Article:

Pusey, Matthew Fairbairn orcid.org/0000-0002-6189-7144 (2018) Robust preparation noncontextuality inequalities in the simplest scenario. Physical Review A. 022112. ISSN 1094-1622

https://doi.org/10.1103/PhysRevA.98.022112

Robust preparation noncontextuality inequalities in the simplest scenario

Matthew F. Pusey
Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford OX1 3QD, United Kingdom and Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada, N2L 2Y5

(Received 17 May 2018; published 7 August 2018)

Contextuality is the leading notion of nonclassicality for a single system. However, an experimental demonstration requires finding procedures that are operationally equivalent, which might seem impossible to achieve exactly. Here I focus on the simplest nontrivial case, four preparations and two tomographically complete binary measurements. Exploiting a subtle connection to the Clauser-Horne-Shimony-Holt scenario gives eight nonlinear inequalities which are together necessary and sufficient for the experimental statistics to admit a preparation noncontextual model in such a scenario. No fixed operational equivalences are required, removing a key difficulty with experimental tests of older preparation noncontextuality inequalities.

DOI: 10.1103/PhysRevA.98.022112

I. INTRODUCTION

The gold standard for an experiment that defies classical explanation is the violation of a Bell inequality [1,2]. In the case of a single quantum system, this is not a possibility, and so attention has focused on contextuality.

Contextuality was first identified by Bell, Kochen, and Specker [3,4]. Although this was a profound insight into quantum mechanics, the definition they used is stated in quantum terms and applies only to the ideal of projective measurements. It is therefore not amenable to experimental test. A generalized definition due to Spekkens [5] is stated operationally and applies to arbitrary procedures. As shown in Ref. [6], this definition, or even just one component of it known as *preparation noncontextuality*, is thus well suited to experiment. (For an alternative perspective on experimental contextuality, not based on Spekkens's generalization, see for example [7–12].)

However, Ref. [6] used an assumption that two preparation procedures were indistinguishable, which was not satisfied exactly in the reported experiment and never will be in any experiment. Here I show how this problem can be eliminated by providing a full characterization of the preparation noncontextual statistics in the simplest scenario to which the concept applies. The shift in approach is that, with the help of tomographically complete measurements, indistinguishable preparation procedures are inferred from the statistics, rather than posited *a priori*.

II. DEFINITIONS

Consider an experiment where one implements a preparation procedure $\mathcal{P}_i$ followed by a measurement procedure $\mathcal{M}_j$ with outcome k, characterized by the probabilities $P(k|\mathcal{P}_i,\mathcal{M}_j)$. An *ontological model* seeks to explain these results via an *ontic state* λ that screens off the preparation from the measurement result:

$$P(k|\mathcal{P}_i,\mathcal{M}_j) = \int P(k|\lambda,\mathcal{M}_j)\mu_i(\lambda)\,d\lambda, \tag{1}$$

where we use the shorthand $\mu_i(\lambda) = P(\lambda|\mathcal{P}_i)$.

The explanation proffered by an ontological model is compelling only if it does justice to important features of the observed statistics. For example, in a bipartite scenario, Bell's locally causal models would provide a natural explanation for the observed no signaling [13].

Preparation noncontextuality concerns a closely related feature, namely, *operational equivalence* among preparations. Of particular relevance are operationally equivalent mixtures: suppose there exist probability distributions $\{p_i\}$ and $\{q_i\}$ such that for all j,k

$$\sum_i p_i P(k|\mathcal{P}_i,\mathcal{M}_j) = \sum_i q_i P(k|\mathcal{P}_i,\mathcal{M}_j). \tag{2}$$

Then we say that the probabilistic mixtures, which might be written $\sum_i p_i\mathcal{P}_i$ and $\sum_i q_i\mathcal{P}_i$, are operationally equivalent. In principle it is possible that somebody invents a new measurement procedure $\mathcal{M}'$ such that $\sum_i p_i P(k|\mathcal{P}_i,\mathcal{M}') \neq \sum_i q_i P(k|\mathcal{P}_i,\mathcal{M}')$, in which case the apparent operational equivalence would evaporate. For the time being we assume no such $\mathcal{M}'$ exists. In other words, we assume that the $\mathcal{M}_j$ are *tomographically complete* or *fiducial* [14] for the $\mathcal{P}_i$.

This assumption can be made without specifying any particular operational theory, but as an example: in the language of quantum theory, where preparations $\mathcal{P}_i$ are associated with density operators ρ_i, we need that $\sum_i p_i\rho_i = \sum_i q_i\rho_i$. This follows from Eq. (2) if and only if the positive operator-valued measure (POVM) elements associated with the $\mathcal{M}_j$ span the space of operators defined by the ρ_i [15].

What can explain the inability of any measurement to distinguish $\sum_i p_i\mathcal{P}_i$ from $\sum_i q_i\mathcal{P}_i$? The most natural explanation is that this "distinction without a difference" is no distinction at all:

$$\sum_i p_i\mu_i(\lambda) = \sum_i q_i\mu_i(\lambda). \tag{3}$$

The inference from the operational equivalence (2) to the ontic equivalence (3) constitutes the assumption of preparation noncontextuality [5]. (For comparison, measurement noncontextuality is the assumption that measurements that cannot be distinguished by the statistics for any preparation are

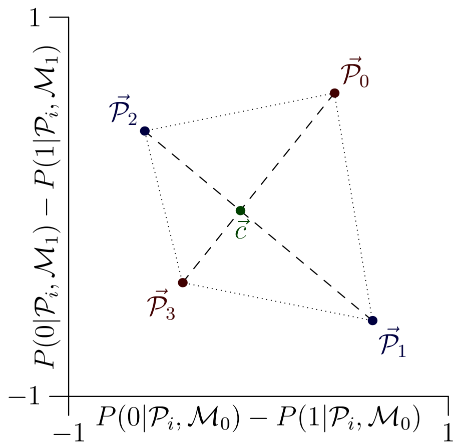

FIG. 1. Example statistics. The preparations have been labeled in accordance with the conventions that $\vec{\mathcal{P}}_0$ is opposite $\vec{\mathcal{P}}_3$ and $\{\vec{\mathcal{P}}_0-\vec{\mathcal{P}}_3,\vec{\mathcal{P}}_2-\vec{\mathcal{P}}_1\}$ is positively oriented. Also shown is $\vec{c}$ as defined by Eq. (5).

equivalent in the ontological model.) Note that this assumption is presented slightly differently in Refs. [5,6]; for readers familiar with the latter presentation the connection is made in Appendix A.

This article focuses on four preparations and two binary measurements. This is the simplest nontrivial scenario because, as shown in Appendix B, there is a noncontextual model for *any* operational probabilities in any simpler scenario. In a scenario with two binary measurements the operational probabilities for a single preparation $\mathcal{P}_i$ are given by the two-dimensional real vector $\vec{\mathcal{P}}_i=(P(0|\mathcal{P}_i,\mathcal{M}_0)-P(1|\mathcal{P}_i,\mathcal{M}_0),P(0|\mathcal{P}_i,\mathcal{M}_1)-P(1|\mathcal{P}_i,\mathcal{M}_1))$, which (along with normalization) fixes all four probabilities.

III. LABELING CONVENTIONS

Denoting the four preparations $\{\mathcal{P}_0,\mathcal{P}_1,\mathcal{P}_2,\mathcal{P}_3\}$, the $\vec{\mathcal{P}}_i=(x_i,y_i)$ must be the vertices of a convex quadrilateral, since any degeneracy will lead to a simplex and hence an immediate preparation noncontextual model by the argument in Appendix B. As in Fig. 1, we adopt the conventions that $\vec{\mathcal{P}}_0$ is opposite to $\vec{\mathcal{P}}_3$, and the $\vec{\mathcal{P}}_0-\vec{\mathcal{P}}_3$ and $\vec{\mathcal{P}}_2-\vec{\mathcal{P}}_1$ diagonals are positively oriented:

$$\begin{vmatrix} x_0-x_3 & x_2-x_1 \\ y_0-y_3 & y_2-y_1 \end{vmatrix} > 0. \tag{4}$$

IV. A PIVOTAL EQUIVALENCE

One example of an operational equivalence is given by the point $\vec{c}$ at which the $\{\vec{\mathcal{P}}_0,\vec{\mathcal{P}}_3\}$ diagonal intersects the $\{\vec{\mathcal{P}}_1,\vec{\mathcal{P}}_2\}$ diagonal, giving probabilities p,q such that

$$p\vec{\mathcal{P}}_0+(1-p)\vec{\mathcal{P}}_3=q\vec{\mathcal{P}}_1+(1-q)\vec{\mathcal{P}}_2=\vec{c}. \tag{5}$$

Preparation noncontextuality then demands that

$$p\mu_0(\lambda)+(1-p)\mu_3(\lambda)=q\mu_1(\lambda)+(1-q)\mu_2(\lambda). \tag{6}$$

I now show that in the current scenario, this single equivalence is in fact sufficient for a model to be preparation noncontextual.

Suppose we have a preparation contextual model, i.e., there exists p_i,q_i such that Eq. (2) holds yet Eq. (3) fails. We want to prove that Eq. (6) must also fail. The first step is to show that Eq. (3) must fail for some p_i',q_i' with $\sum_i p_i'\vec{\mathcal{P}}_i=\sum_i q_i'\vec{\mathcal{P}}_i=\vec{c}$. To see this, denote $\vec{p}=\sum_i p_i\vec{\mathcal{P}}_i=\sum_i q_i\vec{\mathcal{P}}_i$, and notice that, since $\vec{c}$ is in the interior of the quadrilateral, $\vec{c}=\sum_i r_i\vec{\mathcal{P}}_i+r_*\vec{p}$ for some probability distribution $\{r_0,r_1,r_2,r_3,r_*\}$ with $r_*>0$. But then $p_i'=r_i+r_*p_i$ and $q_i'=r_i+r_*q_i$ give the required instance, with the failure of Eq. (3) ensured by $r_*>0$ and the fact that Eq. (3) fails for the p_i,q_i.

Now I argue that there exist probabilities s,t such that $\sum_i p_i'\mathcal{P}_i$ amounts to preparing $p\mathcal{P}_0+(1-p)\mathcal{P}_3$ with probability s and $q\mathcal{P}_1+(1-q)\mathcal{P}_2$ with probability $(1-s)$ (similarly for the q_i' with t in place of s). Formally, this means $\{p_0',p_1',p_2',p_3'\}=\{sp,(1-s)q,(1-s)(1-q),s(1-p)\}$ and $\{q_0',q_1',q_2',q_3'\}=\{tp,(1-t)q,(1-t)(1-q),t(1-p)\}$. If that is the case then Eq. (6) implies $\sum_i p_i'\mu_i(\lambda)=\sum_i q_i'\mu_i(\lambda)$ and so the failure of the latter requires the failure of the former, which is what we wanted to prove.

To find s and t it is useful to make an affine transformation to a new coordinate system in which $\vec{c}=(0,0)$, $\vec{\mathcal{P}}_0=(1-p,0)$, and $\vec{\mathcal{P}}_1=(0,1-q)$. Then $p\vec{\mathcal{P}}_0+(1-p)\vec{\mathcal{P}}_3=\vec{c}$ gives $\vec{\mathcal{P}}_3=(-p,0)$ and $q\vec{\mathcal{P}}_1+(1-q)\vec{\mathcal{P}}_2=\vec{c}$ gives $\vec{\mathcal{P}}_2=(0,-q)$. Now $\sum_i p_i'\vec{\mathcal{P}}_i=\vec{c}$ becomes $(p_0'(1-p)-p_3'p,p_1'(1-q)-p_2'q)=(0,0)$. Defining $s=p_0'/p\geqslant 0$ we have $p_3'=s(1-p)$. Similarly defining $\bar{s}=p_1'/q\geqslant 0$ we have $p_2'=\bar{s}(1-q)$. $\sum_i p_i'=1$ gives $s+\bar{s}=1$, and repeating the same argument with the q_i' gives t.

V. THE CONNECTION TO CLAUSER-HORNE-SHIMONY-HOLT (CHSH)

It was first shown by Barrett that the existence of a preparation noncontextual model for a single system implies the existence of a locally causal model for any bipartite scenario involving that system [16]; in other words, any bipartite proof of Bell's theorem is a proof of preparation contextuality. The converse is not expected to hold in general (although certain proofs of preparation contextuality can be converted into bipartite proofs of Bell's theorem [17]). Nevertheless, with the reduction of the previous section in hand, Barrett's argument can be extended to see that the existence of a preparation noncontextual model for four preparations and two tomographically complete binary measurements is *equivalent* to the existence of a Bell local model in the scenario considered by CHSH [18].

In the relevant Bell scenario two parties choose between two measurements, their choices labeled x and y. They both obtain a binary outcome, labeled a and b. Their statistics $P(a,b|x,y)$ are related to the preparation noncontextuality scenario by

$$\begin{pmatrix} P(0,k|0,j) & P(1,k|0,j) \\ P(0,k|1,j) & P(1,k|1,j) \end{pmatrix} = \begin{pmatrix} pP(k|\mathcal{P}_0,\mathcal{M}_j) & (1-p)P(k|\mathcal{P}_3,\mathcal{M}_j) \\ qP(k|\mathcal{P}_1,\mathcal{M}_j) & (1-q)P(k|\mathcal{P}_2,\mathcal{M}_j) \end{pmatrix}. \tag{7}$$

These statistics are normalized, and are no signaling due to the operational equivalence (5). If we have a preparation noncontextual model, then set $\mu(\lambda) = p\mu_0(\lambda) + (1-p)\mu_3(\lambda)$,

$$\begin{pmatrix} P_A(0|\lambda,0) & P_A(1|\lambda,0) \\ P_A(0|\lambda,1) & P_A(1|\lambda,1) \end{pmatrix} = \begin{pmatrix} p\mu_0(\lambda)/\mu(\lambda) & (1-p)\mu_3(\lambda)/\mu(\lambda) \\ q\mu_1(\lambda)/\mu(\lambda) & (1-q)\mu_2(\lambda)/\mu(\lambda) \end{pmatrix} \quad (8)$$

[which is normalized by Eq. (6)], and $P_B(k|\lambda, j) = P(k|\lambda, \mathcal{M}_j)$. Then Eq. (1) gives the locally causal model

$$p(a,b|x,y) = \int P_A(a|\lambda,x)P_B(b|\lambda,y)\mu(\lambda)\,d\lambda. \quad (9)$$

If, conversely, we start with a locally causal model, then inverting Eq. (8) gives an ontological model for $P(k|\mathcal{P}_i, \mathcal{M}_j)$, where Eq. (6) is guaranteed by the normalization of the locally causal model. Since we have seen that Eq. (6) is sufficient for a preparation noncontextual model, we have established the desired equivalence.

Fine [19] has shown that the eight versions of the CHSH inequality [18] are necessary and sufficient for the existence of a locally causal model. Hence given $P(k|\mathcal{P}_i, \mathcal{M}_j)$, one can calculate the corresponding Bell scenario probabilities (7) and then use the eight CHSH inequalities to determine whether or not a preparation noncontextual model exists.

VI. A CLOSED EXPRESSION

The above argument completely characterizes the noncontextual statistics in our scenario. However, it might appear that if one were to calculate p and q explicitly and substitute Eq. (7) into the CHSH inequalities the result would be extremely convoluted. In fact it can be written in a remarkably simple form, thanks to following lemma.

Lemma 1. Suppose $\vec{\mathcal{P}}_i = (x_i, y_i)$ for $i = 0,1,2,3$ satisfies Eq. (4), and p and q are defined as the solutions of Eq. (5). Then for any real numbers $\{z_0, z_1, z_2, z_3\}$,

$$pz_0 + (1-p)z_3 \leqslant qz_1 + (1-q)z_2 \quad (10)$$

if and only if

$$\begin{vmatrix} x_0 & y_0 & z_0 & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} \leqslant 0. \quad (11)$$

Furthermore, equality in Eqs. (10) and (11) is also equivalent.

Geometrically this lemma concerns a tetrahedron with vertices (x_i, y_i, z_i). Equation 10 asks whether, when the 0-3 edge meets the 1-2 edge in the (x,y) plane, it is below in the z direction. Subject to the convention (4), that is equivalent to the statement (11) about the signed volume of the tetrahedron. A purely algebraic proof can be given as follows.

Proof. Denoting the left-hand side (LHS) of Eq. (4) by D, Cramer's rule gives

$$p = \begin{vmatrix} x_2 - x_3 & x_2 - x_1 \\ y_2 - y_3 & y_2 - y_1 \end{vmatrix} \Big/ D, \quad (12)$$

$$q = \begin{vmatrix} x_0 - x_3 & x_2 - x_3 \\ y_0 - y_3 & y_2 - y_3 \end{vmatrix} \Big/ D. \quad (13)$$

Substituting these into (10) and multiplying through by $D > 0$ gives, upon expanding all the determinants, (11). The reader can avoid an algebraic quagmire by referring to the MATHEMATICA notebook (whose output is also provided in a pdf file) [20]. ■

If we substitute Eq. (7) into the CHSH inequalities we obtain (10) with $z_0 = c_0x_0 + d_0y_0 - 1$, $z_1 = -c_1x_1 - d_1y_1 + 1$, $z_2 = c_1x_2 + d_1y_2 + 1$, and $z_3 = -c_0x_3 - d_0y_3 - 1$, where (c_0, d_0, c_1, d_1) is a column of

$$\begin{pmatrix} 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \end{pmatrix}. \quad (14)$$

(There is one column for each version of the CHSH inequality.) Substituting the z_i into (11) then gives the desired inequality. For example the $\langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle \leqslant 2$ version of the CHSH inequality corresponds to the first column of (14), and (11) becomes

$$\begin{vmatrix} x_0 & y_0 & x_0 + y_0 - 1 & 1 \\ x_1 & y_1 & -x_1 + y_1 + 1 & 1 \\ x_2 & y_2 & x_2 - y_2 + 1 & 1 \\ x_3 & y_3 & -x_3 - y_3 - 1 & 1 \end{vmatrix} \leqslant 0. \quad (15)$$

[Subject to (4), this can still serve as a Bell inequality, but now in terms of $p(a|b,x,y)$ rather than $p(a,b|x,y)$.]

VII. QUANTUM VIOLATION

Suppose $\mathcal{M}_0$ and $\mathcal{M}_1$ correspond to X and Z measurements of a qubit. Let $\mathcal{P}_0$ correspond to preparing the $+1$ eigenstate of $(X+Z)/\sqrt{2}$ and $\mathcal{P}_3$ the -1 eigenstate. Similarly let $\{\mathcal{P}_1, \mathcal{P}_2\}$ be the $\{+1,-1\}$ eigenstates of $(X-Z)/\sqrt{2}$. Denoting $v = 1/\sqrt{2}$ the LHS of (15) is

$$\begin{vmatrix} v & v & 2v-1 & 1 \\ v & -v & 1-2v & 1 \\ -v & v & 1-2v & 1 \\ -v & -v & 2v-1 & 1 \end{vmatrix} = 16v^2(2v-1) \approx 3.31. \quad (16)$$

This is the same proof of the preparation contextuality of a qubit that appeared in Ref. [6]. However, that proof assumed $\frac{1}{2}\mathcal{P}_1 + \frac{1}{2}\mathcal{P}_3 = \frac{1}{2}\mathcal{P}_1 + \frac{1}{2}\mathcal{P}_2$, which will never hold exactly in a realistic experiment. Since no such assumption entered into (15), the proof presented here is more experimentally robust.

In Appendix C a correspondence is established between quantum strategies in the preparation contextuality scenario and quantum strategies in the CHSH scenario. Since the above corresponds with the strategy for maximally [21] violating the CHSH inequality, it is the quantum maximum for the contextuality scenario.

VIII. TOMOGRAPHIC COMPLETENESS

A difficulty remains in the present approach, namely, our assumption that $\mathcal{M}_0$ and $\mathcal{M}_1$ are tomographically complete for the $\mathcal{P}_i$. An obvious objection in the qubit example of the previous section is the Y measurement. Suppose we have a preparation noncontextual model for three measurements $\{\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}'\}$, and consider a set of preparations that all give the same probability for $\mathcal{M}'$. Since Eq. (2) holds for all

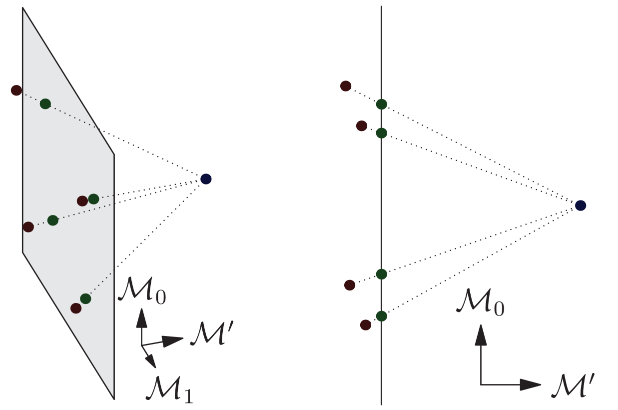

FIG. 2. Dealing with failure of tomographic completeness. The statistics for three binary measurements $\{\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}'\}$ are defined by a three-dimensional vector for each preparation. Two views of this space are shown. The four nonplanar preparations in brown, together with additional preparation in blue, imply the existence of the four planar preparations in green. The statistics for the green preparations can be calculated with simple trigonometry and then tested against the inequalities derived here for two binary measurements.

three measurements if and only if it holds for $\{\mathcal{M}_0, \mathcal{M}_1\}$, the problem reduces to finding a preparation noncontextual model for $\mathcal{M}_0$ and $\mathcal{M}_1$. Since the quantum states mentioned above all give uniformly random outcomes for the Y measurement, there is in principle nothing wrong with using (15).

In practice, however, the preparations in a real experiment will each give slightly different probabilities for the Y measurement. The simplest way to deal with this is to add a fifth preparation that gives one outcome of the Y measurement with high probability (e.g., the $+1$ eigenstate of Y). If, as in Fig. 2, we then consider a plane perpendicular to the Y axis with the four original preparations on one side and the new preparations on the other, convexity implies the existence of, and determines the probabilities for, four preparations in the plane that can then be tested against (15). Notice there is no need to actually implement the four new preparations. Any further measurements that reveal small amounts of information about the preparations can be dealt with similarly.

This idea has been generalized [22] into a technique for identifying experimental violations of other noncontextuality inequalities, even those requiring fixed equivalences.

These techniques deal with additional measurements in the tomographically complete set, but it requires that those additional measurements are actually performed. Hence any test based on these techniques will still require the assumption that there are not any unknown measurements that would be required to construct a tomographically complete set, i.e., unknown measurements whose statistics cannot be inferred from the measurements that have been performed. No test based on Ref. [5] can avoid such an assumption (or some other assumption that restricts possible measurements), because it is always logically possible that there exists a measurement that simply reads out a complete description of the preparation. In that case, no two distinct preparations would be equivalent, and so any model would be trivially noncontextual.

It should be noted that an assumption about the tomographically complete set is much weaker than assuming all of quantum theory. Examples of other theories where the simplest system has three binary measurements in the tomographically complete set include Spekkens's toy theory [23] (which is noncontextual), quantum theory with fundamental decoherence [24,25] (whose contextuality would depend on the amount of decoherence been preparation and measurement), and a version of generalized no-signaling theory or "boxworld" [26] [which could violate (15) *more* than quantum theory]. One way to lend credence to such an assumption without reference to quantum theory would be to experimentally test whether the statistics of a large number of measurements can be inferred from a small subset [27]. More speculatively, it may be possible to support the assumption via physical principles independent of quantum theory, for example thermodynamical principles (cf. Ref. [28]).

IX. CONCLUSIONS

The eight inequalities derived here fully classify the preparation noncontextual statistics in the simplest nontrivial scenario. No assumptions of unattainable operational equivalences were made. No new assumptions on the representation of approximately operationally equivalent procedures [11,12] were made either. Furthermore, it was not assumed that the measurements are represented deterministically in the ontological model, which (in quantum terms) means it does not matter whether the measurements are projective [5,29]. Hence a violation can only be explained by a failure of noncontextuality or a failure of the tomographic completeness of the measurements. Since observed failures of the latter can be dealt with using convexity, nonclassicality may be left as the only plausible explanation.

The main extension of these results would be to classify scenarios with more preparations and measurements. Additional (performed) measurements in the tomographically complete set could then be dealt with more elegantly than above, by fully incorporating the extra procedures into the contextuality scenario. In Appendix D some of the results above are generalized to such scenarios, and the limitations of these generalizations are discussed. It would also be interesting to apply similar ideas to measurement and transformation noncontextuality. More broadly, the status of tomographic completeness assumptions in tests of contextuality deserves further study.

ACKNOWLEDGMENTS

Thanks to Rob Spekkens for invaluable discussions, to Lídia del Rio for feedback on a previous version, and to Jon Barrett for sharing his result [16] and allowing me to include it here. Research at Perimeter Institute is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MRI. I am now supported by the Royal Commision for the Exhibition of 1851.

APPENDIX A: FORMULATING PREPARATION NONCONTEXTUALITY

In Ref. [5], preparation noncontextuality is defined as the requirement that if $P(k|\mathcal{P}, \mathcal{M}) = P(k|\mathcal{P}', \mathcal{M})$ for all $\mathcal{M}$, then $P(\lambda|\mathcal{P}) = P(\lambda|\mathcal{P}')$. Here I show if we are only interested in a finite set of preparations $\{\mathcal{P}_i\}$ (and convex combinations thereof), and we assume the $\mathcal{M}_j$ are tomographically complete, then this definition is equivalent to the definition in Sec. II, i.e., the requirement that if Eq. (2) holds then Eq. (3) holds.

Define $\mathcal{P}$ as the procedure of preparing $\mathcal{P}_i$ with probability p_i, and $\mathcal{P}'$ as the procedure of preparing $\mathcal{P}_i$ with probability q_i, and notice that any of the preparations we are interested in are of this form (in particular $\{p_i\}$ can assign probability 1 to a preparation). By definition $P(k|\mathcal{P},\mathcal{M}) = \sum_i p_i P(k|\mathcal{P}_i,\mathcal{M})$ and $P(k|\mathcal{P}',\mathcal{M}) = \sum_i q_i P(k|\mathcal{P}_i,\mathcal{M})$. Hence Eq. (2) is exactly the statement that $P(k|\mathcal{P},\mathcal{M}_j) = P(k|\mathcal{P}',\mathcal{M}_j)$ for all j. Tomographic completeness means that this is equivalent to $P(k|\mathcal{P},\mathcal{M}) = P(k|\mathcal{P}',\mathcal{M})$ for all $\mathcal{M}$. So the "if" conditions of the two definitions are equivalent.

Again by the definitions of $\mathcal{P}$ and $\mathcal{P}'$, $P(\lambda|\mathcal{P}) = \sum_i p_i\mu_i(\lambda)$ and $P(\lambda|\mathcal{P}') = \sum_i q_i\mu_i(\lambda)$ and so Eq. (3) is exactly $P(\lambda|\mathcal{P}) = P(\lambda|\mathcal{P}')$. So the "then" implications of each definition are also equivalent, and we are done.

APPENDIX B: IDENTIFYING THE SIMPLEST SCENARIO

Here I consider scenarios simpler than the four preparations and two binary measurements discussed in the main text, and show that they all trivially admit noncontextual models for any values of the operational probabilities $P(k|\mathcal{P}_i,\mathcal{M}_j)$. The basic ideas are as follows. For one measurement, a noncontextual model can be obtained by having the ontic state encode the outcome. On the other hand, if we have so few preparations that they form a simplex in the space of operational probabilities then the ontic state can simply encode the identity of the preparation.

In detail, suppose we consider only a single measurement $\mathcal{M}_1$. There is a simple "$\lambda = k$" model, where the ontic state simply specifies the outcome of $\mathcal{M}_1$, i.e., $P(k|\lambda,\mathcal{M}_1) = \delta_{k\lambda}$. Equation 1 is ensured by distributing the ontic states according to $\mu_i(\lambda) = P(k=\lambda|\mathcal{P}_i, M_1)$. This is manifestly preparation noncontextual because it makes Eq. (3) the same as Eq. (2). Hence the simplest nontrivial scenario must have at least two measurements, the simplest such case being two binary (two-outcome) measurements.

Now suppose we consider three or fewer preparations with two binary measurements. As noted in Sec. II, we can associate preparations in such a scenario with vectors $\vec{\mathcal{P}}_i = (P(0|\mathcal{P}_i,\mathcal{M}_0) - P(1|\mathcal{P}_i,\mathcal{M}_0), P(0|\mathcal{P}_i,\mathcal{M}_1) - P(1|\mathcal{P}_i,\mathcal{M}_1))$. The convex hull of any one to three $\vec{\mathcal{P}}_i$ is a simplex: a point, a line segment, or a triangle. Recall that every point in a simplex has exactly one decomposition into extremal points. Hence we consider the ontological model in which there is an ontic state λ for each extreme point $\vec{\lambda}$ of that simplex, with $\mu_i(\lambda)$ being the unique distribution ensuring that $\sum_{\lambda}\mu_i(\lambda)\vec{\lambda} = \vec{\mathcal{P}}_i$. Naturally $P(k|\lambda,\mathcal{M}_i)$ is defined so that $\vec{\lambda} = (P(0|\lambda,\mathcal{M}_0) - P(1|\lambda,\mathcal{M}_0), P(0|\lambda,\mathcal{M}_1) - P(1|\lambda,\mathcal{M}_1))$, so that Eq. (1) is satisfied. Expanding Eq. (2) as

$$\sum_{i,\lambda} p_i\mu_i(\lambda)\vec{\lambda} = \sum_{i,\lambda} q_i\mu_i(\lambda)\vec{\lambda} \tag{B1}$$

and applying uniqueness again immediately gives Eq. (3), so our model is preparation noncontextual.

APPENDIX C: QUANTUM CONNECTION WITH THE BELL SCENARIO

In Sec. V it is shown that for the scenario considered, the $P(k|\mathcal{P}_i,\mathcal{M}_j)$ admit a preparation noncontextual model if and only if the $P(a,b|x,y)$ defined in Eq. (7) admit a locally causal model. Here I sketch an argument that, similarly, the $P(k|\mathcal{P}_i,\mathcal{M}_j)$ admit a quantum realization (satisfying the tomographic completeness assumption) if and only if the $P(a,b|x,y)$ are quantum realizable in the usual sense that

$$P(a,b|x,y) = \mathrm{Tr}((E_{a|x}\otimes F_{b|y})\rho_{AB}) \tag{C1}$$

for a bipartite quantum state ρ_{AB} and sets of POVMs $\{E_{a|x}\}$ and $\{E_{b|y}\}$.

For the "if" part, let $\rho_i \propto \mathrm{Tr}_A((E_i\otimes I)\rho_{AB})$ be the normalized steered state when Alice measures $E_0 = E_{0|0}, E_1 = E_{0|1}, E_2 = E_{1|1}$, and $E_3 = E_{1|0}$, respectively. The basic idea is that in the noncontextuality scenario $\mathcal{P}_i$ will correspond to preparing ρ_i and $\mathcal{M}_j$ to measuring $\{F_{k|j}\}$. Certainly this will achieve the correct $P(k|\mathcal{P}_i,\mathcal{M}_j)$ according to Eq. (7). However, we also need that the $\mathcal{M}_j$ are tomographically complete for the $\mathcal{P}_i$. Plotting the $P(k|\mathcal{P}_i,\mathcal{M}_j)$ as in Fig. 1, there are two cases to consider. The first is that the $\mathcal{P}_i$ form a two-dimensional shape. In this case the $\mathcal{M}_j$ must be tomographically complete because the ρ_i certainly live in *some* two-dimensional affine subspace of the quantum states by the no-signaling condition $p\rho_0 + (1-p)\rho_3 = q\rho_1 + (1-q)\rho_2 = \mathrm{Tr}_A(\rho_{AB})$, and so the only way for the projection onto the $\mathcal{M}_j$ to be two-dimensional is if they span that space. The second case is that the $\mathcal{P}_i$ form a one-dimensional shape, which is necessarily a simplex. Hence the statistics can be reproduced using the preparation noncontextual "extreme point" model of the previous section, which can be implemented with the ontic state encoded in a qubit and the $\mathcal{M}_j$ will be tomographically complete by construction.

For the "only if" part, there are again two cases that arise on consideration of the geometry in Fig. 1. The first case is that the $\mathcal{P}_i$ indeed form a nondegenerate convex quadrilateral as shown in the figure. Adopting the specified labeling convention, we then have $p\rho_0 + (1-p)\rho_3 = q\rho_1 + (1-q)\rho_2 =: \rho_B$ for some probabilities p and q (we are using tomographic completeness to derive that the two mixtures having the same statistics for the $\mathcal{M}_j$ implies they correspond to the same density operator). Let ρ_{AB} be a purification of ρ_B. By the Schrödinger-Hughston-Jozsa-Wootters theorem [30,31] there exist measurements for Alice that steer Bob onto the two decompositions of ρ_B, giving the required measurements for the correct $P(a,b|x,y)$ in the Bell scenario. The second case is that the convex hull of the $\mathcal{P}_i$ is a simplex, in which case there is a noncontextual model by the argument in the previous section, and hence a Bell-local model for the corresponding $P(a,b|x,y)$ by the argument in Sec. V. Any Bell-local $P(a,b|x,y)$ is also a quantum $P(a,b|x,y)$.

APPENDIX D: GENERALIZATIONS TO OTHER SCENARIOS

Here I provide generalizations of some of the steps in the main text to scenarios involving more preparations or measurements, and discuss the difficulties in using these techniques to provide a full classification of such scenarios.

Notice that if any finite number of measurements are tomographically complete, the preparations $\mathcal{P}_i$ are characterized by a finite number of probabilities and hence can be considered as points in a finite-dimensional vector space. This generalizes the two-dimensional space of $\vec{\mathcal{P}}_i$ discussed in the main text; for notational simplicity we do not distinguish a preparation from its vector here.

1. Only need to look at decompositions of one state

The following is a straightforward generalization to arbitrary numbers of preparations and measurements of the first step in Sec. IV.

Lemma 2. Let $\mathcal{P}^*$ be in the interior of the convex hull of the $\{\mathcal{P}_i\}$. An ontological model for the $\{\mathcal{P}_i\}$ is preparation noncontextual if and only if there exists a $\mu^*(\lambda)$ such that for all probability distributions $\{p_i\}$ such that

$$\sum_i p_i\mathcal{P}_i = \mathcal{P}^*, \tag{D1}$$

we have

$$\sum_i p_i\mu_i(\lambda) = \mu^*(\lambda). \tag{D2}$$

Proof. The "only if" part is trivial; simply take $\mu(\lambda) = \sum_i p_i\mu_i(\lambda)$ for one decomposition of $\mathcal{P}^*$ and note that preparation noncontextuality ensures this works for any other decompositions.

For the "if" part, suppose that an ontological model is preparation contextual. In particular, there exist distributions $\{p_i\}$ and $\{q_i\}$ such that Eq. (2) holds and yet Eq. (3) fails. Define $\tilde{\mathcal{P}} = \sum_i p_i\mathcal{P}_i$ $(= \sum_i q_i\mathcal{P}_i)$. Since $\mathcal{P}^*$ is in the interior of the state space, $\mathcal{P}^* = \sum_i r_i\mathcal{P}_i + \tilde{r}\tilde{\mathcal{P}}$ for some probability distribution $\{r_i, \tilde{r}\}$ with $\tilde{r} > 0$. But then letting $p_i' = r_i + \tilde{r}p_i$ and $q_i' = r_i + \tilde{r}q_i$ give two distributions that satisfy Eq. (D1), yet the failure of Eq. (3) for p_i, q_i and $r_* > 0$ means that $\sum_i p_i'\mu_i(\lambda) \neq q_i'\mu_i(\lambda)$. Hence Eq. (D2) cannot be satisfied for both $\{p_i'\}$ and $\{q_i'\}$. ■

2. Only need to look at a finite number of decompositions

A somewhat less straightforward generalization of the second step in Sec. IV is presented.

Theorem 1. Let $\mathcal{P}^*$ be in the interior of the state space. An ontological model is preparation noncontextual if and only if there exists a $\mu^*(\lambda)$ such that for all probability distributions $\{p_i\}$, with $\{\mathcal{P}_i : p_i > 0\}$ forming a simplex, and such that

$$\sum_i p_i\mathcal{P}_i = \mathcal{P}^*, \tag{D3}$$

we have

$$\sum_i p_i\mu_i(\lambda) = \mu^*(\lambda). \tag{D4}$$

Furthermore, there are a finite number of such $\{p_i\}$.

We need the following:

Lemma 3. Consider a real d-dimensional vector space V, a finite set of points $v_i \in V$, and a further point $v \in V$. The set of probability distributions $\{p_i\}$ such that $\sum_i p_i v_i = v$ is a closed convex polytope, the extreme points of which are exactly those elements supported on i such that the v_i form a simplex.

Proof. Such p_i are defined by the finite set of linear constraints $p_i \geqslant 0$, $\sum_i p_i = 1$, and $\sum_i p_i v_i = v$, hence forming a closed convex polytope.

Suppose $\{p_i^*\}$ is an element of the polytope, and that its support $S^* = \{i : p_i^* > 0\}$ defines a nonsimplex v_i for $i \in S^*$. Hence the size of S^* must be at least $d^* + 2$, where d^* is the dimension of the affine span of the v_i with $i \in S^*$. By Carathéodory's theorem in convex geometry there exists another element $\{p_i'\}$ of the polytope, supported on a proper subset of S. But then for a sufficiently small value of q, with $0 < q < 1$, $p_i'' = (p_i^* - qp_i')/(1-q)$ will satisfy $0 \leqslant p_i'' \leqslant 1$. Noting that p_i'' is then another element of the polytope, and that $p_i^* = qp_i' + (1-q)p_i''$, we see that $\{p_i^*\}$ is not extremal.

On the other hand, suppose $\{p_i^*\}$ has support $S^* = \{i : p_i^* > 0\}$ such that the v_i form a simplex. Any other elements of the polytope that $\{p_i^*\}$ can be decomposed into must have the same (or smaller) support. But there is only one way to write an element of a simplex as a convex combination of the vertices of the simplex, and so the decomposition must be trivial. That is to say, $\{p_i^*\}$ is extremal. ■

Proof of Theorem 1. The "only if" part is again trivial.

For the "if" part, by Lemma 2 we only need to ensure that μ^* satisfies Eq. (D2) whenever $\{p_i\}$ satisfies Eq. (D1). Letting v_i and v represent $\mathcal{P}_i$ and $\mathcal{P}^*$ respectively in Lemma 3, we see that any $\{p_i\}$ satisfying Eq. (D1) is an element of the described polytope. Any element of a polytope can be written as a convex combination of its extreme points, and so there exists a distribution q_j such that

$$\sum_j q_j p_i^{(j)} = p_i, \tag{D5}$$

where for each j, $\{p_i^{(j)}\}$ is a distribution supported on i such that $\mathcal{P}_i$ form a simplex. But then by assumption we have

$$\sum_i p_i^{(j)}\mu_i(\lambda) = \mu^*(\lambda). \tag{D6}$$

Combining Eq. (D5) with Eq. (D6) we find

$$\sum_i p_i\mu_i(\lambda) = \sum_{i,j} q_j p_i^{(j)}\mu_i(\lambda) = \sum_j q_j\mu^*(\lambda) = \mu^*(\lambda) \tag{D7}$$

as required. ■

In the main text the special case of four preparations forming a quadrilateral with $\mathcal{P}^*$ at the intersection of the diagonals was used. In that case Theorem 1 shows that you only need to check the two decompositions onto opposite pairs of corners.

3. The connection with Bell's theorem

In Sec. V the preparation contextuality scenario under consideration was linked with the CHSH scenario based on the ideas of Ref. [16]. In light of the above results, one might hope that more complicated preparation contextuality scenarios can be also be fully characterized by translating to Bell scenarios. Unfortunately that appears not to be the case: even though Theorem 1 reduces the problem to a finite number of decompositions of a single preparation, the same $\mathcal{P}_i$ will normally appear in multiple decompositions. As far as I can see there is no way to enforce that the corresponding μ_i are the same when converting to a Bell scenario without breaking the linearity that is so important computationally. This leaves the cases where the sets of $\mathcal{P}_i$ happen to be disjoint:

Theorem 2. Suppose there exists a $\mathcal{P}^*$ such that each i appears in the support of exactly one of the decompositions $\{p_i^{(1)}\}$, $\{p_i^{(2)}\}$, ..., described in Theorem 1. Then there exists a preparation noncontextual model if and only if the bipartite probabilities

$$P(i,k|x,j) = p_i^{(x)} P(k|\mathcal{P}_i,\mathcal{M}_j) \tag{D8}$$

admit a locally causal model.

Proof. For the "if" part: starting with a locally causal model

$$P(i,k|x,j) = \int P(i|\lambda,x)P(k|\lambda,\mathcal{M}_j)\mu^*(\lambda)\,d\lambda, \tag{D9}$$

we can define an ontological model for the contextuality scenario by

$$\mu_i(\lambda) = \frac{\mu^*(\lambda)P(i|\lambda,x(i))}{p_i^{(x(i))}}, \tag{D10}$$

where $x(i)$ is uniquely defined by $p_i^{(x(i))} > 0$.

By construction this model satisfies the requirement of Theorem 1 and so is preparation noncontextual. By Eqs. (D8)–(D10) its operational predictions are

$$\begin{aligned}\int P(k|\lambda,\mathcal{M}_j)\mu_i(\lambda)\,d\lambda &= \int P(k|\lambda,\mathcal{M}_j)\frac{\mu^*(\lambda)P(i|\lambda,x(i))}{p_i^{(x(i))}}\,d\lambda \\ &= \frac{P(i,k|x(i),j)}{p_i^{(x(i))}} = P(k|\mathcal{P}_i,\mathcal{M}_j)\end{aligned} \tag{D11}$$

as required.

For the "only if" part: starting with a preparation noncontextual model $\mu_i(\lambda)$ and $P(k|\lambda,\mathcal{M}_j)$ we can define a locally causal model using the μ^* from Theorem 1 and

$$P(i|\lambda,x) = \frac{\mu_i(\lambda)p_i^{(x)}}{\mu^*(\lambda)}, \tag{D12}$$

which gives

$$\begin{aligned}&\int P(i|\lambda,x)P(k|\lambda,\mathcal{M}_j)\mu^*(\lambda)\,d\lambda \\ &= \int \mu_i(\lambda)p_i^{(x)}P(k|\lambda,\mathcal{M}_j) = p_i^{(x)}P(k|\mathcal{P}_i,\mathcal{M}_j) \\ &= P(i,k|x,j)\end{aligned} \tag{D13}$$

as required. ■

The "only if" part does not require the assumption that each i appears in only one support. Hence, in any contextuality scenario, one way to show the impossibility of a preparation noncontextual model is to show that Eq. (D8) violates a Bell inequality. The special thing about scenarios where each i appears in only one support is that this technique is powerful enough to capture *every* failure of preparation noncontextuality.

A generalization of Sec. VI is left open.

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[3] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[4] S. Kochen and E. Specker, Indiana Univ. Math. J. **17**, 59 (1968).
[5] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).
[6] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Phys. Rev. Lett. **102**, 010401 (2009).
[7] A. A. Klyachko, M. A. Can, S. Binicioğlu, and A. S. Shumovsky, Phys. Rev. Lett. **101**, 020403 (2008).
[8] G. Kirchmair, F. Zahringer, R. Gerritsma, M. Kleinmann, O. Guhne, A. Cabello, R. Blatt, and C. F. Roos, Nature (London) **460**, 494 (2009).
[9] O. Gühne, M. Kleinmann, A. Cabello, J.-A. Larsson, G. Kirchmair, F. Zähringer, R. Gerritsma, and C. F. Roos, Phys. Rev. A **81**, 022121 (2010).
[10] R. Lapkiewicz, P. Li, C. Schaeff, N. K. Langford, S. Ramelow, M. Wiesniak, and A. Zeilinger, Nature (London) **474**, 490 (2011).
[11] A. Winter, J. Phys. A: Math. Theor. **47**, 424031 (2014).
[12] J. V. Kujala, E. N. Dzhafarov, and J.-Å. Larsson, Phys. Rev. Lett. **115**, 150401 (2015).
[13] C. J. Wood and R. W. Spekkens, New J. Phys. **17**, 033002 (2015).
[14] L. Hardy, arXiv:quant-ph/0101012.
[15] V. Bužek, R. Derka, G. Adam, and P. Knight, Ann. Phys. **266**, 454 (1998).
[16] J. Barrett, unpublished note received in 2010.
[17] M. S. Leifer and O. J. E. Maroney, Phys. Rev. Lett. **110**, 120401 (2013).
[18] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[19] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[20] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevA.98.022112 for a mathematica notebook for verifying the proof of Lemma 1, and a PDF file showing the output thereof.
[21] B. S. Cirel'son, Lett. Math. Phys. **4**, 93 (1980).
[22] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, Nat. Commun. **7**, 11780 (2016).
[23] R. W. Spekkens, Phys. Rev. A **75**, 032110 (2007).
[24] L. Diósi, Phys. Rev. A **40**, 1165 (1989).
[25] R. Penrose, Gen. Relat. Gravit. **28**, 581 (1996).
[26] J. Barrett, Phys. Rev. A **75**, 032304 (2007).

[27] M. D. Mazurek, M. F. Pusey, K. J. Resch, and R. W. Spekkens, arXiv:1710.05948.

[28] A. Cabello, M. Gu, O. Gühne, J.-A. Larsson, and K. Wiesner, Phys. Rev. A **94**, 052127 (2016).

[29] R. W. Spekkens, Found. Phys. **44**, 1125 (2014).

[30] E. Schrödinger, Proc. Camb. Philos. Soc. **32**, 446 (1936).

[31] L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A **183**, 14 (1993).