in \Lambda : \mu(s_1, \lambda|S_1) = \mu(s_2, \lambda|S_2) \equiv \mu(s, \lambda|\mathcal{S}). \quad (21)$$

Note that an individual preparation can be understood as a special kind of source, one wherein the outcome is trivial (i.e. taking a value in a singleton set), and for such sources the definition of noncontextuality provided above reduces to the standard one for preparations articulated in [8].

We can also express this assumption in terms of expectation values, as we did for the case of measurements, but with a critical difference, as noted above equation (16): for sources, the relevant expectation values concern retrodictions rather than predictions.

If two sources, $S_1$ and $S_2$, are operationally equivalent then not only does preparation noncontextuality imply that the distributions $\mu(s_1, \lambda|S_1)$ and $\mu(s_2, \lambda|S_2)$ are equal (and hence the marginals $\mu(\lambda|S_1)$ and $\mu(\lambda|S_2)$ are equal as well), it also implies, via equation (14), that the conditional distributions $\mu(s_1|\lambda, S_1)$ and $\mu(s_2|\lambda, S_2)$ are equal as well. That is, applying equations (14)–(21), we find that we can express the assumption of preparation noncontextuality as
If \( S_1, S_2 \in \mathcal{S} \), then \( \forall \lambda \in \Lambda : \mu(s | \lambda, S_1) = \mu(s | \lambda, S_2) \equiv \mu(s | \lambda, \mathcal{S}) \) \hspace{1cm} (22)

or, translating this into expectation values using equation (16), we can express it as

\[
\text{If } S_1, S_2 \in \mathcal{S}, \text{ then } \forall \lambda \in \Lambda : \langle s | \lambda, S_1 \rangle = \langle s | \lambda, S_2 \rangle \equiv \langle s | \lambda, \mathcal{S} \rangle.
\] \hspace{1cm} (23)

It is apparent, therefore, that the assumption of noncontextuality for sources is a kind of retrodictive analogue of the assumption of noncontextuality for measurements.

An ontic state \( \lambda \) will be said to make a deterministic assignment to an equivalence class of sources \( S \) if

\[
\mu(s | \lambda, S) \in \{0, 1\},
\] \hspace{1cm} (24)

or equivalently, if \( \langle s | \lambda, S \rangle \in \text{Range}(s) \).

We again pause to illustrate these notions by specializing to the quantum case. Consider a quantum source associated to the ensemble \( \mathcal{E} = \{ P_1, P_2 \} \). Suppose that \( \mathcal{E} \) is compatible with an ensemble \( \mathcal{E}_1 = \{ P_1, P_2 \} \) and that it is also compatible with an ensemble \( \mathcal{E}_2 = \{ P_1, P_2 \} \) but that \( \mathcal{E}_1 \) and \( \mathcal{E}_2 \) are not compatible with one another. In this case, there are two operationally equivalent ways of implementing the source associated to \( \mathcal{E} \), namely, by implementing it jointly with \( \mathcal{E}_1 \) and by implementing it jointly with \( \mathcal{E}_2 \). If \( \langle s_1 | \lambda, \mathcal{E}_1 \rangle \) denotes the expectation value for the outcome of the source \( \mathcal{E} \) when it is implemented jointly with \( \mathcal{E}_1 \) and \( \langle s_2 | \lambda, \mathcal{E}_2 \rangle \) denotes the expectation value for the outcome of the source \( \mathcal{E} \) when it is implemented jointly with \( \mathcal{E}_2 \), then preparation noncontextuality implies that \( \forall \lambda \in \Lambda : \langle s_1 | \lambda, \mathcal{E}_1 \rangle = \langle s_2 | \lambda, \mathcal{E}_2 \rangle \equiv \langle s | \lambda, \mathcal{E} \rangle \).

### 2.2.4. Universal noncontextuality

An operational theory will be said to admit of a universally noncontextual ontological model if it admits of an ontological model that is noncontextual for all experimental procedures, and therefore for both the preparations and the measurements [8].

### 3. Quantum no-go theorems based on the Peres–Mermin square

KS-noncontextuality can be understood as the conjunction of the assumption of measurement noncontextuality defined above and the assumption that the ontic state assigns outcomes to projective measurements deterministically. It is well known that quantum theory does not admit of a KS-noncontextual model. The Peres–Mermin proof [16, 17] is particularly intuitive and has therefore become a paradigm example.

Reference [8] provided several reasons for adopting the notion of universal noncontextuality described in the previous section rather than KS-noncontextuality. First of all, the reference to projective measurements in the definition of KS-noncontextuality makes it clear that the notion of KS-noncontextuality can only be applied to quantum theory and leaves open the question of how to apply it to other operational theories or to experimental data. Furthermore, it has been argued in [8, 9] that the notion of universal noncontextuality stands to the notion of KS-noncontextuality as the notion of local causality stands to that of local determinism (defined by Bell in [24]). The problem with both KS-noncontextuality and local determinism is that in the face of a contradiction, one can always salvage the spirit of noncontextuality or locality by simply abandoning determinism. Such an option is not available if one derives a contradiction from universal noncontextuality or local causality.

It was shown in section VIII of [8] that one can turn any no-go theorem for a KS-noncontextual model of quantum theory into a no-go theorem for a universally noncontextual model of quantum theory\(^{10}\). In this section, we carry out this translation for the proof based on the Peres–Mermin square. This will serve to clarify the contrast between KS-noncontextuality and universal noncontextuality. However, the main purpose of this section is to provide the reader with some intuition about how the contradiction arises, so that she may better follow our technique for deriving noncontextuality inequalities from the Peres–Mermin construction.

#### 3.1. No-go theorem for KS-noncontextuality based on the Peres–Mermin square

The Peres–Mermin magic square construction [16, 17] consists of nine observables, each defined on two qubits and each expressible as a product of Pauli operators. Denoting the four Pauli operators by:

\[\begin{array}{cccc}
I & I & I & I \\
I & i \sigma_x & -i \sigma_y & -i \sigma_z \\
i \sigma_x & -i \sigma_y & i \sigma_z & I \\
i \sigma_z & i \sigma_y & -i \sigma_z & I \\
\end{array}\]

\[^{10}\]The proof of this result relies on preparation noncontextuality, together with two facts about quantum theory: (i) for every observable, there is a basis of quantum states each element of which makes its outcome perfectly predictable and (ii) the uniform mixture of any basis of quantum states is operationally equivalent to the uniform mixture of any other such basis.
the nine observables relevant to the construction are as follows:

\[
\begin{align*}
X \otimes 1 & \quad I \otimes X & \quad X \otimes X \\
1 \otimes Z & \quad Z \otimes 1 & \quad Z \otimes Z \\
X \otimes Z & \quad Z \otimes X & \quad Y \otimes Y.
\end{align*}
\]  

These are organized into a 3 \( \times \) 3 grid (the 'square') to visually represent their commutativity properties: the three observables on any row or column of the square commute and therefore are jointly measurable.

KS-noncontextuality implies that every observable in the square is assigned a value deterministically by the ontic state \( \lambda \) and independently of whether that observable is measured together with the other observables in its row or whether it is measured together with the other observables in its column. The deterministic value assigned to observable \( O \) by the ontic state \( \lambda \) we denote by \( \langle O \rangle_\lambda \). Because an observable is only ever found to take values from the eigenspectrum of the associated operator, it follows that the deterministic assignments to \( O \) by \( \lambda \) can only take values in this set,

\[ \langle O \rangle_\lambda \in \text{spec}(O). \]  

Finally, for any set of observables that can be jointly measured, the functional relations that hold among the observables in the set must also hold among the values assigned to them by the ontic state. This follows from the fact that if a given functional relation \( \text{failed} \) to hold for the values assigned by the ontic state, then the ontological model would predict that it failed to hold for the values obtained in a joint measurement. In the Peres–Mermin square, one can show that the product of the observables along each of the rows and along each of the first two columns is \( 1 \otimes 1 \), while the product of the observables on the last column is \( -1 \otimes 1 \). Therefore, the functional relations are

\[
\begin{align*}
(X \otimes 1)(1 \otimes X)(X \otimes X) & = 1 \otimes 1, \\
(1 \otimes Z)(Z \otimes 1)(Z \otimes Z) & = 1 \otimes 1, \\
(X \otimes Z)(Z \otimes X)(Y \otimes Y) & = 1 \otimes 1, \\
(X \otimes 1)(1 \otimes Z)(X \otimes Z) & = 1 \otimes 1, \\
(1 \otimes X)(Z \otimes 1)(Z \otimes X) & = 1 \otimes 1, \\
(X \otimes X)(Z \otimes Z)(Y \otimes Y) & = -1 \otimes 1.
\end{align*}
\]  

Together with the fact, inferred from equation (27), that

\[ \langle 1 \otimes 1 \rangle_\lambda = +1; \quad \langle -1 \otimes 1 \rangle_\lambda = -1, \]  

we conclude that the functional relations holding among the deterministic assignments to the nine observables in the Peres–Mermin square are

\[
\begin{align*}
\langle X \otimes 1 \rangle_\lambda \langle 1 \otimes X \rangle_\lambda \langle X \otimes X \rangle_\lambda & = +1, \\
\langle 1 \otimes Z \rangle_\lambda \langle Z \otimes 1 \rangle_\lambda \langle Z \otimes Z \rangle_\lambda & = +1, \\
\langle X \otimes Z \rangle_\lambda \langle Z \otimes X \rangle_\lambda \langle Y \otimes Y \rangle_\lambda & = +1, \\
\langle X \otimes 1 \rangle_\lambda \langle 1 \otimes Z \rangle_\lambda \langle X \otimes Z \rangle_\lambda & = +1, \\
\langle 1 \otimes X \rangle_\lambda \langle Z \otimes 1 \rangle_\lambda \langle Z \otimes X \rangle_\lambda & = +1, \\
\langle X \otimes X \rangle_\lambda \langle Z \otimes Z \rangle_\lambda \langle Y \otimes Y \rangle_\lambda & = -1.
\end{align*}
\]  

By equation (27), we also have that, for all \( O, O' \in \{ X, Y, Z \} \),

\[ \langle O \otimes 1 \rangle_\lambda \in \{-1, +1\}, \quad \langle 1 \otimes O' \rangle_\lambda \in \{-1, +1\}, \quad \langle O \otimes O' \rangle_\lambda \in \{-1, +1\}. \]  

However, given this constraint, the set of equations (30a)–(30f) has no solution. To see this, it suffices to note that the product of the left-hand sides of the six equations is \( +1 \) (because every term appears squared in this product), while the product of the right-hand sides of the six equations is \( -1 \). We have thereby arrived at a contradiction.

3.2. No-go theorem for universal noncontextuality based on the Peres–Mermin square

As noted in section 2.2.2, unlike KS-noncontextuality, measurement noncontextuality allows for measurement outcomes to be assigned \( \text{indeterministically} \) by the ontic state. Obtaining the contradiction in the previous section relied critically on this assumption of deterministic assignments. Indeed, once one allows
indeterministic assignments, one finds that there are, in fact, many noncontextual assignments to the observables in the Peres–Mermin square. For instance, every quantum state defines such an assignment through the Born rule, and there are other valid assignments as well which do not arise from the Born rule (but could arise in some putative post-quantum theory). At first glance, therefore, it may seem that by replacing the notion of KS-noncontextuality with the generalized notion of noncontextuality proposed in [8], one has lost the possibility of deriving a contradiction. However, although the generalized notion of noncontextuality does indeed weaken the constraints on how one represents measurements in a noncontextual ontological model, it also introduces a novel constraint on how one represents preparations. By availing oneself of the assumption of preparation noncontextuality, one can again derive a no-go theorem for a noncontextual model of quantum theory.

Each observable on a two-level quantum system represents an operational equivalence class of binary-outcome measurements. We denote the observable in position $i, j$ of the Peres–Mermin square by $O_{ij}$. The projector-valued measure associated to the observable $O_{ij}$ consists of the pair of orthogonal rank-2 projectors:

$$\Pi^+_j = \frac{1}{2}(I \otimes I + O_{ij}), \quad \Pi^-_j = \frac{1}{2}(I \otimes I - O_{ij}),$$

(32)

which correspond respectively to the $+1$ and $-1$ eigenspaces of $O_{ij}$.

We also define an equivalence class of binary-outcome quantum sources for each of the observables as follows. For each observable $O_{ij}$, we consider the quantum source associated to the 2-element ensemble

$$\mathcal{E}_{ij} = \left\{ \frac{1}{2} \rho^+_j, \frac{1}{2} \rho^-_j \right\},$$

where

$$\rho^+_j = \frac{1}{2} \Pi^+_j = \frac{1}{4}(I \otimes I + O_{ij}), \quad \rho^-_j = \frac{1}{2} \Pi^-_j = \frac{1}{4}(I \otimes I - O_{ij}),$$

(33)

are normalized density operators. Note that each of these quantum sources defines the same average state, namely, the completely mixed state $\frac{1}{2}I$.

If we arrange these nine quantum sources into a square, then they are compatible along the rows and the columns. For example, in the first row of the square, the three sources are seen to be compatible by virtue of the fact that they can all be obtained by post-processing of the outcome of a single 4-outcome source, namely, the one associated to the uniform ensemble of joint eigenstates of the set of commuting observables associated to that row. The other rows and columns are analogous. We will speak of the source version of the Peres–Mermin square to refer to the compatibility relations among the sources.

Given these definitions, it is clear that when the measurement associated to the observable $O_{ij}$ is implemented on the source associated to the ensemble $\mathcal{E}_{ij}$ (i.e., the source at the same location in the square), the outcome $m$ of the measurement is perfectly correlated with the outcome $s$ of the source and the marginal distribution over either outcome is uniform:

$$\forall i, j : \text{pr}(m, s | O_{ij}, \mathcal{E}_{ij}) = \frac{1}{2} \text{tr} \{ \Pi^m_j \rho^m_j \} = \frac{1}{2} \delta_{m,s}.$$  

(34)

This can be expressed equivalently as

$$\forall i, j : \langle ms \rangle_{O_{ij},E_{ij}} = 1.$$  

(35)

Now consider what this implies for any putative noncontextual ontological model of the experiment. Denoting the expectation value for the outcome $m$ of the measurement of observable $O_{ij}$ given ontic state $\lambda$ by $\langle m \rangle_{\lambda, O_{ij}}$, and the expectation value for the outcome $s$ of the source associated to the ensemble $\mathcal{E}_{ij}$ given ontic state $\lambda$ by $\langle s \rangle_{\lambda, E_{ij}}$ (note that the latter is a retrodictive expectation), then given equation (17), we have

$$\langle ms \rangle_{O_{ij},E_{ij}} = \sum_{\lambda \in \Lambda} \langle m \rangle_{\lambda, O_{ij}} \langle s \rangle_{\lambda, E_{ij}} \mu(\lambda | E_{ij}).$$  

(36)

An assumption of measurement noncontextuality has been made at this stage because we have assumed that the expectation value $\langle ms \rangle_{O_{ij},E_{ij}}$ depends only on $O_{ij}$ and not on what other observables were measured together with it, that is, we have assumed that this expectation value is independent of whether we measure $O_{ij}$ with the other observables in the same row of the Peres–Mermin square or with the other observables in the same column. Similarly, an assumption of preparation noncontextuality has been made at this stage because we have assumed that $\langle s \rangle_{\lambda, E_{ij}}$ depends only on $\mathcal{E}_{ij}$ and not on which set of compatible ensembles are implemented jointly with it, those on the same row of the source version of the Peres–Mermin square or those on the same column.

Finally, the assumption of preparation noncontextuality has an additional consequence that equation (36) does not yet fully incorporate. For each of the nine binary-outcome quantum sources, if one marginalizes over its outcome, one obtains the source that simply prepares $\frac{1}{2}I$, the average state associated to the ensemble. Recall
that no quantum measurement can distinguish the different ensembles by which the completely mixed state might have been prepared. Therefore, for any given pair of quantum sources, \( E_{ij} \) and \( E_{j'j''} \), the pair of quantum sources one obtains by marginalizing over their outcomes are operationally equivalent.

If \( \mu(\lambda, s|E_{ij}) \) is the representation in the ontological model of the quantum source \( E_{ij} \), then the quantum source that one obtains by marginalizing over its outcomes \( s \) is represented in the ontological model by

\[
\mu(\lambda|E_{ij}) = \sum_s \mu(\lambda, s|E_{ij}).
\]

Applying the assumption of preparation noncontextuality, equation (21), to the operational equivalence of \( E_{ij} \) and \( E_{j'j''} \), we obtain

\[
\forall (ij), (j'j''): \mu(\lambda|E_{ij}) = \mu(\lambda|E_{j'j''}) \equiv \mu(\lambda|\frac{1}{2}I).
\]

Substituting this into equation (36), we finally obtain

\[
\langle m \sigma \rangle_{O_{ij}E_{ij}} = \sum_{\lambda \in \Lambda} \langle m \rangle_{\lambda, O_{ij}} \langle \sigma \rangle_{\lambda, E_{ij}} \mu(\lambda|\frac{1}{2}I).
\]

Now we are in a position to derive a contradiction. The only way to reproduce the perfect correlations of equation (35) is if for all \( \lambda \) in the support of \( \mu(\lambda|\frac{1}{2}I) \) and for all \( i, j \),

\[
\langle m \rangle_{\lambda, O_{ij}} \langle \sigma \rangle_{\lambda, E_{ij}} \in \{+1, -1\}, \quad \text{and} \quad \langle m \rangle_{\lambda, O_i} = \langle \sigma \rangle_{\lambda, E_i}.
\]

In other words, every ontic state in the support of \( \mu(\lambda|\frac{1}{2}I) \) must assign perfectly correlated outcomes to the source and measurement when these are associated to the same observable, and the only way to achieve this is if it assigns these outcomes deterministically. However, any deterministic assignment to all of the measurements in the Peres–Mermin square must satisfy the functional relationships that hold among the outcomes of the compatible subsets of those measurements, that is, it must satisfy equations (36a)–(36f). But following the standard argument (reviewed in the previous section), there are no such deterministic assignments, so we have arrived at our contradiction\(^\text{11}\).

It follows that if one entertains the hypothesis that a given experiment is, in fact, described by a noncontextual ontological model, then one expects that for some subset of the nine source–measurement pairs, the correlations will be imperfect. The noncontextuality inequalities that we derive for the Peres–Mermin scenario will capture the precise tradeoffs among the strengths of these nine correlations.

First, however, we must operationalize our description of the experiment, which is to say that we must purge it of any reference to the quantum formalism.

4. From the quantum no-go theorem to noncontextuality inequalities

4.1. A purely operational description of the Peres–Mermin square

In section 2.1, we defined operational equivalence relations and compatibility relations among experimental procedures in a manner that made reference only to experimental statistics, without appeal to the quantum formalism. Here, we use these notions to express the relations must hold among a set of measurements and sources in an operational version of the quantum Peres–Mermin construction. Any experiment satisfying all of these relations will be termed an operational Peres–Mermin scenario.

We start with the measurements. There are 9 distinct equivalence classes of binary-outcome measurements, which we label by \( M_{ij} \) where \( i \in \{1, 2, 3\} \) and \( j \in \{1, 2, 3\} \). Laying these out in a \(3 \times 3\) square, where the measurement \( M_{ij} \) appears at the \( i\)th row and \( j\)th column, each triple of measurements making up a row or a column of the square constitutes a compatible set of measurements. This is depicted in the compatibility hypergraph of figure 1.

By the definition of compatibility for measurements, equation (7), this implies that for every row and column there exists a measurement that simulates all the measurements on that row or column. We denote the measurement that simulates the triple of measurements in row 1 by \( M_{R1} \), the one that simulates the triple in column 1 by \( M_{C1} \), and so forth. We denote their outcomes by \( m_{R1} \), \( m_{C1} \), and so forth.

We now turn to the nature of the particular relation that holds between the measurements on a given row or column and the measurement that simulates them. Consider the measurements in the first row. The outcomes of the simulating measurement \( m_{R1} \) is presumed to be 4-valued, such that it can be presented as an ordered pair of binary outcomes, which we denote by \( m_{R1,1} \) and \( m_{R1,2} \). In terms of this notation, the three measurements in the first row are presumed to be obtained from the simulating measurement by the following identification of

\[^{11}\text{Note that the contradiction could have been obtained equally well by considering the impossibility of finding deterministic assignments to all of the sources while respecting the functional relations that hold among compatible subsets of these.}\]
outcomes,

\begin{align}
m_{11} &= m_{R_1,1}, \\
m_{12} &= m_{R_1,2}, \\
m_{13} &= m_{R_1,1} \cdot m_{R_1,2},
\end{align}

which in terms of the conditional probabilities in equation (7) corresponds to the following post-processings of the simulating measurement:

\begin{align}
pr\{m_{11}|m_{R_1}\} &= \delta_{m_{11},m_{R_1,1}}, \\
pr\{m_{12}|m_{R_1}\} &= \delta_{m_{12},m_{R_1,2}}, \\
pr\{m_{13}|m_{R_1}\} &= \delta_{m_{13},m_{R_1,1} \cdot m_{R_1,2}}.
\end{align}

Analogous compatibility relations hold for the second and third rows and for the first and second column. The relations are slightly different for the third column:

\begin{align}
m_{13} &= m_{C_1,1}, \\
m_{23} &= m_{C_1,2}, \\
m_{33} &= -m_{C_1,1} \cdot m_{C_1,2},
\end{align}

or in terms of the conditional probabilities,

\begin{align}
pr\{m_{13}|m_{C_1}\} &= \delta_{m_{13},m_{C_1,1}}, \\
pr\{m_{23}|m_{C_1}\} &= \delta_{m_{23},m_{C_1,2}}, \\
pr\{m_{33}|m_{C_1}\} &= \delta_{m_{33},-m_{C_1,1} \cdot m_{C_1,2}}.
\end{align}

A similar story holds for the sources. There are 9 distinct equivalence classes of binary-outcome sources, which we label by \(S_{ij}\) where \(i \in \{1, 2, 3\}\) and \(j \in \{1, 2, 3\}\), with compatibility relations described by the hypergraph of figure 2.

By the definition of compatibility for sources, equation (10), this implies that for every row and column there exists a source that simulates all the sources on that row or column. We denote the source that simulates the triple of measurements in row 1 by \(S_{R_1}\) and its outcome by \(s_{R_1}\), the one that simulates the triple in column 1 by \(S_{C_1}\) and its outcome by \(s_{C_1}\), and so forth. Each such outcome is presumed to be 4-valued, such that it can be presented as an ordered pair of binary variables, so that \(s_{R_1} = (s_{R_1,1}, s_{R_1,2})\), etcetera. The conditional probabilities, which, by equation (10), define the precise nature of the compatibility relations are exactly the same as for the measurements. For the first row, they are

\begin{align}
pr\{a_1|s_{R_1}\} &= \delta_{a_1, s_{R_1,1}}, \\
pr\{a_2|s_{R_1}\} &= \delta_{a_2, s_{R_1,2}}, \\
pr\{a_3|s_{R_1}\} &= \delta_{a_3, s_{R_1,1} \cdot s_{R_1,2}},
\end{align}

with analogous relations holding for the other rows and the first and second column, while for the third column, they are

![Figure 1. The hypergraph representing compatibility relations among the nine binary-outcome measurements.](image)
4.1.1. Noiseless quantum realization

It is straightforward to verify that the quantum measurements and quantum sources appearing in the no-go theorem described in section 3.2 instantiate all of the compatibility relations that were described in the previous section.

We begin with the quantum measurements. Take, for example, the three observables in the first row of the Peres–Mermin square. These are associated to the projector-valued measures \( m_{11}, m_{12}, m_{13} \) where

\[
\begin{align*}
\text{pr}(s_{13}|a_{C}) & = \delta_{a_{C},c_{1}}, \\
\text{pr}(s_{23}|a_{C}) & = \delta_{a_{C},c_{2}}, \\
\text{pr}(s_{33}|a_{C}) & = \delta_{a_{C},c_{3}}.
\end{align*}
\]

(46a) \hspace{1cm} (46b) \hspace{1cm} (46c)

The simulation of each of the three measurements in the row is achieved by implementing this PVM and then post-processing its outcome using the three conditional probability distributions specified in equations (42a)–(42c), that is,

\[
\begin{align*}
\Pi_{11}^{(m_{11} \cdot m_{12})} & = \sum_{m_{11},m_{12}} \delta_{m_{11},m_{12}} \Pi_{R_{1}}^{(m_{11} \cdot m_{12})}, \\
\Pi_{12}^{(m_{12} \cdot m_{13})} & = \sum_{m_{12},m_{13}} \delta_{m_{12},m_{13}} \Pi_{R_{1}}^{(m_{12} \cdot m_{13})}, \\
\Pi_{13}^{(m_{13} \cdot m_{12})} & = \sum_{m_{13},m_{12}} \delta_{m_{13},m_{12}} \Pi_{R_{1}}^{(m_{13} \cdot m_{12})}.
\end{align*}
\]

(48a) \hspace{1cm} (48b) \hspace{1cm} (48c)

In a similar fashion, one can verify that the other rows and columns of the Peres–Mermin square of quantum measurements have the compatibility relations described in the previous section.

The nine quantum sources appearing in the no-go theorem of section 3.2 also have the compatibility relations described in the previous section. Consider the first row of the source version of the Peres–Mermin square as an example. The three sources on this row are associated to the ensembles \( s_{11}, s_{12}, s_{13} \) where

\[
\begin{align*}
\text{pr}(s_{13}|a_{C}) & = \delta_{a_{C},c_{1}}, \\
\text{pr}(s_{23}|a_{C}) & = \delta_{a_{C},c_{2}}, \\
\text{pr}(s_{33}|a_{C}) & = \delta_{a_{C},c_{3}}.
\end{align*}
\]

(46a) \hspace{1cm} (46b) \hspace{1cm} (46c)

(47)

The simulation of each of the three measurements in the row is achieved by implementing this PVM and then post-processing its outcome using the three conditional probability distributions specified in equations (42a)–(42c), that is,

\[
\begin{align*}
\Pi_{11}^{(m_{11} \cdot m_{12})} & = \sum_{m_{11},m_{12}} \delta_{m_{11},m_{12}} \Pi_{R_{1}}^{(m_{11} \cdot m_{12})}, \\
\Pi_{12}^{(m_{12} \cdot m_{13})} & = \sum_{m_{12},m_{13}} \delta_{m_{12},m_{13}} \Pi_{R_{1}}^{(m_{12} \cdot m_{13})}, \\
\Pi_{13}^{(m_{13} \cdot m_{12})} & = \sum_{m_{13},m_{12}} \delta_{m_{13},m_{12}} \Pi_{R_{1}}^{(m_{13} \cdot m_{12})}.
\end{align*}
\]

(48a) \hspace{1cm} (48b) \hspace{1cm} (48c)

In a similar fashion, one can verify that the other rows and columns of the Peres–Mermin square of quantum measurements have the compatibility relations described in the previous section.

The nine quantum sources appearing in the no-go theorem of section 3.2 also have the compatibility relations described in the previous section. Consider the first row of the source version of the Peres–Mermin square as an example. The three sources on this row are associated to the ensembles \( s_{11}, s_{12}, s_{13} \) where

\[
\begin{align*}
\text{pr}(s_{13}|a_{C}) & = \delta_{a_{C},c_{1}}, \\
\text{pr}(s_{23}|a_{C}) & = \delta_{a_{C},c_{2}}, \\
\text{pr}(s_{33}|a_{C}) & = \delta_{a_{C},c_{3}}.
\end{align*}
\]

(46a) \hspace{1cm} (46b) \hspace{1cm} (46c)

(47)

The simulation of each of the three measurements in the row is achieved by implementing this PVM and then post-processing its outcome using the three conditional probability distributions specified in equations (42a)–(42c), that is,
probabilities appearing in the simulation are precisely those given in equations (45a)–(45c). This fact follows
from equations (48a)–(48c). The compatibility relations for the other rows and columns are verified similarly.

4.1.2. Noisy quantum realization
In deriving noncontextuality inequalities, it is critical that one not base these on assumptions that are only valid
when the measurements or the sources are noiseless because this ideal is never achieved in real experiments. The
compatibility relations outlined in section 4.1 satisfy this desideratum. In the quantum case, for instance, they
can be satisfied even if the measurements and sources are not sharp (i.e. not associated to an orthogonal set of
projectors).

A specific example helps to clarify the point.

Suppose the nine sharp measurements appearing in the Peres–Mermin square are replaced by noisy versions
thereof, that is, by the nine unsharp measurements that are the images of the sharp measurements under a partially
depolarizing channel \( \mathcal{D} \). In this case, the projector-valued measures are replaced by POVMs that are not projective.
For instance, the three measurements in the first row of the Peres–Mermin square are associated to the binary-
outcome POVMs \( \{ \mathcal{D}(\Pi_{1i}^{(m)}) \}_{m} \) and \( \{ \mathcal{D}(\Pi_{13}^{(m)}) \}_{m} \), where \( m_{11}, m_{12}, m_{13} \in \{ -1, +1 \} \). These can be jointly implemented using the 4-outcome POVM \( \{ \mathcal{D}(\Pi_{R_{1i}^{(m)}, m_{13}}) \}_{m_{1i}, m_{13}} \) where \( (m_{R_{1i}}, m_{R_{13}}) \in \{ -1, +1 \} \). The three binary-outcome POVMs are simulated by the 4-outcome POVM using the conditional probabilities in equations (42a)–(42c); to see this, it suffices to apply \( \mathcal{D} \) to equations (48a)–(48c) and recall that it is a linear map.

Similarly, suppose that the nine sources appearing in the source version of the Peres–Mermin square are replaced by partially depolarized versions thereof. (For simplicity, we will assume that the strength of the noise on the sources is equal to that on the measurements.) In this case, the ensembles associated to the three sources in
the first row are \( \{ \frac{1}{2} \mathcal{D}(\rho_{11}^{(c)}) \}_{s_{11}} \), \( \{ \frac{1}{2} \mathcal{D}(\rho_{12}^{(c)}) \}_{s_{12}} \) and \( \{ \frac{1}{2} \mathcal{D}(\rho_{13}^{(c)}) \}_{s_{13}} \), where \( s_{11}, s_{12}, s_{13} \in \{ -1, +1 \} \). The source that simulates all of these is then simply the partially depolarized version of the one that simulated the sharp sources, that is, \( \{ \frac{1}{4} \mathcal{D}(\rho_{R_{1i}^{(c)}, s_{13}^{(c)}}) \}_{(s_{1i}, s_{13}^{(c)})} \), which again follows from the linearity of equations (48a)–(48c).

Recall that a partial depolarization map \( \mathcal{D} \) can be written as a convex mixture of the identity channel, \( I \), and the channel that traces over the system and reprepares the completely mixed state. An element of the
1-parameter family of such maps is

\[
\mathcal{D}_{r} = r I + (1 - r) \frac{1}{4} I_{Tr},
\]

where \( r \in [0, 1] \). The strength of depolarization is specified by the probability \( r \) of realizing the identity map
(with lower values of \( r \) corresponding to stronger noise).

It follows that the degree of correlation that can be observed between sources and measurements is a function of \( r \). For \( r < 1 \), one no longer achieves the perfect correlations of the noiseless quantum realization, equations (34), (35), but rather imperfect correlations. Denoting the \( r \)-depolarized versions of the observables \( O_{y} \) and the ensembles \( E_{y} \) by \( O_{y}^{(r)} \) and \( E_{y}^{(r)} \) respectively, we have

\[
\forall i, j : \quad \Pr(m_{y}, a_{y}|O_{y}^{(r)}, E_{y}^{(r)}) = \frac{1}{2} \text{tr}(\mathcal{D}(\Pi_{y}^{(m)}) \mathcal{D}(\rho_{y}^{(c)})) = \frac{1}{2} r^{2} \delta_{m_{y}, a_{y}}.
\]

This can be expressed equivalently as

\[
\forall i, j : \quad \langle m_{y} a_{y}|O_{y}^{(r)}, E_{y}^{(r)} \rangle = r^{2}.
\]

Because the no-go theorem of section 3.2 relied on having perfect correlations, it is not applicable to the
noisy quantum realization of the operational Peres–Mermin scenario. Nonetheless, one expects that for values
of \( r \) sufficiently close to 1, a noncontextual model should still be ruled out. The noncontextual inequalities that
we derive confirm this expectation. They are robust to noise in the sense that they can be violated by values of \( r \)
strictly less than 1. The lower bound on \( r \) they imply is determined in section 4.6.2. This bound specifies how
much noise one can tolerate in the noisy quantum realization of the Peres–Mermin scenario and still rule out a
noncontextual model of the experiment.

4.2. Expressing operational correlations in terms of noncontextual ontic assignments
Consider an experiment that can realize the nine equivalence classes of measurements and the nine equivalence
classes of sources having the compatibility structures of figures 1 and 2 respectively and having the compatibility
relations specified in the text, such as equations (42a)–(42c) and (45a)–(45c). There are 81 possible pairings of a
source with a measurement. For a given such pairing, say \( S_{ij} \) with \( M_{ij} \), the experiment yields a joint probability
distribution over outcomes, \( \Pr(m_{ij}, a_{ij}|M_{ij}, S_{ij}) \). Equivalently, the experimental data can be summarized by the
expectation values \( \langle s_{ij}|s_{ij}, \langle m_{ij}|M_{ij}, S_{ij} \rangle \) and \( \langle m_{ij}|S_{ij} \rangle \).
In this article, we limit our focus to deriving constraints which do not refer to the marginal expectations \((s_i|m_{ij})_{S,M_{ij}}\) and \((m_{ij})_{S,M_{ij}}\), i.e. we focus on deriving inequalities which refer only to the correlations \(\langle s_i m_{ij} \rangle_{S,M_{ij}}\). Furthermore, we consider only 9 of the 81 possible pairings of a source with a measurement, namely those wherein the source and the measurement are associated with a common label in their respective compatibility hypergraphs. That is, we hereafter consider only those correlations \(\langle s_i m_{ij} \rangle_{S,M_{ij}}\), wherein \((i',j') = (i,j)\). We will derive the necessary and sufficient conditions—with respect to these nine correlations—for an experiment to admit a noncontextual model.

For the equivalence class of measurements \(M_{ij}\), there are two associated measurement procedures, which we denote by \(M_{ij}^R\) and \(M_{ij}^C\), with outcomes denoted by \(m_{ij}^R\) and \(m_{ij}^C\), and which correspond to whether \(M_{ij}\) is implemented jointly with the other measurements in its row or with the other measurements in its column. Similarly, the equivalence class of sources \(S_{ij}\) is associated with two sources, \(S_{ij}^R\) and \(S_{ij}^C\), with outcomes denoted by \(s_{ij}^R\) and \(s_{ij}^C\). The operational equalities imply that \(\langle s_{ij}^R m_{ij}^R \rangle_{S_{ij}^R M_{ij}^R} = \langle s_{ij}^C m_{ij}^C \rangle_{S_{ij}^C M_{ij}^C} = \langle s_{ij}^C m_{ij}^R \rangle_{S_{ij}^C M_{ij}^R} = \langle s_{ij}^R m_{ij}^C \rangle_{S_{ij}^R M_{ij}^C}\), which we can simply denote by \(\langle s_i m_{ij} \rangle_{S_{ij} M_{ij}}\).

Recall equation (17), which specifies how correlations such as \(\langle s_i m_{ij} \rangle_{S_{ij} M_{ij}}\) are expressed in an ontological model. Under the assumption of measurement noncontextuality, every measurement in the equivalence class \(M_{ij}\) is assigned the same expectation value by the ontic state. Therefore, because \(M_{ij}^R, M_{ij}^C \in M_{ij}\), it follows that

\[
\langle m_{ij}^R \rangle_{\lambda, M_{ij}} = \langle m_{ij}^C \rangle_{\lambda, M_{ij}} = \langle m_{ij} \rangle_{\lambda, M_{ij}}.
\]

(52)

In other words, measurement noncontextuality warrants the assumption that the expectation value for the outcome of \(M_{ij}\) does not depend on the measurement context, that is, whether it is implemented with the measurements in the same row or in the same column of figure 1. Similarly, under the assumption of preparation noncontextuality, every source in the equivalence class \(S_{ij}\) is assigned the same retrodictive expectation value by the ontic state, such that because \(S_{ij}^R, S_{ij}^C \in S_{ij}\), it follows that

\[
\langle s_{ij}^R \rangle_{\lambda, S_{ij}} = \langle s_{ij}^C \rangle_{\lambda, S_{ij}} = \langle s_i \rangle_{\lambda, S_{ij}}.
\]

(53)

In other words, the expectation value for the outcome of \(S_{ij}\) does not depend on the source context, that is, whether it is implemented with the sources in the same row or in the same column of figure 2.

We conclude that

\[
\langle s_i m_{ij} \rangle_{S_{ij} M_{ij}} = \sum_{\lambda \in \Lambda} \langle s_i \rangle_{\lambda, S_{ij}} \langle m_{ij} \rangle_{\lambda, M_{ij}} \mu(\lambda | S_{ij}).
\]

(54)

The operational equivalence relations among the sources and the assumption of preparation noncontextuality together imply one further simplification of this expression, namely, that \(\mu(\lambda | S_{ij})\) is independent of \((i,j)\),

\[
\forall (i;j';j''): \mu(\lambda | S_{ij}) = \mu(\lambda | S_{i'j'}) \equiv \mu(\lambda).
\]

(55)

To see why this is the case, consider the triple of sources in the first row of figure 2, \(S_{11}, S_{12}\) and \(S_{13}\). By assumption, these are each simulatable by a single source, namely, \(S_R\), by post-processing its outcome in the manner specified by the compatibility relations, equations (45a)–(45c). Marginalizing over the outcome of \(S_{11}, S_{12}\) or \(S_{13}\) is simply a further post-processing of \(S_R\) and consequently the outcome-marginalized versions of these three sources are each operationally equivalent to the outcome-marginalized version of \(S_R\) and therefore operationally equivalent to one another. The assumption of preparation noncontextuality then implies that the distributions over ontic states associated to these, namely that the three ontic state distributions

\[
\mu(\lambda | S_{11}) = \sum_{s_{11}} \mu(s_{11}, \lambda | S_{11}),
\]

\[
\mu(\lambda | S_{12}) = \sum_{s_{12}} \mu(s_{12}, \lambda | S_{12}),
\]

\[
\mu(\lambda | S_{13}) = \sum_{s_{13}} \mu(s_{13}, \lambda | S_{13})
\]

are equal,

\[
\mu(\lambda | S_{11}) = \mu(\lambda | S_{12}) = \mu(\lambda | S_{13}).
\]

(56)

The same argument repeated for the other rows and the columns yields analogous equalities. Together these imply equation (55).

Denoting \(\langle m_{ij} \rangle_{\lambda, M_{ij}}\) simply as \(\langle m_i \rangle_{\lambda}\) and \(\langle s_i \rangle_{\lambda, S_{ij}}\) simply as \(\langle s_i \rangle_{\lambda}\), and using equation (55), we find that equation (54) becomes

\[
\langle s_i m_{ij} \rangle_{S_{ij} M_{ij}} = \sum_{\lambda \in \Lambda} \langle s_i \rangle_{\lambda} \langle m_{ij} \rangle_{\lambda} \mu(\lambda).
\]

(57)

We pause here to note that this expression for the correlation between the measurement outcome and the source outcome in a noncontextual model has the same form as the expression for the correlation between the
measurement outcomes at the two wings of a Bell experiment in a locally causal model of the latter. This provides a particularly intuitive demonstration of the isomorphism between the assumption of local causality and the assumption of preparation noncontextuality for the outcome-marginalized sources articulated in equation (55). Note, however, that the assumptions of noncontextuality articulated in equations (52), (53) cannot be inferred from an assumption of local causality in the corresponding Bell scenario, so that the noncontextuality inequalities that we derive here are not isomorphic to Bell inequalities.12

The compatibility relations among the measurements imply constraints on the $\langle m_3 \rangle_\lambda$. We will refer to any 9-tuple of expectation values, $(\langle m_1 \rangle_\lambda, \langle m_2 \rangle_\lambda, \ldots, \langle m_9 \rangle_\lambda)$, satisfying these constraints as a noncontextual ontic assignment to the measurements. We will see that the set of all such 9-tuples defines a polytope in a 9-dimensional space, which we term the noncontextual measurement-assignment polytope. Similarly, the compatibility relations among the sources imply constraints on the $\langle s_3 \rangle_\lambda$. We will refer to any 9-tuple of expectation values, $(\langle s_1 \rangle_\lambda, \langle s_2 \rangle_\lambda, \ldots, \langle s_9 \rangle_\lambda)$, satisfying these constraints as a noncontextual ontic assignment to the sources. These also form a polytope, which we term the noncontextual source-assignment polytope. The vertices of a polytope, i.e. the extremal noncontextual ontic assignments, can be deduced from that polytope’s defining constraints using standard convex hull algorithms [25–27].

Every ontic state $\lambda$ specifies some noncontextual assignment to measurements, but not every noncontextual assignment corresponds to a vertex of the noncontextual measurement-assignment polytope. Nevertheless, those non-vertex noncontextual assignments to measurements can be simulated by a distribution over ontic states that do correspond to vertices; suppose $\kappa$ is a variable that runs over the vertices of the noncontextual measurement-assignment polytope. Then, for any $\lambda$ in the polytope, there exists a distribution $p(\kappa | \lambda)$ such that

$$
(\langle m_1 \rangle_\lambda, \langle m_2 \rangle_\lambda, \langle m_3 \rangle_\lambda) = \sum_\kappa (\langle m_1 \rangle_\kappa, \langle m_2 \rangle_\kappa, \langle m_3 \rangle_\kappa) p(\kappa | \lambda),
$$

where we have presented the 9-tuple as a $3 \times 3$ array.

A similar arguments holds for the noncontextual source-assignment polytope. Denoting its vertices by $\kappa'$, for any point $\lambda$ in the noncontextual source-assignment polytope, there exists a distribution $p(\kappa' | \lambda)$ such that

$$
(\langle s_1 \rangle_\lambda, \langle s_2 \rangle_\lambda, \langle s_3 \rangle_\lambda) = \sum_{\kappa'} (\langle s_1 \rangle_{\kappa'}, \langle s_2 \rangle_{\kappa'}, \langle s_3 \rangle_{\kappa'}) p(\kappa' | \lambda).
$$

It is useful to introduce a simplified notation for the nine operational correlations in which we are interested, namely,

$$
\omega_{ij} \equiv \langle s_j m_i \rangle_{S_i M_j}.
$$

The set of 9-dimensional vectors $(\omega_{ij})$ that can arise in an operational theory that admits of a noncontextual ontological model will be termed the noncontextual correlation polytope. Recalling equation (57), and representing the $\omega_{ij}$ as a $3 \times 3$ array, it is defined in terms of the noncontextual measurement-assignment polytope and the noncontextual source-assignment polytope as follows:

$$
\omega_{ij} = \sum_\lambda (\langle m_1 \rangle_\lambda, \langle m_2 \rangle_\lambda, \langle m_3 \rangle_\lambda) \odot (\langle s_1 \rangle_\lambda, \langle s_2 \rangle_\lambda, \langle s_3 \rangle_\lambda) \mu(\lambda),
$$

where $\odot$ denotes the entry-wise product of the arrays (also known as the Hadamard or Schur product).

Substituting equations (58) and (59), we have

$$
\omega_{ij} = \sum_{\kappa, \kappa'} (\langle m_1 \rangle_{\kappa}, \langle s_1 \rangle_{\kappa'}) (\langle m_2 \rangle_{\kappa}, \langle s_2 \rangle_{\kappa'}) (\langle m_3 \rangle_{\kappa}, \langle s_3 \rangle_{\kappa'}) \times p(\kappa, \kappa')
$$

where $p(\kappa, \kappa') \equiv \sum_\lambda p(\kappa | \lambda) p(\kappa' | \lambda) \mu(\lambda)$.

Therefore, the noncontextual correlation polytope is the convex hull of the correlations one obtains for all possible pairings $(\kappa, \kappa')$ of a vertex $\kappa$ from the noncontextual measurement-assignment polytope and a vertex

12 In particular, the noncontextuality inequalities we derive here are not isomorphic to the Bell inequality derived in [38] (and experimentally tested in [39]) even though the latter is inspired by a consideration of the Peres–Mermin construction (one such construction on each wing of the Bell experiment). This is because the inequality of [38] is based on the assumption of local causality alone. The analogue, for our prepare-and-measure scenario, of this inequality would be a constraint that follows from the assumption of preparation noncontextuality for the outcome-marginalized sources alone, equation (55).
\( \kappa' \) from the noncontextual source-assignment polytope, that is, the convex hull of the 9-tuples \( \{ (m_{11}, s_{11}), \ldots, (m_{33}, s_{33}) \} \), as one varies over \( (\kappa, \kappa') \).

Not every pairing of \( (\kappa, \kappa') \) corresponds to a unique vertex of the noncontextual correlation polytope; the fact that we consider correlations for only 9 source-measurement pairings and not the full set of 81 such pairings, and the fact that we do not consider any of the marginal expectations, implies that (i) more than one choice of \( (\kappa, \kappa') \) can yield the same 9-tuple of correlations, and (ii) one choice of \( (\kappa, \kappa') \) can yield a 9-tuple of correlations that lies in the convex hull of the 9-tuples associated to several other choices of \( (\kappa, \kappa') \). It is therefore convenient to re-express equation (62) simply as

\[
\begin{pmatrix}
\omega_1 & \omega_2 & \omega_3 \\
\omega_1' & \omega_2' & \omega_3' \\
\omega_1'' & \omega_2'' & \omega_3''
\end{pmatrix} = \sum_{\gamma} \begin{pmatrix}
(m_{11})\gamma & (s_{11})\gamma & (m_{22})\gamma & (s_{22})\gamma & (m_{33})\gamma & (s_{33})\gamma
\end{pmatrix} p(\gamma),
\]

where instead of ranging over all pairings of \( (\kappa, \kappa') \) we restrict \( \gamma \) to range over the vertices of the noncontextual correlation polytope without loss of generality.

We ultimately seek to derive noncontextuality inequalities, that is, the nontrivial facet inequalities of the noncontextual correlation polytope. We begin by characterizing the noncontextual measurement-assignment polytope and the noncontextual source-assignment polytope. We will see that the nature of the compatibility relations among the measurements/sources determines their respective facet inequalities. From these, we infer the two sets of vertices (measurements/sources) using standard convex hull algorithms [25–28]. Subsequently, by considering every possible pairing between those two sets of vertices, we determine the set of vertices of the noncontextual correlation polytope. Finally, using standard convex hull algorithms again, we obtain all of the facet inequalities of the noncontextual correlation polytope. The nontrivial facet inequalities define the set of noncontextuality inequalities for our problem. In the following sections, we proceed through these various steps explicitly.

### 4.3. Facets of the noncontextual measurement-assignment and noncontextual source-assignment polytopes

We will begin with the measurements. The compatibility relations holding among the measurements in a given row or column must also hold for the response functions representing these in the ontological model. (This is a constraint on any ontological model, rather than one arising from the assumption of measurement noncontextuality. See section 5 for further discussion.)

Consider the response functions associated to the three equivalence classes of measurements in the first row of figure 1. We denote the set of response functions associated to each of these by \( \{ \xi_1(m_{11}|\lambda) \}_{m_{11}} \), \( \{ \xi_2(m_{12}|\lambda) \}_{m_{12}} \), and \( \{ \xi_3(m_{13}|\lambda) \}_{m_{13}} \), where \( m_{11}, m_{12}, m_{13} \in \{-1, +1\} \), and we denote the set of response functions associated to the measurement that simulates these \( \{ \xi_R(m_{R,1}, m_{R,2}|\lambda) \}_{m_{R,1},m_{R,2}} \), where \( m_{R,1}, m_{R,2} \in \{-1, +1\} \). The fact that the simulation is achieved with the three conditional probability distributions specified in equations (42a)–(42c) implies that

\[
\begin{align*}
\xi_1(m_{11}|\lambda) &= \sum_{m_{R,1},m_{R,2}} \delta_{m_{11},m_{R,1}} \xi_R(m_{R,1}, m_{R,2}|\lambda), \\
\xi_2(m_{12}|\lambda) &= \sum_{m_{R,1},m_{R,2}} \delta_{m_{12},m_{R,1}} \xi_R(m_{R,1}, m_{R,2}|\lambda), \\
\xi_3(m_{13}|\lambda) &= \sum_{m_{R,1},m_{R,2}} \delta_{m_{13},m_{R,1}} \xi_R(m_{R,1}, m_{R,2}|\lambda).
\end{align*}
\]

Recalling equations (13) and (41), we infer that

\[
\begin{align*}
\langle m_{11}\rangle &= \xi_R(++)|\lambda| + \xi_R(+-|\lambda|) - \xi_R(-+|\lambda|) - \xi_R(---|\lambda|), \\
\langle m_{12}\rangle &= \xi_R(++)|\lambda| - \xi_R(+-|\lambda|) + \xi_R(-+|\lambda|) - \xi_R(---|\lambda|), \\
\langle m_{13}\rangle &= \xi_R(++)|\lambda| - \xi_R(+-|\lambda|) - \xi_R(-+|\lambda|) + \xi_R(---|\lambda|).
\end{align*}
\]

Using the fact that \( \xi_R \) is a normalized probability distribution,

\[
1 = \xi_R(++)|\lambda| + \xi_R(+-|\lambda|) + \xi_R(-+|\lambda|) + \xi_R(---|\lambda|),
\]

the relations equations (66a)–(66c), (67) can be inverted to write down the probabilities in terms of the expectation values: for \( a, b \in \{-1, +1\} \),

\[
\xi_R(a, b|\lambda) = \frac{1}{4}(1 + a \langle m_{11}\rangle + b \langle m_{12}\rangle + ab \langle m_{13}\rangle).
\]
Finally, from the fact that
\[ \forall a, b : \xi_{R_{1}}(a, b|\lambda) \geq 0, \]  
we infer that \( \forall a, b : \)
\[ -a \langle m_{1}\lambda \rangle - b \langle m_{2}\lambda \rangle - ab \langle m_{3}\lambda \rangle \leq 1. \]

This logic can be repeated for the other rows and the first two columns. For the third column, it is slightly different. We find
\[ \xi_{C_{3}}(a, b|\lambda) = \frac{1}{4} (1 + a \langle m_{1}\lambda \rangle + b \langle m_{2}\lambda \rangle - ab \langle m_{3}\lambda \rangle) \]
from which we infer that \( \forall a, b : \)
\[ -a \langle m_{1}\lambda \rangle - b \langle m_{2}\lambda \rangle + ab \langle m_{3}\lambda \rangle \leq 1. \]

In all then, we find that the 9 expectation values \( \{ \langle m_{ij}\lambda \rangle : (i, j) \in \{1, 2, 3\}^2 \} \) must satisfy
\[ \forall i \in \{1, 2, 3\} : \langle m_{i1}\lambda \rangle + \langle m_{i2}\lambda \rangle - \langle m_{i3}\lambda \rangle \leq 1, \quad \langle m_{i1}\lambda \rangle - \langle m_{i2}\lambda \rangle + \langle m_{i3}\lambda \rangle \leq 1, \quad \langle m_{i1}\lambda \rangle + \langle m_{i2}\lambda \rangle + \langle m_{i3}\lambda \rangle \leq 1, \quad \langle m_{i1}\lambda \rangle - \langle m_{i2}\lambda \rangle - \langle m_{i3}\lambda \rangle \leq 1, \]
and
\[ \forall j \in \{1, 2, 3\} : \langle m_{j1}\lambda \rangle + \langle m_{j2}\lambda \rangle - \eta_j \langle m_{j3}\lambda \rangle \leq 1, \quad \langle m_{j1}\lambda \rangle - \langle m_{j2}\lambda \rangle + \eta_j \langle m_{j3}\lambda \rangle \leq 1, \quad \langle m_{j1}\lambda \rangle + \langle m_{j2}\lambda \rangle + \eta_j \langle m_{j3}\lambda \rangle \leq 1, \quad \langle m_{j1}\lambda \rangle - \langle m_{j2}\lambda \rangle - \eta_j \langle m_{j3}\lambda \rangle \leq 1, \]
where \( \eta_j = \begin{cases} +1 & \text{for } j \in \{1, 2\} \\ -1 & \text{for } j = 3 \end{cases} \).

Note that these inequalities subsume the constraint that \(-1 \leq \langle m_{ij}\lambda \rangle \leq +1 \) for all \( i, j \). Equations (71a) and (71b) capture all of the consequences of the compatibility relations for the ontic expectation values. They are the facet inequalities for the noncontextual measurement-assignment polytope.

Similar constraints hold for the nine retrodictive expectation values \( \{ \langle s_{ij}\lambda \rangle : (i, j) \in \{1, 2, 3\}^2 \} \), as we now show.

Consider the the three equivalence classes of sources in the first row of figure 2. We will denote the probability distributions associated to each of these by \(\mu_{11}(s_{11}, \lambda), \mu_{12}(s_{12}, \lambda), \) and \(\mu_{13}(s_{13}, \lambda),\) where \(s_{11}, s_{12}, s_{13} \in \{-1, +1\}\), and the probability distribution associated to the source that simulates these by \(\mu_{R}(s_{R1}, s_{R2}, \lambda)\) where \(s_{R1}, s_{R2} \in \{-1, +1\}\). The fact that these sources obey the compatibility relations given in equations (45a)–(45c) implies that
\[
\mu_{11}(s_{11}, \lambda) = \sum_{s_{R1}, s_{R2}} \delta_{s_{11}, s_{R1}} \mu_{R}(s_{R1}, s_{R2}, \lambda),
\]
\[
\mu_{12}(s_{12}, \lambda) = \sum_{s_{R1}, s_{R2}} \delta_{s_{12}, s_{R2}} \mu_{R}(s_{R1}, s_{R2}, \lambda),
\]
\[
\mu_{13}(s_{13}, \lambda) = \sum_{s_{R1}, s_{R2}} \delta_{s_{13}, s_{R1} + s_{R2}} \mu_{R}(s_{R1}, s_{R2}, \lambda).
\]

Note that these expressions allow one to confirm the equality of \(\sum_{s_{11}} \mu(\lambda | S_{11}), \sum_{s_{12}} \mu(\lambda | S_{12}), \) and \(\sum_{s_{13}} \mu(\lambda | S_{13}),\) noted in equation (56), via their equality with \(\sum_{s_{R1}, s_{R2}} \mu_{R}(s_{R1}, s_{R2}, \lambda)\). Recalling equation (55), these outcome-marginalized probability distributions are in fact equal to those associated to every other source in the problem, and we have denoted this unique distribution by \(\mu(\lambda)\). It follows that we can Bayesian invert all of the terms in equations (72a)–(72c) by dividing each equation by \(\mu(\lambda)\). We thereby obtain
\[
\mu_{11}(s_{11}|\lambda) = \sum_{s_{R1}, s_{R2}} \delta_{s_{11}, s_{R1}} \mu_{R}(s_{R1}, s_{R2}|\lambda),
\]
\[
\mu_{12}(s_{12}|\lambda) = \sum_{s_{R1}, s_{R2}} \delta_{s_{12}, s_{R2}} \mu_{R}(s_{R1}, s_{R2}|\lambda),
\]
\[
\mu_{13}(s_{13}|\lambda) = \sum_{s_{R1}, s_{R2}} \delta_{s_{13}, s_{R1}, s_{R2}} \mu_{R}(s_{R1}, s_{R2}|\lambda).
\]

Using these relations, together with equation (16), we can express the expectation values \(\{ \langle s_{ij}\lambda \rangle : (i, j) \in \{1, 2, 3\}^2 \} \) in terms of the \(\mu_{R}(s_{R1}, s_{R2}|\lambda)\). By appealing to the fact that \(\mu_{R}\) is a normalized probability distribution, we can invert these equations and then use \(\forall a, b : \mu_{R}(a, b|\lambda) \geq 0\) to obtain inequality constraints on the \(\{ \langle s_{ij}\lambda \rangle : (i, j) \in \{1, 2, 3\}^2 \} \). The analysis proceeds precisely in analogy with the case of measurements, and we obtain:
In this section, we describe the conversion from the facet representation of the noncontextual measurement-assignment polytope, denoted by the facet inequalities of equations \(71a\) and \(71b\), to its vertex representation. We use standard numerical algorithms to do so\(^{25–28}\), the details of which are provided in Appendix A. In addition to providing a description of this set of vertices, it is our aim here to provide some intuitions about their form.

To begin with, note that all of the points within the noncontextual measurement-assignment polytope are indeterministic assignments—in the sense of violating equation \(20\)—for one or more of the measurements. To see that there are no noncontextual ontic assignments that are deterministic for all of the measurements, it suffices to note that for deterministic assignments, the constraints \((71a)\) and \((71b)\) simplify to

\[
\begin{align*}
|m_{11}|_{\lambda}, |m_{12}|_{\lambda}, |m_{13}|_{\lambda} &= +1, \\
|m_{21}|_{\lambda}, |m_{22}|_{\lambda}, |m_{13}|_{\lambda} &= +1, \\
|m_{31}|_{\lambda}, |m_{32}|_{\lambda}, |m_{33}|_{\lambda} &= +1, \\
|m_{11}|_{\lambda}, |m_{12}|_{\lambda}, |m_{33}|_{\lambda} &= +1, \\
|m_{12}|_{\lambda}, |m_{22}|_{\lambda}, |m_{33}|_{\lambda} &= +1, \\
|m_{13}|_{\lambda}, |m_{23}|_{\lambda}, |m_{33}|_{\lambda} &= -1,
\end{align*}
\]

where \(|m_{ij}|_{\lambda}\) denotes a deterministic assignment by ontic state \(\lambda\), and that these are equivalent to the constraints specified in equations \((30a)–(30f)\), which, as noted in section 3.1 admit no solution.

To get a feeling for how indeterministic noncontextual ontic assignments to the measurements escape contradiction, it is useful to see a concrete example (one that is a vertex of the noncontextual measurement-assignment polytope). We denote it by \(\nu_1\). We begin by describing it in terms of probabilistic assignments to the 4-outcome measurements associated to each row and column, rather than in terms of the expectation values for each of the nine equivalence classes of binary-outcome measurements, because the correlations that hold between the different binary-outcome measurements are more transparent in this form. Introducing the notation \(p(m) = [\alpha]\) as a shorthand for \(p(m) = \delta_{m,\alpha}\), and \(p(m, m') = [\alpha, \beta]\) for \(p(m, m') = \delta_{m,\alpha}\delta_{m',\beta}\), the vertex \(\nu_1\) is:

\[
\begin{align*}
\xi_R(m_{R_{1,1}}, m_{R_{1,2}} | \nu_1) &= \frac{1}{7}[1, +1, -1] + \frac{1}{7}[-1, +1], \\
\xi_R(m_{R_{2,1}}, m_{R_{2,2}} | \nu_1) &= \frac{1}{7}[1, +1, +1] + \frac{1}{7}[-1, -1], \\
\xi_R(m_{R_{1,1}}, m_{R_{2,1}} | \nu_1) &= [+1, +1], \\
\xi_C(m_{C_{1,1}}, m_{C_{1,2}} | \nu_1) &= \frac{1}{7}[1, +1, +1] + \frac{1}{7}[-1, -1], \\
\xi_C(m_{C_{2,1}}, m_{C_{1,2}} | \nu_1) &= \frac{1}{7}[1, +1, +1] + \frac{1}{7}[-1, -1], \\
\xi_C(m_{C_{1,1}}, m_{C_{2,2}} | \nu_1) &= [-1, +1].
\end{align*}
\]

Using equations \((65a)–(65c)\) and analogues thereof, one can compute from these the response functions for each of the nine equivalence classes of binary-outcome measurements. They are
It is easy to verify that the two ways of defining the value of the response function for $\mathcal{M}_k$ at $\kappa_i$ (via simulation by $M_R$ or via simulation by $M_C$) yield the same result, so that this is indeed a noncontextual assignment satisfying the compatibility relations. In terms of expectation values, this assignment corresponds to

$$
\begin{pmatrix}
    (m_1)_{\kappa_1} & (m_2)_{\kappa_1} & (m_3)_{\kappa_1} \\
    (m_1)_{\kappa_2} & (m_2)_{\kappa_2} & (m_3)_{\kappa_2} \\
    (m_1)_{\kappa_3} & (m_2)_{\kappa_3} & (m_3)_{\kappa_3}
\end{pmatrix} = \begin{pmatrix}
    0 & 0 & -1 \\
    0 & 0 & +1 \\
    +1 & +1 & +1
\end{pmatrix}.
$$

Note that it makes four of the nine measurements outcome-indeterministic.

A second concrete example of a vertex of the polytope, denoted $\kappa_{2s}$ is

$$
\begin{pmatrix}
    (m_1)_{\kappa_1} & (m_2)_{\kappa_1} & (m_3)_{\kappa_1} \\
    (m_1)_{\kappa_2} & (m_2)_{\kappa_2} & (m_3)_{\kappa_2} \\
    (m_1)_{\kappa_3} & (m_2)_{\kappa_3} & (m_3)_{\kappa_3}
\end{pmatrix} = \begin{pmatrix}
    0 & 0 & +1 \\
    +1 & 0 & 0 \\
    0 & -1 & 0
\end{pmatrix},
$$

where six of the nine measurements are outcome-indeterministic.

By considering the set of all deterministic processings of the measurements that preserve the compatibility relations holding among these, defined in appendix B, one can determine the symmetries of the noncontextual measurement-assignment polytope. Specifically, each such deterministic processing induces a bijective mapping of the set of vertices to itself. The full symmetry group is specified in appendix B. It is straightforward to verify that it can be generated by the following three deterministic processings:

$$
\begin{align*}
\langle m_1 \rangle_{\kappa_1} & \leftrightarrow \langle m_2 \rangle_{\kappa_1}, \\
\langle m_2 \rangle_{\kappa_1} & \leftrightarrow \langle m_2 \rangle_{\kappa_2}, \quad \text{and} \\
\langle m_3 \rangle_{\kappa_1} & \leftrightarrow \langle m_3 \rangle_{\kappa_3}, \quad \text{and} \\
\langle m_3 \rangle_{\kappa_2} & \leftrightarrow \langle m_3 \rangle_{\kappa_2}.
\end{align*}
$$

Note that the number of measurements that are assigned outcomes deterministically is preserved by these symmetry operations. Consequently, our two examples above, equations (78) and (79), are in different symmetry classes. In fact, we find that there are only these two symmetry classes.

The symmetry class wherein six of the nine measurements are indeterministic contains 48 vertices. As $3 \times 3$ matrices, they correspond to those with elements in $\{-1, 0, +1\}$ having the property that every row and every column contains precisely one non-zero element. The symmetry class wherein four of the nine measurements are indeterministic contains 72 vertices, and corresponds to those $3 \times 3$ matrices with elements in $\{-1, 0, +1\}$ having a single row of nonzero elements and a single column of nonzero elements such that the overall parity of the row is $+1$, and the overall parity of the column is $\eta$, where $\eta = -1$ if it is the third column and $\eta = +1$ otherwise.

4.5. Vertices of the noncontextual correlation polytope

To determine the vertices of the noncontextual correlation polytope, from the vertices of the noncontextual measurement-assignment polytope and those of the noncontextual source-assignment polytope, we preserve only those pairings which lead to extremal 9-tuples, as noted above equation (64). Specifically, for each of the $120^2 = 14400$ pairings $(\kappa_i, \kappa_i')$, one computes the 9-tuple $(\langle m_1 \rangle_{\kappa_1}, \langle s_1 \rangle_{\kappa_1'}, \ldots, \langle m_3 \rangle_{\kappa_3}, \langle s_3 \rangle_{\kappa_3'})$. By eliminating duplicate and non-extremal points from this set, we obtain the vertices of the noncontextual correlation polytope.

A concrete example of a vertex of the noncontextual correlation polytope is obtained by pairing the vertex $\kappa_3$ of the noncontextual measurement-assignment polytope, described in equation (78), with a vertex $\kappa_3'$ of the noncontextual source-assignment polytope having precisely the same components as $\kappa_3$. This pairing yields

$$
\begin{pmatrix}
    (m_1)_{\kappa_1} & (s_1)_{\kappa_1'} & (m_2)_{\kappa_1} & (s_2)_{\kappa_1'} & (m_3)_{\kappa_1} & (s_3)_{\kappa_1'} \\
    (m_1)_{\kappa_2} & (s_2)_{\kappa_1'} & (m_2)_{\kappa_2} & (s_2)_{\kappa_2'} & (m_3)_{\kappa_2} & (s_3)_{\kappa_2'} \\
    (m_1)_{\kappa_3} & (s_3)_{\kappa_1'} & (m_2)_{\kappa_3} & (s_3)_{\kappa_2'} & (m_3)_{\kappa_3} & (s_3)_{\kappa_3'}
\end{pmatrix} = \begin{pmatrix}
    0 & 0 & +1 \\
    0 & 0 & +1 \\
    +1 & +1 & +1
\end{pmatrix}.
$$

Note that this vertex can also be constructed by pairing $\kappa_5$ with a corresponding $\kappa_5'$, where $\kappa_5$ is defined as $\kappa_1$ per equation (78) but with the first two rows permuted, i.e. such that the $-1$ appears in the second row instead of the...
first. The distinct pairings \((\kappa_1, \kappa'_1)\) and \((\kappa_3, \kappa'_3)\) therefore yield duplicate noncontextual correlation points under entry-wise product.

Another vertex of the noncontextual correlation polytope is obtained by pairing the vertex \(\kappa_2\) of the noncontextual measurement-assignment polytope, described in equation (79), with a vertex \(\kappa'_2\) of the noncontextual source-assignment polytope having precisely the same components as \(\kappa_2\):

\[
\begin{bmatrix}
  \langle m_1 \rangle, \langle s_1 \rangle, \langle s_1 \rangle' \\
  \langle m_2 \rangle, \langle s_2 \rangle, \langle s_2 \rangle' \\
  \langle m_3 \rangle, \langle s_3 \rangle, \langle s_3 \rangle'
\end{bmatrix}
\]

By contrast, if we pair \(\kappa_0\) with \(\kappa'_0\), we obtain a point that is nonextremal in the noncontextual correlation polytope. We find that the noncontextual correlation polytope has 120 vertices\(^{13}\).

By considering the deterministic processings of the measurements and sources which preserve the operational Peres–Mermin scenario—that is, the processings which preserve the compatibility relations among the measurements, the compatibility relations among the sources, and the manner in which the sources and the measurements are paired—we can infer the symmetries of the noncontextual correlation polytope, as discussed in appendix B. Specifically, every such symmetry bijectively maps the set of vertices of this polytope to itself. The full symmetry group is described in appendix B, and it is straightforward to verify that it can be generated by the following three processings (which we also describe as operations on the \(3 \times 3\) matrix):

\[
\begin{bmatrix}
  \alpha_{11} & \alpha_{12} & \alpha_{13} \\
  \alpha_{21} & \alpha_{22} & \alpha_{23} \\
  \alpha_{31} & \alpha_{32} & \alpha_{33}
\end{bmatrix}
\]

Columns 1→2

Rows 2→3

Modified transpose

Note that the modified transpose operation appearing in equation (86) is distinct from the modified transpose appearing in equation (80).

4.6. Facets of the noncontextual correlation polytope: noncontextuality inequalities

To determine the facet inequalities for the noncontextual correlation polytope from its vertices, we proceed (again) by solving the convex hull problem \([25–28]\), discussed more fully in appendix A.

Facet inequalities of the noncontextual correlation polytope have the form

\[
\sum_{i,j=1}^{3} \omega_{ij} \alpha_{ij} \leq \beta,
\]

where \([\omega_{ij}]\) and \(\beta\) are integers. Arranging the \(\alpha_{ij}\) and the \(\omega_{ij}\) into \(3 \times 3\) matrices and denoting the entry-wise matrix product by \(\odot\) and the sum of the elements of a matrix \(A\) by \(\text{su}(A)\), we can express this as

\[
\text{su}\left[
\begin{bmatrix}
  \alpha_{11} & \alpha_{12} & \alpha_{13} \\
  \alpha_{21} & \alpha_{22} & \alpha_{23} \\
  \alpha_{31} & \alpha_{32} & \alpha_{33}
\end{bmatrix}
\odot
\begin{bmatrix}
  \omega_{11} & \omega_{12} & \omega_{13} \\
  \omega_{21} & \omega_{22} & \omega_{23} \\
  \omega_{31} & \omega_{32} & \omega_{33}
\end{bmatrix}
\right] \leq \beta.
\]

We will refer to the matrix of \(\alpha_{ij}\)'s as the coefficient matrix for the inequality.

We find that there are 184 inequalities, all of which are expressed using coefficient matrices where \(\forall i, j : \alpha_{ij} \in \{0, -1, +1\}\) and \(\beta \in \{1, 3, 5\}\).

Note that the deterministic processings of the experiment that bijectively map the set of vertices of the noncontextual correlation polytope to itself also bijectively map the set of facet inequalities of the noncontextual correlation polytope to itself, and vice versa. Consequently, the symmetry group of the set of facet inequalities is the same as the symmetry group of the set of vertices. The value of \(\beta\) in a facet inequality is invariant under the symmetry group, so that only the matrix of \(\alpha\) coefficients transforms nontrivially. Given that the facet inequalities can be expressed in the form of equation (85), any map on the \(\omega\) matrix can be transferred onto the matrix of \(\alpha\) coefficients. Consequently, the action of the symmetry group on the coefficient matrix is precisely

\(^{13}\)To be clear, the 120 vertices of the noncontextual correlation polytope should not be confused with the 120 vertices of the noncontextual measurement-assignment polytope; they are distinct sets, subject to different symmetry classifications.
parallel to that described in equation (83), namely, the group generated by

\[
\begin{align*}
\alpha_{11} &\leftrightarrow \alpha_{12} & &\alpha_{11} &\leftrightarrow -\alpha_{11} \\
\alpha_{21} &\leftrightarrow \alpha_{22} & &\alpha_{12} &\leftrightarrow \alpha_{21} \\
\alpha_{31} &\leftrightarrow \alpha_{32} & &\alpha_{22} &\leftrightarrow \alpha_{32} \\
\alpha_{33} &\leftrightarrow \alpha_{33} & &\alpha_{33} &\leftrightarrow -\alpha_{33} \\
\end{align*}
\]

(86)

Thus, if the \(\alpha\) coefficient matrix of one inequality is related to that of another inequality by one of the symmetry operations we have identified, then these two inequalities are in the same symmetry class. An efficient description of all of the facet inequalities is achieved by describing representatives of each of the symmetry classes, and closing under the action of the symmetries. We find that there are just three symmetry classes of facet inequalities, conveniently distinguished by their values of \(\beta\).

The first symmetry class is trivial in the sense that the facet inequalities therein hold for all correlations that are logically possible in the operational Peres–Mermin scenario, and consequently they are not sensitive to whether or not the correlations admit of a noncontextual model.

**Trivial Class**: a representative of this class is

\[
\text{su}\begin{pmatrix}
-1 & +1 & +1 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix} \circ \begin{pmatrix}
\omega_{11} & \omega_{12} & \omega_{13} \\
\omega_{21} & \omega_{22} & \omega_{23} \\
\omega_{31} & \omega_{32} & \omega_{33}
\end{pmatrix} \leq 1.
\]

(87)

Closing under the symmetries, one finds that there are 24 such inequalities, corresponding to coefficient matrices where only one row or column has all nonzero elements, and such that the overall parity of these is \(-1\). We justify the claim that these inequalities are trivial in appendix C.

The term **noncontextuality inequality** is reserved for those facet inequalities of the noncontextual correlation polytope that are nontrivial. There are two symmetry classes of these.

**Nontrivial Class I**: a representative of this class is

\[
\text{su}\begin{pmatrix}
-1 & 0 & +1 \\
+1 & 0 & +1 \\
0 & -1 & 0
\end{pmatrix} \circ \begin{pmatrix}
\omega_{11} & \omega_{12} & \omega_{13} \\
\omega_{21} & \omega_{22} & \omega_{23} \\
\omega_{31} & \omega_{32} & \omega_{33}
\end{pmatrix} \leq 3.
\]

(88)

Closing under the symmetries, one finds that there are 144 such inequalities, all of which can be constructed as follows: choose a special position in the matrix of coefficients, say element \(\alpha_{ij}\), and make it \(+1\) or \(-1\). Let all other elements in the same row or column of the coefficients matrix be zero. Finally, choose any assignment of \(\pm 1\) for the remaining four elements such that the overall parity of the five nonzero elements is \(+1\). The example inequality of equation (88) is one of the eight inequalities that follow by starting with \(\alpha_{32} = -1\) as the special element.

An example of an inequality from this class that is maximally violated by the noiseless quantum realization of the operational Peres–Mermin scenario (described in section 4.1.1) is

\[
\text{su}\begin{pmatrix}
+1 & +1 & 0 \\
+1 & +1 & 0 \\
0 & 0 & +1
\end{pmatrix} \circ \begin{pmatrix}
\omega_{11} & \omega_{12} & \omega_{13} \\
\omega_{21} & \omega_{22} & \omega_{23} \\
\omega_{31} & \omega_{32} & \omega_{33}
\end{pmatrix} \leq 3,
\]

(89)

as we will demonstrate in the next section.

**Nontrivial Class II**: a representative of this class is

\[
\text{su}\begin{pmatrix}
+1 & -1 & -1 \\
+1 & +1 & +1 \\
+1 & -1 & -1
\end{pmatrix} \circ \begin{pmatrix}
\omega_{11} & \omega_{12} & \omega_{13} \\
\omega_{21} & \omega_{22} & \omega_{23} \\
\omega_{31} & \omega_{32} & \omega_{33}
\end{pmatrix} \leq 5.
\]

(90)

Closing under the symmetries, one finds that there are 16 inequalities in this class, all of which have only nonzero elements in the coefficient matrix. The 16 inequalities are precisely those whose coefficient matrices have \(+1\) overall parity for every row and every column.

An example of an inequality from this class that is maximally violated by the noiseless quantum realization of the operational Peres–Mermin scenario (described in section 4.1.1) is

\[
\text{su}\begin{pmatrix}
+1 & +1 & +1 \\
+1 & +1 & +1 \\
+1 & +1 & +1
\end{pmatrix} \circ \begin{pmatrix}
\omega_{11} & \omega_{12} & \omega_{13} \\
\omega_{21} & \omega_{22} & \omega_{23} \\
\omega_{31} & \omega_{32} & \omega_{33}
\end{pmatrix} \leq 5,
\]

(91)

as we will demonstrate in the next section.
4.6.1. Quantum violation of the inequalities

We have already seen in section 3.2 that one can identify sharp quantum sources and sharp quantum measurements that satisfy the Peres–Mermin compatibility structure and whose statistics are inconsistent with a universally noncontextual model. These quantum sources and measurements must therefore violate our inequalities. Indeed, for these quantum sources and measurements, it follows from equation (35) that \( \forall i, j : \omega_{ij} = 1 \), that is, every one of the nine source-measurement pairs exhibits perfect correlation. This implies that the left-hand side of the noncontextuality inequality of equation (89) evaluates to 5 for this quantum realization, thereby exceeding the noncontextual bound of 3. It also implies that the left-hand side of the noncontextuality inequality of equation (89) evaluates to 9, which exceeds the noncontextual bound of 5.

4.6.2. Robustness of the inequalities to noise

We now demonstrate explicitly how the noncontextuality inequalities we have derived are robust to noise by showing that the noisy quantum realization of the operational Peres–Mermin scenario, described in section 4.1.2, can still lead to a violation. Recall that this consisted of quantum sources and quantum measurements that were the image under a partial depolarization map of those appearing in the no-go theorem of section 3.2.

For these noisy sources and measurements, the value of the correlation for each of the nine source-measurement pairs was computed, as a function of the weight \( r \) of the identity map in the partial depolarization, in equation (51). Translating into the \( \omega_{ij} \) notation of equation (60), the result is

\[
\forall i, j: \omega_{ij} = r^2.
\]

Substituting this expression into the noncontextuality inequality of equation (89), we obtain

\[
5r^2 \leqslant 3,
\]

Consequently, as long as the level of noise is such that \( r > \frac{\sqrt{5}}{\sqrt{3}} \approx 0.77460 \), one has a violation of the noncontextuality inequality. (Lower values of \( r \) correspond to stronger noise, so this is an upper bound on the noise.)

Similarly, from the noncontextuality inequality of equation (91), we obtain

\[
9r^2 \leqslant 5,
\]

implying that we require \( r > \frac{\sqrt{5}}{\sqrt{9}} \approx 0.73536 \) to see a violation.

Because \( \frac{\sqrt{5}}{\sqrt{9}} < \frac{\sqrt{5}}{\sqrt{3}} \), the noncontextuality inequality of equation (91) can tolerate more noise than that of equation (89).

We have shown that our noncontextuality inequalities still admit a violation even in the presence of significant depolarizing noise relative to the ideal quantum realization. As such, the fact that any attempt at an experimental implementation of the ideal quantum sources and measurements inevitably only realizes noisy versions of these is not an obstacle to demonstrating an experimental failure of noncontextuality.

5. How to implement an experimental test of these inequalities

The assumption of noncontextuality only has nontrivial consequences if one has experimentally verified that certain operational equivalence relations hold among the measurements and among the sources. This creates a problem for experimental tests of noncontextuality because the definition of operational equivalence for sources is in terms of equivalence of statistics for all measurements, and for measurements it is in terms of equivalence of statistics for all sources, and, strictly speaking, one can never experimentally implement all possible procedures of either type. The problem may be summarized as the physical impossibility of verifying any criterion that involves a universal quantifier. We here explain our view on what is the appropriate attitude to take towards this problem.

A tomographically-complete set of measurements is defined as a set of measurements whose statistics are sufficient to infer the statistics of any other measurement. A tomographically-complete set of sources is defined as a set of sources whose statistics are sufficient to infer the statistics of any other source. It follows that to judge operational equivalence relations among sources, it is sufficient to consider their statistics on a tomographically-complete set of measurements rather than on all measurements, and to judge operational equivalence relations among measurements, it is sufficient to consider their statistics on a tomographically-complete set of sources rather than on all sources. The problem therefore reduces to identifying tomographically complete sets for each.
It is well known that in quantum theory, the set of observables obtained by taking all products of Pauli operators corresponds to a tomographically complete set of measurements for a pair of qubits. Therefore, quantum theory dictates that one must supplement the products of Pauli operators appearing in the Peres–Mermin square with all the other nontrivial products of Pauli operators, that is, with \({ \mathbb{Y} \otimes \mathbb{X}, \mathbb{Y} \otimes \mathbb{Z}, \mathbb{Y} \otimes \mathbb{I}, \mathbb{I} \otimes \mathbb{Y}, \mathbb{X} \otimes \mathbb{Y}, \mathbb{Z} \otimes \mathbb{Y} }\), in order to obtain a tomographically complete set. Consequently, by the lights of quantum theory, in order to test operational equivalence relations among the sources, it is necessary to do more measurements than appear in the Peres–Mermin square construction.

Furthermore, if one seeks to implement a direct experimental test of noncontextuality, then one does not want to presume the correctness of quantum theory. As such, it is inappropriate to presume that the fifteen binary-outcome measurements that one expects to be tomographically complete by the lights of quantum theory are in fact tomographically complete. Instead, one should accumulate experimental evidence for this hypothesis by implementing the greatest diversity of measurements on the system that one can and by verifying that the statistics for each such measurement can be inferred from the statistics of the hypothetical tomographically complete set.

Similar comments apply for the problem of identifying a tomographically complete set of sources for the purpose of evaluating operational equivalence relations among the measurements.

We refer the reader to [29] for more details on how to acquire experimental evidence for a given set of measurements (preparations) being tomographically complete.

Even given good evidence of tomographic completeness, one faces another problem with experimentally verifying operational equivalences, namely, that if one aims to implement a particular set of procedures (termed the target procedures), the unavoidable imprecision of experimental implementations implies that one inevitably fails to do so precisely, and one’s experiment instead realizes a set of procedures that deviate from the target procedures, termed the primary procedures. Given the failure of the primary procedures to coincide with the target procedures, operational equivalence relations that hold among the target procedures need not hold among the primary procedures. The primary procedures therefore do not generally realize the operational equivalence relations that are the starting point for derivations of noncontextuality inequalities. This has been termed the problem of ‘no strict operational equivalences’ in [13].

Before describing how to resolve it, we specify how it arises in the context of the Peres–Mermin construction.

There are eighteen target measurement procedures, corresponding to six compatible triples of binary-outcome measurements, one triple for each row and column of the square. Recall that we denoted the 4-outcome measurement that simulates all of the binary-outcome measurements in the first row by \(M_{R_{1}}\) and its outcome by the pair of binary-outcome variables \((m_{R_{1,1}}, m_{R_{1,2}})\). Recall also that we defined \(M_{C_{1}}\) to be the binary-outcome measurement procedure one obtains by implementing \(M_{R_{1}}\) and outputting the single binary variable \(m_{R_{1}}\) (i.e., by marginalizing over \(m_{R_{1,2}}\)). Similarly, we defined \(M_{C_{1}}\) to be the binary-outcome measurement procedure that one obtains by implementing \(M_{C_{1}}\) and outputting the single binary variable \(m_{C_{1}}\) (i.e., by marginalizing over \(m_{C_{1,2}}\)). \(M_{R_{1}}\) and \(M_{C_{1}}\) are two of the target procedures. Although they are distinct in the sense of involving different physical operations in the laboratory, they are required to be operationally equivalent in order for the assumption of noncontextuality to have any nontrivial consequences. Similar comments hold for the pair of procedures associated to each of the other points of the Peres–Mermin square. Consequently, if instead of considering the hypergraph where the nodes are operational equivalence classes of measurement procedures and the hyperedges are compatibility relations as we did in figure 1, we consider the hypergraph where the nodes are individual measurements, and there are two types of hyperedges, one denoting compatibility relations and the other denoting operational equivalence relations, then we can represent the set of target measurement procedures by figure 3.

Note that all of the relationships of compatibility that hold among the measurements are guaranteed by the manner in which those measurements are implemented. Specifically, for every set of compatible measurements in the experiment, these are implemented by various coarse-grainings of a single measurement. As such, the compatibility relations are ensured by construction, and no further evidence must be accumulated to confirm their presence. Only the operational equivalence relations among these different coarse-grained measurements must be tested explicitly.

This distinguishes our approach from other approaches to experimental tests of noncontextuality [23] wherein the compatibility relations must be tested explicitly. In the latter approaches, the experiment seeks to implement a single measurement procedure \(M\) before or after another measurement \(M'\) which is drawn from a set of possibilities that are compatible with \(M\) but incompatible with one another. It is then critical to demonstrate that the procedures \(M\) and \(M'\) that are actually realized in the experiment are indeed compatible.
Similar comments apply to sources. In a hypergraph where the nodes are individual sources, and there are two types of hyperedges, one denoting compatibility relations and the other denoting operational equivalence relations, we represent the set of target sources by figure 4. If every set of compatible sources is implemented by a coarse-graining of a single source, then the compatibility relations among the sources are achieved by construction and only the operational equivalence relations need to be tested.

We are now in a position to describe the problem of no strict operational equivalences in the case of Peres–Mermin. Given the unavoidable deviation of the actually-realized procedures (the primary procedures) from the target procedures, one expects that for each $ij$, the actually-implemented versions of $M_{ij}^R$ and of $M_{ij}^C$ will not be strictly operationally equivalent, nor will the actually-implemented versions of $S_{ij}^R$ and of $S_{ij}^C$.

The resolution to this problem was provided in [13] and proceeds as follows. When one’s experiment realizes a finite set of primary procedures, it simultaneously provides a characterization of an infinite set of procedures, namely, those that can be obtained by classical post-processing of the set of primary procedures, for instance, any procedure in the convex hull of these. One can then choose a set of secondary procedures from among this infinite set under the constraint of exactly satisfying the desired operational equivalence relations.

Specifically, we select secondary versions of the 4-outcome measurements and 4-outcome sources, under the constraint that the binary-outcome measurements and sources that they define precisely satisfy the operational equivalence relations depicted in the hypergraphs of figures 3 and 4 respectively.
Once this is done, one uses the secondary versions of the nine binary-outcome measurements and nine binary-outcome sources to compute the correlations $\{w_{ij}\}$. Because the operational equivalences hold for these secondary measurements and sources, the assumption of noncontextuality implies a constraint on their ontological representation, and consequently the operational correlations exhibited by these secondary measurements and sources are expected to satisfy the noncontextuality inequalities if the experiment admits of a noncontextual model. A violation of these inequalities by the secondary procedures, therefore, witnesses a failure of noncontextuality.

6. Conclusions

In this paper, we have derived a set of noncontextuality inequalities based on the Peres–Mermin square proof of the impossibility of a KS-noncontextuality in quantum theory. These inequalities are robust to noise and consequently they can be tested directly by experiment. If they are found to be violated by experiment, then not only quantum theory, but any physical theory that can do justice to the experimental data must fail to admit of a noncontextual model. The procedure we have outlined for deriving such inequalities is quite general and can be applied to other state-independent proofs of the Kochen–Specker theorem, particularly those that are expressed algebraically.

In appendix D, we contrast our approach with previous attempts at deriving operationally-meaningful noncontextuality inequalities based on the Peres–Mermin square and we criticize the latter.

The technique we have described here can be applied to deriving inequalities for correlations in the full set of 81 source-measurement pairings, or even to deriving inequalities on these together with the nine marginal expectation values for the measurements alone and the nine marginal expectation values for the sources alone. What prevented us from doing so here was the computational infeasibility of solving the associated convex hull problems, using the best algorithms for this problem that we could identify and given the computational resources we devoted to the task. It is conceivable, however, that by leveraging our knowledge of the symmetries of the polytopes involved in the problem, one might render it computationally feasible using the same algorithms and computational resources. We are also hopeful that the graph-theoretic techniques for analysing contextuality scenarios, described in [30–32], might suggest better algorithms for deriving these inequalities.

It is worth noting that the technique for deriving noncontextuality inequalities we have introduced here, insofar as it reduces to a convex hull problem, is an instance of the problem of quantifier elimination. Recent work in quantum foundations has seen increasing use of quantifier elimination algorithms. Fourier–Motzkin elimination, which is appropriate for problems wherein the dependence on the variables to be eliminated is linear, has been used to derive Bell inequalities [33], and also recently, to derive Bell-like inequalities for novel causal scenarios [34–36]. In [36], where the problem is reduced to what is known as the classical marginals problem—that of determining whether a given set of distributions on various subsets of a set of variables can arise as the marginals of a single joint distribution over all of the variables—and this problem can be solved by performing quantifier elimination on the probabilities in the joint distributions using convex hull algorithms. Nonlinear quantifier elimination using computational algebraic geometry has also found application in deriving Bell-like inequalities in simple scenarios [37]. We anticipate that these more general techniques for quantifier elimination will ultimately also find applications to the derivation of noncontextuality inequalities.

Acknowledgments

The authors thank D Schmid for helpful comments. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

Appendix A. Polytope representation conversion

The problem of converting between polytope representations is essential to our technique for deriving noncontextuality inequalities. We describe the abstract mathematical problem here, and refer the reader to various software packages which are capable of solving it efficiently. A polytope may be uniquely characterized in terms of two dual descriptions: (1) as a collection of facet-defining inequalities, which is called the halfspace
representation of the polytope, denoted \( H\)-rep, or (2) as a collection of vertices, which is called the vertex representation of the polytope, denoted \( V\)-rep. It turns out that facet enumeration from vertices and vertex enumeration from facets are exactly the same computational problem. To understand the representation conversion problem in general, it is imperative to first understand the computational data structures used to represent a collection of points or a collection of inequalities pertaining to \( d\)-dimensional vector spaces.

We elect to represent the inequality

\[
0 \leq c_0 + \sum_{i=1}^{d} c_i x_i
\]

by the \((d + 1)\)-dimensional vector

\[
h := \left\{ \frac{c_0}{g}, \frac{c_1}{g}, \ldots, \frac{c_d}{g} \right\},
\]

where \(g\) is chosen to efficiently store \( h\) in computer memory. For rational coefficients, taking \( g = \text{GCD} [c_0, \ldots, c_d] \) allows \( h\) to be stored as a vector of integers\(^{14}\). Indeed, for the sort of contextuality problems that involve linear rational compatibility relationships (such as the operational Peres–Mermin scenario), we can always represent the pertinent inequalities by vectors of integers.

We elect to represent the point

\[
x := \{x_0, \ldots, x_d\}
\]

by the \((d + 1)\)-dimensional vector

\[
v := \left\{ \frac{1}{g'}, \frac{x_1}{g'}, \ldots, \frac{x_d}{g'} \right\},
\]

where \(g'\) is chosen to efficiently store \( v\) in computer memory. For rational components, taking \( g' = \text{GCD} [x_0, \ldots, x_d] \) allows \( v\) to be stored as a vector of integers.

If a polytope has \( n\) facet inequalities, then given that each of these can be represented by a \((d + 1)\)-dimensional vector, its halfspace representation can be given as an \( n \times (d + 1)\) matrix

\[
H := \begin{pmatrix} h^{(1)} \\ \vdots \\ h^{(n)} \end{pmatrix}.
\]

We refer to \( H = \{h^{(i)}\} \), as the \( H\)-rep matrix.

Similarly, if a polytope has \( m\) vertices, then given that each of these can be represented by a \((d + 1)\)-dimensional vector, its vertex representation can be given as an \( m \times (d + 1)\) matrix

\[
V := \begin{pmatrix} v^{(1)} \\ \vdots \\ v^{(m)} \end{pmatrix}.
\]

We refer to \( V = \{v^{(i)}\} \), as the \( V\)-rep matrix.

It follows that a facet enumeration algorithm takes a \( V\)-rep matrix as input and returns an \( H\)-rep matrix as output, while a vertex enumeration algorithm takes an \( H\)-rep matrix as input and returns a \( V\)-rep matrix as output. We now show that the algorithms are equivalent.

A set of point-vectors \( V = \{v^{(i)}\} \) defines the \( V\)-rep of a polytope relative to some \( H\)-rep inequalities-matrix \( H = \{h^{(i)}\} \), if and only if the following conditions hold:

**Consistent:** every element of \( \{v^{(i)}\} \) must satisfy all of the facet inequalities,

\[
\forall i, j : h^{(i)} \cdot v^{(j)} \geq 0.
\]

**Complete:** every point \( x \) which cannot be expressed as a positive combination of the elements of \( \{v^{(j)}\} \), must violate some facet inequality,

\[
\forall x \notin \text{PosLinSpan}\{\{v^{(j)}\}\}, \exists j : h^{(j)} \cdot x < 0.
\]

---

\(^{14}\) GCD stands for Greatest Common Divisor; here we are applying GCD to the field of rational numbers, see functions.wolfram.com/IntegerFunctions/GCD and math.stackexchange.com/q/131081.
Concise: no element of \( \{v^{(j)}\} \) is a positive combination of the others,

\[
\forall j : v^{(j)} \not\in \text{PosLinSpan}[\{v^{(i)}\}_{k \neq j}].
\]

A set of inequality-vectors \( H = \{h^{(i)}\} \) defines the \( H \)-rep of a polytope relative to some \( V \)-rep vertices-matrix \( V = \{v^{(j)}\} \) if and only if the following conditions hold:

**Consistent:** every inequality in the set \( \{h^{(i)}\} \) is satisfied by all of the vertices of the polytope,

\[
\forall i, j : h^{(i)} \cdot v^{(j)} \geq 0.
\]

**Complete:** any inequality \( e \) which cannot be expressed as a positive combination of the inequalities \( \{h^{(i)}\} \), must be violated by at least one vertex of the polytope,

\[
\forall e \not\in \text{PosLinSpan}[\{h^{(i)}\}], \exists j : e \cdot v^{(j)} < 0.
\]

Concise: no inequality in \( \{h^{(i)}\} \) is a positive combination of the others,

\[
\forall i : h^{(i)} \not\in \text{PosLinSpan}[\{h^{(j)}\}_{k \neq i}].
\]

It should now be clear that the conditions which define the dual representation are symmetric with respect to the interchange of \( H \) and \( V \). Consequently, the polytope representation conversion task is a single computational problem, oblivious to the subjective interpretation of the task as facet enumeration or vertex enumeration. The common underlying computational task is conventionally referred to as the *convex hull* problem.

In the general procedure for deriving robust operational noncontextuality inequalities, one encounters the convex hull problem three times:

1. When one enumerates the vertices of the noncontextual measurement-assignment polytope starting from its facet inequalities (71a) and (71b), one is performing \( H \)-rep to \( V \)-rep conversion.
2. When one enumerates the vertices of the noncontextual source-assignment polytope starting from its facet inequalities (74a) and (74b), one is performing \( H \)-rep to \( V \)-rep conversion.
3. When one enumerates the facets of the noncontextual correlation polytope starting from its vertices, one is performing \( V \)-rep to \( H \)-rep conversion.

For the problem of interest in this article, the Peres–Mermin scenario, the noncontextual measurement-assignment polytope coincides with the noncontextual source-assignment polytope, so the first two instances of the convex hull problem are the same, converting 24 facet inequalities into 120 vertices. The third instance concerns the noncontextual correlation polytope, and converts the 120 vertices thereof into 184 facet inequalities.

Solving the convex hull problem can, in practice, be accomplished by a variety of algorithms; the most prominent implementation is the dual description method. For details on the various algorithms used to solve the convex hull problem see [25–28]. Efficient convex hull solvers are abundant in current software packages\(^{15}\).

---

\(^{15}\) **PORTA**: [Executable] older software, limited to moderate dimension polytopes \((d \leq 100)\), nevertheless **PORTA** is quite fast. **CDD, LRS** and **SymPol**: [Executionables and C-libraries] all accept the same style of input file. The **CDD** binary is no longer being updated, but it is quite fast. **LRS** uses a different internal algorithm than **CDD**; sometimes it is much faster, sometimes much slower. **SymPol** can automatically discover symmetries to reduce the size of the computation, and is the authors’ preferred tool for solving large convex hull problems in the presence of high symmetry, such as is very often the case in noncontextuality polytopes. **Qhull**: [Multiple interfaces] widely used, but suboptimal for high dimensional polytopes. **Skeleton** and **Qskeleton**: [Executionables] the fastest convex hull solvers in the authors’ informal benchmarking; both accept the same style of input file. **polytop2**: [Python package] polyhedral geometry software with good convex hull performance if used with GLPK bindings. **polymake**: [Software for Linux and Mac] versatile polyhedral geometry software, capable of much more than just convex hull, **polymake** offers selecting amongst multiple internal convex hull algorithms. **PANDA**: [Executable and C-library] given a description of known symmetries, **PANDA** is capable of converting polytope representations extremely efficiently; the symmetries are specified relative to the polytope’s coordinate system.
Note that when we infer the vertices of the noncontextual correlation polytope from the superset of all possible pairings of vertices from the noncontextual measurement-assignment polytope and the noncontextual source-assignment polytope, the computational problem is one of redundancy removal. Redundancy removal is the task of isolating the vertices from a set of non-extremal points, or, computationally equivalently, isolating the irredundant linear constraints from some inconcise collection of inequalities. Standard algorithms have been developed for redundancy removal \cite{10.1007/978-3-642-36038-1_13,40}, and a variety of current software tools are also available\textsuperscript{16}.

### Appendix B. Deterministic processings of the experimental procedures that preserve the compatibility relations of the Peres–Mermin scenario

When we express our noncontextuality inequalities in section 4.6, the description of the full set of such inequalities is simplified by appealing to the symmetries of the operational construction. These symmetries correspond to processings of the experiment that permute source-measurement pairs (i.e., deterministically process the setting variables in a correlated fashion), and that deterministically process the outcomes of the measurements (and independently) the outcomes of the sources. It is therefore useful to explicitly characterize these symmetries.

Suppose one has an experiment satisfying the operational features of the Peres–Mermin scenario. This means that there are nine equivalence classes of binary-outcome measurements, \( \{ M_{11}, \ldots, M_{13} \} \), satisfying the compatibility hypergraph of figure 1 and the compatibility relations defined around equations (42a)–(42c). Similarly, there are nine equivalence classes of binary-outcome sources, \( \{ S_{11}, \ldots, S_{33} \} \), satisfying the compatibility hypergraph of figure 2 and the compatibility relations defined around equations (45a)–(45c).

Now suppose that from these one defined a new pair of sets, \( \{ M'_{11}, \ldots, M'_{13} \} \) and \( \{ S'_{11}, \ldots, S'_{33} \} \), as processings of \( \{ M_{11}, \ldots, M_{13} \} \) and \( \{ S_{11}, \ldots, S_{33} \} \) as follows. To implement the source-measurement pair \( (S'_{ij}, M'_{ij}) \), one begins by implementing the source-measurement pair \( (S_{ij}, M_{ij}) \), where \( \pi \) is some permutation of the nine positions in the Peres–Mermin square, and then for each \( M_{ij} \), one multiplies its outcome by \( C_{i} \in \{ -1, +1 \} \) (thereby either flipping its value or leaving it the same) and for each \( S_{ij} \), one multiplies its outcome by \( \gamma_{i} \in \{ -1, +1 \} \).

The deterministic processings by which the outcomes of the original measurements and sources are mapped to those of the primed measurements and sources can be expressed as follows. Denoting component-wise product of matrices by \( \odot \),

\[
\begin{pmatrix}
    m'_{11} & m'_{12} & m'_{13} \\
    m'_{21} & m'_{22} & m'_{23} \\
    m'_{31} & m'_{32} & m'_{33}
\end{pmatrix} =
\begin{pmatrix}
    \zeta_{11} & \zeta_{12} & \zeta_{13} \\
    \zeta_{21} & \zeta_{22} & \zeta_{23} \\
    \zeta_{31} & \zeta_{32} & \zeta_{33}
\end{pmatrix} \odot
\begin{pmatrix}
    m_{\pi(11)} & m_{\pi(12)} & m_{\pi(13)} \\
    m_{\pi(21)} & m_{\pi(22)} & m_{\pi(23)} \\
    m_{\pi(31)} & m_{\pi(32)} & m_{\pi(33)}
\end{pmatrix}
\]

(B.1)

and

\[
\begin{pmatrix}
    s'_{11} & s'_{12} & s'_{13} \\
    s'_{21} & s'_{22} & s'_{23} \\
    s'_{31} & s'_{32} & s'_{33}
\end{pmatrix} =
\begin{pmatrix}
    \gamma_{11} & \gamma_{12} & \gamma_{13} \\
    \gamma_{21} & \gamma_{22} & \gamma_{23} \\
    \gamma_{31} & \gamma_{32} & \gamma_{33}
\end{pmatrix} \odot
\begin{pmatrix}
    s_{\pi(11)} & s_{\pi(12)} & s_{\pi(13)} \\
    s_{\pi(21)} & s_{\pi(22)} & s_{\pi(23)} \\
    s_{\pi(31)} & s_{\pi(32)} & s_{\pi(33)}
\end{pmatrix}
\]

(B.2)

We must determine what constraints on \( \pi \), \( \{ C_{i} \} \) and \( \{ \gamma_{i} \} \) are implied by the requirement that the primed measurements have the compatibility relations of the Peres–Mermin scenario.

We begin by deriving some necessary conditions on the permutation \( \pi \). To have the right compatibility structure, it is necessary that each of the rows and columns of primed measurements in the Peres–Mermin square (and those of the primed sources) must constitute a compatible triple. It is necessary, therefore, that each must be the image, under the permutation \( \pi \), of a compatible triple in the original set of nine measurements (respectively sources), and given that the only triples that are compatible in the original set are the rows and columns, it follows that each must be the image of a row or column of the original. The overall permutation must therefore be generated by permutations of the rows, permutations of the columns, and the transpose of the Peres–Mermin square.

But this condition is not sufficient. We also require the particular compatibility relations among the nine primed measurements (and sources) to be those specified in the operational Peres–Mermin scenario. That is, we require that there exist six 4-outcome measurements \( M'_{1}, M'_{2}, \ldots, M'_{6} \) such that primed versions of equations (42a)–(42c) hold for the first row, and its analogues hold for the other rows and the first two columns, and primed versions of equations (44a)–(44c) hold for the third column, and we require that there exist six 4-outcome sources \( S'_{1}, S'_{2}, \ldots, S'_{6} \) such that primed versions of equations (45a)–(45c) hold for the first row,

\textsuperscript{16} The authors customized a redundancy removal script to utilize MOSEK’s efficient linear programming features; code available upon request. Filtering for extremality is also available natively in many polyhedral geometry software tools. A single-purpose freely-available tool is the \texttt{redund} binary which ships with \texttt{LR3}.
and its analogues hold for the other rows and the first two columns, and primed versions of equations (46a)--(46c) hold for the third column.

We require that for any triple corresponding to a row or a column in the primed Peres–Mermin square, the product of the outcomes in any given run of the experiment must be the same as the product of the outcomes from the corresponding triple in the original square. In other words, the deterministic processing is required to satisfy

$$m'_a m'_b m'_c = m_a m_b m_c$$

for all \((a, b, c)\) corresponding to compatible triples. Given equation (B.1), it follows that

$$\zeta_a \zeta_b \zeta_c = \frac{m_a m_b m_c}{m_{\pi(a)} m_{\pi(b)} m_{\pi(c)}}.$$  \hspace{1cm} (B.4)

Now note that the compatibility relations imply that for each of the nine compatible triples of measurements, the associated triple of outcomes that one can obtain in any run of the experiment has a product, namely,

$$m_1 m_2 m_3 = +1, \quad m_{12} m_{22} m_{33} = +1, \quad m_{13} m_{23} m_{31} = +1,$$

$$m_{12} m_{23} m_{32} = +1, \quad m_{23} m_{31} m_{12} = -1.$$  \hspace{1cm} (B.5)

(To be clear, we are not here assuming that the particular triple of values that are assigned to, say, \(m_{11}, m_{12}\), and \(m_{13}\), need to be assigned deterministically. Even if the assignment is indeterministic, so that this triple of values is sampled from some probability distribution, the compatibility relations defining the Peres-Mermin construction dictate that only those triples of values that satisfy \(m_1 m_2 m_3 = 1\) have a nonzero probability of occurrence.)

Consequently, if the permutation \(\pi\) is such that the third column is mapped to itself, then for all triples,

$$\zeta_a \zeta_b \zeta_c = +1.$$  \hspace{1cm} (B.6)

If, on the other hand, the permutation \(\pi\) maps the third column to a different triple, then

$$\zeta_a \zeta_b \zeta_c = \begin{cases} -1 & \text{for this distinguished pair of triples} \\ +1 & \text{for the other four triples.} \end{cases}$$  \hspace{1cm} (B.7)

The same constraints hold for the \(\gamma\)'s.

It is useful at this stage to distinguish three sorts of permutations of the Peres–Mermin square. Some notation is required to do so. Let the 2-cycle that interchanges the \((ij)\) element of the square with the \((i'j')\) element be denoted \(((ij)(i'j'))\). Let the permutation of the first and second row be denoted by \((R_i R_j)\), and similarly for other cases, that is, \((R_i R_j) = ((11)(21)(12)(22))(13)(23)\), etcetera. Let the transpose of the Peres–Mermin square be denoted \(T\), that is, \(T = ((13)(31))(12)(21))(23)(32)\). Finally, let the group closure of a set of permutations be denoted ‘clos’.

The set of all permutations of the Peres–Mermin square that map the third column to itself is:

$$\Pi_3 \equiv \{ \pi = \pi_C \pi_R : \pi_R \in \text{clos} \{(R_1 R_2), (R_2 R_3)\}, \quad \pi_C \in \text{clos} \{(C_1 C_2)\} \}.$$  \hspace{1cm} (B.8)

Here, \(\pi_R\) runs over all permutations of the rows, and \(\pi_C\) is either identity or the swap of the first two columns.

Of those permutations that do not take the third column to itself, it is useful to distinguish those that take it to another column, denoted \(\Pi_3\), and those that take it to a row, denoted \(\Pi_C\). These are defined as follows:

$$\Pi_3 \equiv \{ \pi = \pi_C \pi_R (C_2 C_3) : \pi_R \in \text{clos} \{(R_1 R_2), (R_2 R_3)\}, \quad \pi_C \in \text{clos} \{(C_1 C_2)\} \},$$  \hspace{1cm} (B.9)

where again \(\pi_R\) runs over all permutations of the rows, and \(\pi_C\) is either identity or the swap of the first two columns, and

$$\Pi_C \equiv \{ \pi = \pi_C \pi_R : \pi_R \in \text{clos} \{(R_1 R_2), (R_2 R_3)\}, \quad \pi_C \in \text{clos} \{(C_1 C_2), (C_2 C_3)\} \},$$  \hspace{1cm} (B.10)

where \(\pi_R\) runs over all permutations of the rows, and \(\pi_C\) runs over all permutations of the columns.

We can now specify what sorts of deterministic processings of the outcomes (for measurements and for sources) are allowed for each of these types of permutations.

The fact noted at equation (B.6) implies that if \(\pi \in \Pi_3\), then the \(\zeta\)'s must satisfy

$$\zeta_{11} \zeta_{12} \zeta_{13} = +1, \quad \zeta_{21} \zeta_{22} \zeta_{23} = +1, \quad \zeta_{31} \zeta_{32} \zeta_{33} = +1, \quad \zeta_{12} \zeta_{21} \zeta_{31} = +1,$$

$$\zeta_{13} \zeta_{32} \zeta_{31} = +1, \quad \zeta_{11} \zeta_{32} \zeta_{32} = +1, \quad \zeta_{13} \zeta_{23} \zeta_{33} = +1.$$  \hspace{1cm} (B.11)

The solutions of these equations include the case of \(\zeta_{ij} = +1\) for all \(ij\) as well as those of the form

$$\begin{pmatrix} \zeta_{11} & \zeta_{12} & \zeta_{13} \\ \zeta_{21} & \zeta_{22} & \zeta_{23} \\ \zeta_{31} & \zeta_{32} & \zeta_{33} \end{pmatrix} = \begin{pmatrix} 1 & -1 & +1 \\ -1 & 1 & +1 \\ +1 & +1 & +1 \end{pmatrix}.$$  \hspace{1cm} (B.12)

and any matrix of \(\zeta\)'s obtained from this one by permutation of the rows and columns. The set of such \(\zeta\) matrices, denoted \(\mathcal{Z}_3\), can be expressed as
Precisely the same analysis holds for the \( \gamma \)'s, so the set of \( \gamma \) matrices that preserve the compatibility relations if \( \pi \in \mathcal{P}_A \) is also \( \mathcal{A}_A \).

It follows that if
\[
(\pi, \{\zeta_j\}, \{\gamma_j\}) \in \mathcal{P}_A \times \mathcal{A}_A \times \mathcal{A}_A,
\]
then the primed measurements and sources have the compatibility relations of the Peres–Mermin scenario.

Next, consider the case where \( \pi \in \mathcal{P}_B \). It is useful to focus on a specific example, namely, \( \pi = (C_2 C_3) \). The fact noted at equation (B.7) implies that the \( \zeta \)'s must satisfy
\[
\zeta_{11} \zeta_{12} \zeta_{13} = +1, \quad \zeta_{21} \zeta_{22} \zeta_{23} = +1, \quad \zeta_{31} \zeta_{32} \zeta_{33} = +1, \quad \zeta_{11} \zeta_{21} \zeta_{31} = +1, \quad \zeta_{12} \zeta_{22} \zeta_{32} = -1, \quad \zeta_{13} \zeta_{23} \zeta_{33} = -1. \tag{B.19}
\]
Solutions of these equations are of the form
\[
\begin{pmatrix}
\zeta_{11} \\
\zeta_{21} \\
\zeta_{31}
\end{pmatrix} = \begin{pmatrix}
+1 & +1 & +1 \\
+1 & +1 & +1 \\
+1 & -1 & -1
\end{pmatrix} \; Z \in \mathcal{A}_A,
\]
and any matrix of \( \zeta \)'s obtained from this one by component-wise multiplication with an element of \( \mathcal{A}_A \). The set of such \( \zeta \)'s, denoted \( \mathcal{A}_B \), is therefore
\[
\mathcal{A}_B = \left\{ Z \circ \begin{pmatrix}
+1 & +1 & +1 \\
+1 & +1 & +1 \\
+1 & -1 & -1
\end{pmatrix} : Z \in \mathcal{A}_A \right\}. \tag{B.17}
\]

Similarly, the set of \( \gamma \) matrices that preserve the compatibility relations if \( \pi \in \mathcal{P}_B \) are those in \( \mathcal{A}_B \).

It follows that if
\[
(\pi, \{\zeta_j\}, \{\gamma_j\}) \in \mathcal{P}_B \times \mathcal{A}_B \times \mathcal{A}_B,
\]
then the primed measurements and sources have the compatibility relations of the Peres–Mermin scenario.

Finally, we consider the third class of permutations, \( \mathcal{P}_C \). Again, we begin by focusing on a particular permutation in that class, namely, the transpose, \( \pi = T \). The fact noted at equation (B.7) implies that in this case the \( \zeta \)'s must satisfy
\[
\zeta_{11} \zeta_{12} \zeta_{13} = +1, \quad \zeta_{21} \zeta_{22} \zeta_{23} = +1, \quad \zeta_{31} \zeta_{32} \zeta_{33} = -1, \quad \zeta_{11} \zeta_{21} \zeta_{31} = +1, \quad \zeta_{12} \zeta_{22} \zeta_{32} = -1, \quad \zeta_{13} \zeta_{23} \zeta_{33} = -1. \tag{B.19}
\]
Solutions of these equations are of the form
\[
\begin{pmatrix}
\zeta_{11} \\
\zeta_{21} \\
\zeta_{31}
\end{pmatrix} = \begin{pmatrix}
+1 & +1 & +1 \\
+1 & +1 & +1 \\
+1 & -1 & -1
\end{pmatrix} \; Z \in \mathcal{A}_A,
\]
and any matrix of \( \zeta \)'s obtained from this one by component-wise multiplication with an element of \( \mathcal{A}_A \). The set of such \( \zeta \)'s, denoted \( \mathcal{A}_C \), is therefore
\[
\mathcal{A}_C = \left\{ Z \circ \begin{pmatrix}
+1 & +1 & +1 \\
+1 & +1 & +1 \\
+1 & -1 & -1
\end{pmatrix} : Z \in \mathcal{A}_A \right\}. \tag{B.17}
\]

The set of \( \gamma \) matrices that preserve the compatibility relations if \( \pi \in \mathcal{P}_C \) are also those in \( \mathcal{A}_C \).

It follows that if
\[
(\pi, \{\zeta_j\}, \{\gamma_j\}) \in \mathcal{P}_C \times \mathcal{A}_C \times \mathcal{A}_C,
\]
then the primed measurements and sources have the compatibility relations of the Peres–Mermin scenario.

We have now exhaustively enumerated all of the choices of symmetry operations \((\pi, \{\zeta_j\}, \{\gamma_j\})\) that preserve the compatibility relations of the Peres–Mermin scenario. We turn to the question of how these symmetry operations transform the noncontextuality inequalities that we derive.

Recall that our noncontextuality inequalities will ultimately only refer to the following nine products of a measurement outcome variable with its corresponding source outcome variable:
\[
\begin{pmatrix}
S_{11} M_{11} & S_{12} M_{12} & S_{13} M_{13} \\
S_{21} M_{21} & S_{22} M_{22} & S_{23} M_{23} \\
S_{31} M_{31} & S_{32} M_{32} & S_{33} M_{33}
\end{pmatrix}. \tag{B.23}
\]
If we perform a deterministic processing \((\pi, \{\zeta_i\}, \{\gamma_j\})\) then the relevant products of outcome variables for the primed experiment are given by

\[
\begin{pmatrix}
  s'_{11} m_{11}^I & s'_{12} m_{12}^I & s'_{13} m_{13}^I \\
  s'_{21} m_{21}^I & s'_{22} m_{22}^I & s'_{23} m_{23}^I \\
  s'_{31} m_{31}^I & s'_{32} m_{32}^I & s'_{33} m_{33}^I
\end{pmatrix} = \begin{pmatrix}
  \eta_{11} \zeta_{11} & \eta_{12} \zeta_{12} & \eta_{13} \zeta_{13} \\
  \eta_{21} \zeta_{21} & \eta_{22} \zeta_{22} & \eta_{23} \zeta_{23} \\
  \eta_{31} \zeta_{31} & \eta_{32} \zeta_{32} & \eta_{33} \zeta_{33}
\end{pmatrix} \circ \begin{pmatrix}
  s_{11}(\pi_{11}) m_{11}(\pi_{11}) & s_{12}(\pi_{12}) m_{12}(\pi_{12}) & s_{13}(\pi_{13}) m_{13}(\pi_{13}) \\
  s_{21}(\pi_{21}) m_{21}(\pi_{21}) & s_{22}(\pi_{22}) m_{22}(\pi_{22}) & s_{23}(\pi_{23}) m_{23}(\pi_{23}) \\
  s_{31}(\pi_{31}) m_{31}(\pi_{31}) & s_{32}(\pi_{32}) m_{32}(\pi_{32}) & s_{33}(\pi_{33}) m_{33}(\pi_{33})
\end{pmatrix}.
\]

(B.24)

Note that only the component-wise product of the \(\zeta\) matrix and the \(\gamma\) matrix is relevant for the products of outcomes of interest. The sets of such products, in each of the three cases considered above, are

\[
\mathcal{D}_A := \{ Z \circ Z' : Z, Z' \in \mathcal{Z}_A \}, \mathcal{D}_B := \{ Z \circ Z' : Z, Z' \in \mathcal{Z}_B \}, \mathcal{D}_C := \{ Z \circ Z' : Z, Z' \in \mathcal{Z}_C \}.
\]

(B.25)

However, it is straightforward to verify that these three sets are equivalent,

\[
\mathcal{D}_A = \mathcal{D}_B = \mathcal{D}_C,
\]

and easy to determine. They are simply the union of \(\mathcal{Z}_A\) and all matrices of the form

\[
\begin{pmatrix}
  -1 & -1 & +1 \\
  -1 & +1 & -1 \\
  +1 & -1 & -1
\end{pmatrix}
\]

or obtainable from this one by permutations of the rows and columns.

Appendix C. The trivial facet inequalities of the noncontextual correlation polytope

Here we justify the claim that the inequalities in the symmetry class of equation (87) are trivial in the sense that they hold for any correlations arising in the operational Peres–Mermin scenario, independently of whether they admit of a noncontextual model or not. Consider the binary-outcome variables defined as products of outcome variables for source-measurement pairs in the first row of the Peres–Mermin square, \(s_{11} m_{11}, s_{12} m_{12}\) and \(s_{13} m_{13}\).

Recalling equations \((42a)–(42c)\), for every run of the experiment we have \(m_{11} m_{12} m_{13} = 1\), and recalling equations \((45a)–(45c)\), for each run we have \(s_{11} s_{12} s_{13} = 1\). This implies that for each run we have \((s_{11} m_{11})(s_{12} m_{12})(s_{13} m_{13}) = 1\). Given this constraint, it follows that the expectation values of the three variables, \(s_{11} m_{11}, s_{12} m_{12}\) and \(s_{13} m_{13}\), are determined by a joint probability distribution over two of them. But then, positivity of the probabilities in this distribution implies, by an argument articulated around equations (68), (69), that

\[
\text{for all } a, b \in \{-1, +1\} : -a \langle s_{11} m_{11} \rangle - b \langle s_{12} m_{12} \rangle - ab \langle s_{13} m_{13} \rangle \leq 1.
\]

(C.1)

Equation (87) is of this form.

Note that for the third column of the Peres–Mermin square, equations \((44a)–(44c)\) imply that for every run of the experiment we have \(m_{13} m_{23} m_{33} = -1\), and equations \((46a)–(46c)\) imply that for each run we have \(s_{13} s_{23} s_{33} = -1\), but these together yield \((s_{13} m_{13})(s_{23} m_{23})(s_{33} m_{33}) = 1\), exactly as we have for the other columns and the rows.

Appendix D. Criticism of a previous proposal for a noncontextuality inequality based on the Peres–Mermin square

In this section, a prior proposal for an experimental test of noncontextuality based on the Peres–Mermin proof of the Kochen–Specker theorem, that of Cabello [42], will be reviewed and criticized. Our criticisms here parallel those provided in appendix C of [14] for a similar proposal that was based on a different proof of the Kochen–Specker theorem.

We describe the proposal of [42] using the notation introduced in this article. There are nine operational quantities appearing in the inequality derived therein, corresponding to the expectation value, relative to an arbitrary preparation \(P\), of the product of the outcomes of each of the six triples of compatible measurements in the Peres–Mermin scenario: \((m_{11} m_{12} m_{31})_P, (m_{12} m_{22} m_{23})_P, \ldots, (m_{13} m_{23} m_{33})_P\). Only the average of these expectation values is considered in [42]; we denote it

\[
R(P) \equiv \langle m_{11} m_{12} m_{31} \rangle_P + \langle m_{12} m_{22} m_{23} \rangle_P + \langle m_{13} m_{23} m_{33} \rangle_P + \langle m_{11} m_{21} m_{31} \rangle_P + \langle m_{12} m_{22} m_{32} \rangle_P - \langle m_{13} m_{23} m_{33} \rangle_P.
\]

(D.1)

The inequality derived in [42] is

\[
R(P) \leq 4.
\]

(D.2)
We dispute the notion that this inequality delimits the boundary between contextual and noncontextual theories, and therefore also the notion that experimental violations thereof, such as those achieved in \cite{43–45}, have any bearing on the possibility of noncontextual models, for reasons that we presently outline.

We begin by reviewing the argument presented in favour of equation (D.2). Reference \cite{42} presumes that outcomes are assigned deterministically by the ontic state. Suppose that $R(\lambda)$ denotes a particular noncontextual deterministic assignment to the sum of the triple-products. Recalling our notational convention that $[\ ]_x$ denotes a deterministic assignment (see below equation (75)))

$$R(\lambda) \equiv \prod m_{11}m_{12}m_{13} + \prod m_{21}m_{22}m_{23} + \prod m_{31}m_{32}m_{33} + \prod m_{41}m_{42}m_{43}.$$ \hspace{1cm} (D.3)

Any preparation procedure $P$ corresponds to a convex sum of such assignments,

$$R(P) = \sum_\lambda R(\lambda)\mu(\lambda|P).$$ \hspace{1cm} (D.4)

The inequality of equation (D.2) is inferred from the inequality

$$R(\lambda) \leq 4,$$ \hspace{1cm} (D.5)

and the argument presented in favour of the latter is simply this: recalling that each of the nine $m_{ij}$ is assigned a value in $\{-1, +1\}$ in an outcome-deterministic assignment, there are $2^9$ possible joint assignments of values to $(m_{11}, \ldots, m_{33})$, and for each of these, one can verify that $R(\lambda) \leq 4$.

We now explain the problem with this argument. The measurement compatibility relations that serve to define the operational Peres–Mermin scenario, described around equations (42a)–(42c), imply that for each of the six compatible triples of measurements, the associated triple of outcomes that one can obtain in any run of the experiment has a fixed product. Specifically, this product is $-1$ for the triple corresponding to the third column, and $+1$ for the other five triples, as noted in equation (B.5).

However, for each of the $2^9$ joint assignments to $(m_{11}, \ldots, m_{33})$, one can identify a compatible triple which are assigned values that are inconsistent with the compatibility relation that they are meant to satisfy. For instance, the assignment of the $+1$ outcome to each of the nine measurements fails to give the correct product of outcomes for the third column. In other words, if a set of nine measurements satisfy the compatibility relations of the operational Peres–Mermin construction, then none of the $2^9$ joint assignments of outcomes to these measurements are logically possible assignments.

More formally, the fact that the assignments are assumed to be deterministic implies that the assignment to the product of the outcome is the product of the assignments to each outcome, $\prod m_{11}m_{12}m_{13} = \prod m_{11}, \prod m_{12}, \prod m_{13}$, and the compatibility relations that are assumed to hold among the measurements imply that the $([m_{11}], \ldots, [m_{33}])$ must satisfy the constraints of equations (75a)–(75f), which have no solution.

Furthermore, the compatibility relations imply that for any preparation $P$, whatever triple of compatible measurements one implements, the outcomes always satisfy equation (B.5), and therefore the expectation values satisfy these relations as well:

$$\langle m_{11}m_{12}m_{13}\rangle_P = +1, \langle m_{21}m_{22}m_{23}\rangle_P = +1, \langle m_{31}m_{32}m_{33}\rangle_P = +1, \langle m_{41}m_{42}m_{43}\rangle_P = +1,$$

$$\langle m_{11}m_{12}m_{13}\rangle_P = +1, \langle m_{31}m_{32}m_{33}\rangle_P = -1.$$ \hspace{1cm} (D.6)

In other words, just as the observables $\mathbb{I} \otimes \mathbb{I}$ and $-\mathbb{I} \otimes \mathbb{I}$ in quantum theory are trivial insofar as they take the same value for all states, the triple-product of outcomes $m_{11}m_{12}m_{13}, \ldots, m_{33}m_{33}m_{33}$ are trivial operational quantities insofar as they take the same value for all preparations. Substituting the identities in equation (D.6) into (D.1), one obtains

$$R(P) = 6.$$ \hspace{1cm} (D.7)

Therefore, for any set of measurements satisfying the compatibility relations of the operational Peres–Mermin scenario, one will necessarily find this equality to hold.

In particular, we expect equation (D.7) to hold no matter how noisy the measurements are. To see this, it suffices to note that the equalities of equation (B.5) hold for any set of noisy measurements satisfying the compatibility relations, and this implies that equation (D.7) holds for such measurements as well. For instance, equations (B.5) and (D.7) hold for the noisy quantum realization of the operational Peres–Mermin scenario, described in section 4.1.2, for any amount of depolarization noise. Consequently, if measuring $R(P)$ to have a value greater than $4$ could constitute evidence for the failure of noncontextuality, then this evidence could be obtained even in the presence of arbitrarily large amounts of noise. This is an indictment of the proposal of [42] because a minimal constraint on any reasonable notion of noncontextuality (first articulated in [14]) is that it should not be possible to demonstrate its failure in a completely incoherent experiment.
To summarize then, our criticism is as follows. The inequality $R(P) \leq 4$ should not be expected for any experiment satisfying the compatibility structure of the operational Peres–Mermion scenario, while the equality $R(P) = 6$ (and hence the violation of $R(P) \leq 4$) is expected to hold trivially in all such experiments. And this is the case regardless of whether the experiment admits of a noncontextual ontological model. As such, the operational quantity $R(P)$ contains no information about whether or not a noncontextual model can describe the experiment.

The most significant point of contrast between the proposal of this article and that of [42] is that we assume the notion of universal noncontextuality proposed in [8], rather than the notion of KS-noncontextuality. Because universal noncontextuality, unlike KS-noncontextuality, does not assume outcome determinism, we are led to consider indeterministic noncontextual assignments to the measurements. As we saw above, the fact that there are strictly no deterministic noncontextual assignments respecting the compatibility relations of the operational Peres–Mermion scenario is what makes it futile to attempt to derive constraints on operational statistics from the assumption of such assignments. The assumption is logically ruled out, so an operational test is neither necessary nor conceivable. On the other hand, there are many indeterministic noncontextual assignments that respect the compatibility relations, such as the example provided in equation (78). These, therefore, do impose nontrivial constraints on operational statistics, constraints that are encoded in the noncontextuality inequalities that we have derived.

References

[1] Harrigan N and Spekkens R W 2010 Einstein, incompleteness, and the epistemic view of quantum states Found. Phys. 40 125
[2] Kochen S and Specker E 1967 The problem of hidden variables in quantum mechanics J. Appl. Math. Mech. 17 59
[3] Mann C and Crease R 1988 Interview: John Bell, particle physicist Omni Mag. 10 84
[4] Bravyi S and Kitaev A 2005 Universal quantum computation with ideal clifford gates and noisy ancillas Phys. Rev. A 71 022316
[5] Knill E 2005 Quantum computing with realistically noisy devices Nature 434 39
[6] Howard M, Wallman J, Veitch V and Emerson J 2014 Contextuality supplies the ‘magic’ for quantum computation Nature 510 351
[7] Bermejo-Vega I, Delfosse N, Browne D E, Okay C and Raussendorf R 2017 Contextuality as a resource for models of quantum computation with qubits Phys. Rev. Lett. 119 120505
[8] Spekkens R W 2005 Contextuality for preparations, transformations, and unsharp measurements Phys. Rev. A 71 032108
[9] Spekkens R W 2014 The status of determinism in proofs of the impossibility of a noncontextual model of quantum theory Found. Phys. 44 1125
[10] Spekkens R W, Buzzacott D H, Keehn A J, Toner B and Pryde G J 2009 Preparation contextuality powers parity-oblivious multiplexing Phys. Rev. Lett. 102 010401
[11] Chailoulos A, Kerenidis I, Kundu S and Sikora J 2016 Optimal bounds for parity-oblivious random access codes New J. Phys. 18 045003
[12] Ambainis A, Banik M, Chaturvedi A, Kravchenko D and Rai A 2016 Parity oblivious d-level random access codes and class of noncontextuality inequalities arXiv:1607.05490
[13] Mazurek M D, Pusey M F, Resch K J and Spekkens R W 2016 An experimental test of noncontextuality without unphysical idealizations Nat. Commun. 7 11780
[14] Kunjwal R and Spekkens R W 2015 From the Kochen–Specker theorem to noncontextual inequalities without assuming determinism Phys. Rev. Lett. 115 110403
[15] Pusey M F 2013 The robust noncontextuality inequalities in the simplest scenario arXiv:1506.04178
[16] Peres A 1991 Two simple proofs of the Kochen–Specker theorem J. Phys. A 24 L175
[17] Mermin N D 1993 Hidden variables and the two theorems of John Bell Rev. Mod. Phys. 65 803
[18] Waegell M and Aravind P K 2012 Proofs of the Kochen–Specker theorem based on a system of three qubits J. Phys. A 45 405301
[19] Preskill J 2000 Chapter 10: quantum Shannon theory Quantum Information (Cambridge: Cambridge University Press)
[20] Liang Y-C, Spekkens R W and Wiseman H M 2011 Specker’s parable of the overprotective seer: a road to contextuality, nonlocality and complementarity Phys. Rep. 506 1
[21] Busch P, Grabowski M and Lahti P J 1997 Operational Quantum Physics vol 31 (Berlin: Springer)
[22] Kirchmair G, Zähringer F, Gerritsma R, Gühne O, Cabello A, Blatt R and Roos C F 2009 State-independent experimental test of quantum contextuality Nature 460 494
[23] Bell J S 1966 On the problem of hidden variables in quantum mechanics Rev. Mod. Phys. 38 447
[24] Fukuda K and Prodon A 1996 Double description method revisited Combinatorics and Computer Science (Lecture Notes on Computer Science vol 1120) (Berlin: Springer) 91–111
[25] Zolotykh N Y 2012 New modification of the double description method for constructing the skeleton of a polyhedral cone Comput. Math. Math. Phys. 52 146
[26] Avis D, Bremner D and Seidel R 1997 How good are convex hull algorithms? Comput. Geom. 7 265
[27] Barber C B, Dobkin D P and Huhdanpaa H 1996 The quickhull algorithm for convex hulls ACM Trans. Math. Softw. 22 469
[28] Mazurek M, Pusey M, Resch K J and Spekkens R W 2017 Experimentally bounding deviations from quantum theory in the landscape of generalized probabilistic theories in preparation arXiv:1710.05948
[29] Cabello A, Severini S and Winter A 2014 Graph-theoretic approach to quantum correlations Phys. Rev. Lett. 112 040401
[30] Acín A, Fritz T, Leverrier A and Sainz A B 2015 A combinatorial approach to nonlocality and contextuality Commun. Math. Phys. 334 533
[31] Kunjwal R, Heunen C and Fritz T 2014 Quantum realization of arbitrary joint measurability structures Phys. Rev. A 89 052126
[32] Budroni C and Cabello A 2012 Bell inequalities from variable-elimination methods J. Phys. A: Math. Theor. 45 385304
[33] Chaves R, Luit I. and Gross D 2014 Causal structures from entropic information: geometry and novel scenarios New J. Phys. 16 043001
[34] Chaves R 2016 Polynomial Bell inequalities Phys. Rev. Lett. 116 010402
[35] Wolfe E, Spekkens R W and Fritz T 2016 The inflation technique for causal inference with latent variables arXiv:1609.00672
[37] Lee C M and Spekkens R W 2017 Causal inference via algebraic geometry: feasibility tests for functional causal structures with two binary observed variables J. Causal Inference 5

[38] Cabello A 2010 Proposal for revealing quantum nonlocality via local contextuality Phys. Rev. Lett. 104 220401

[39] Liu B-H, Hu X-M, Chen J-S, Huang Y-F, Han Y-J, Li C-F, Guo G-C and Cabello A 2016 Nonlocality from local contextuality Phys. Rev. Lett. 117 220402

[40] Caron R J, McDonald J F and Ponic C M 1989 A degenerate extreme point strategy for the classification of linear constraints as redundant or necessary J. Optim. Theory Appl. 62 225

[41] Matousek J 1993 Linear optimization queries J. Algorithms 14 432

[42] Cabello A 2008 Experimentally testable state-independent quantum contextuality Phys. Rev. Lett. 101 210401

[43] Amselem E, Rådmark M, Bourennane M and Cabello A 2009 State-independent quantum contextuality with single photons Phys. Rev. Lett. 103 160405

[44] Liu B H, Huang Y F, Gong Y X, Sun F W, Zhang Y S, Li C F and Guo G C 2009 Experimental demonstration of quantum contextuality with nonentangled photons Phys. Rev. A 80 044101

[45] D’Ambrosio V, Herbauts I, Amselem E, Nagali E, Bourennane M, Sciarrino F and Cabello A 2013 Experimental implementation of a Kochen–Specker set of quantum tests Phys. Rev. X 3 011012