Contextuality without incompatibility – Supplemental Material

John H. Selby,[1,*] David Schmid,[1,2,3] Elie Wolfe,[2] Ana Belén Sainz,[1] Ravi Kunjwal,[4] and Robert W. Spekkens[2]

[1] *International Centre for Theory of Quantum Technologies, University of Gdańsk, 80-309 Gdańsk, Poland*
[2] *Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario Canada N2L 2Y5*
[3] *Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[4] *Centre for Quantum Information and Communication, Ecole polytechnique de Bruxelles, CP 165, Université libre de Bruxelles, 1050 Brussels, Belgium*

PROOF OF THEOREM 1

We prove Theorem 1 by demonstrating that the flag-convexified scenario (introduced in the main text) admits of a non-contextual ontological model iff the original scenario does so. Specifically, this requires proving that there exist $P(\lambda|s)$ and $P(y|t,\lambda)$ satisfying the constraints of Eq. (3) from the main text and which reproduce the quantum statistics in the original scenario, i.e.,

$$P^{(\mathrm{q})}(y|s,t)=\sum_\lambda P(y|t,\lambda)P(\lambda|s) \tag{1}$$

if and only if there exist $\widetilde{P}(\lambda|s)$ and $\widetilde{P}(y,t|\lambda)$ satisfying the constraints of Eq. (9) from the main text and reproducing the quantum statistics in the flag-convexified scenario, i.e.,

$$\widetilde{P}^{(\mathrm{q})}(y,t|s)=\sum_\lambda \widetilde{P}(y,t|\lambda)\widetilde{P}(\lambda|s). \tag{2}$$

It then straightforwardly follows that every no-go theorem for noncontextuality in a prepare-measure scenario involving incompatibility can be transformed (via flag-convexification) into a no-go theorem involving a single measurement, that is, into a no-go theorem without incompatibility.

We begin with the forward implication. It suffices to take

$$\widetilde{P}(\lambda|s):=P(\lambda|s), \qquad \widetilde{P}(y,t|\lambda):=\frac{1}{|T|}P(y|t,\lambda). \tag{3}$$

It is straightforward to check that $\widetilde{P}(y,t|\lambda)$ is a valid conditional probability distribution. From the fact that $P(\lambda|s)$ and $P(y|t,\lambda)$ satisfy the noncontextuality constraints in Eq. (3) from the main text, it follows immediately that $\widetilde{P}(\lambda|s)$ and $\widetilde{P}(y,\,t|\lambda)$ satisfy the noncontextuality constraints in Eq. (9) from the main text. Eq. (2) then follows from Eq. (1) by making use of Eqs. (6) from the main text and (3), i.e.,

$$\begin{aligned}\widetilde{P}^{(\mathrm{q})}(y,t|s)&=\frac{1}{|T|}P^{(\mathrm{q})}(y|s,t) &(4)\\ &=\frac{1}{|T|}\sum_\lambda P(y|t,\lambda)P(\lambda|s) &(5)\\ &=\sum_\lambda \widetilde{P}(y|t,\lambda)\widetilde{P}(\lambda|s). &(6)\end{aligned}$$

The reverse implication follows by an analogous argument, but where one defines $P(\lambda|s):=\widetilde{P}(\lambda|s)$ and $P(y|t,\lambda):=|T|\widetilde{P}(y,t|\lambda)$.

PROOF OF NONCONTEXTUALITY INEQUALITY

We seek to put bounds on the correlations in the prepare-measure experiment. These are given in the text in terms of $P(y|s)$, where y runs over the five outcomes of the measurement and s runs over the five preparations. In an ontological model, we have

$$P(y|s)=\sum_\lambda P(y|\lambda)P(\lambda|s). \tag{7}$$

But by implementing a Bayesian inversion between λ and s, this can also be expressed as

$$P(y|s)=\sum_\lambda P(y|\lambda)P(s|\lambda)P(\lambda)P(s)^{-1}. \tag{8}$$

We here assume a uniform prior over s, i.e., $P(s)=1/5$. In this case, whatever linear dependences hold among the $P(\lambda|s)$ for a given λ, the same linear dependences hold among the $P(s|\lambda)$ for a given λ. It will be convenient below to focus on the joint distribution $P(y,s)$, which takes the form

$$P(y,s)=\sum_\lambda P(y|\lambda)P(s|\lambda)P(\lambda). \tag{9}$$

For a uniform prior, this is related to $P(y|s)$ by a constant factor.

We now derive the constraints on $P(y,s)$ that are implied by the assumption of noncontextuality and the operational equivalences that hold among the states and among the effects.

Recall that the assumption of noncontextuality means that operational equivalences of the form of Eq. (2) in the main text imply corresponding constraints on the representations of states and effects in the ontological model, namely, Eqs. (3) from the main text.

The particular form of operational equivalences appearing in the pentagon example are

$$\begin{aligned}\rho_0-q\rho_1+q\rho_2-\rho_3&=0,\\ \rho_1-q\rho_2+q\rho_3-\rho_4&=0,\end{aligned} \tag{10}$$

where $q = 2\cos(\pi/5) = (1+\sqrt{5})/2$ (the golden ratio), and

$$\begin{aligned} E_0 - qE_1 + qE_2 - E_3 &= 0, \\ E_1 - qE_2 + qE_3 - E_4 &= 0. \end{aligned} \tag{11}$$

Given the five-fold symmetry of the set of states, the linear dependence relations described by the pair of equations in Eq. (10) can be equivalently expressed by the image of this pair of equations under any cyclic permutation of the five states. Similarly, the five-fold symmetry of the set of effects implies that the linear dependence relations described by the pair of equations in Eq. (11) are equivalent to those obtained by cyclic permutations of the effects.

Note that there is an intuitive justification of the operational equivalence in the first equation of Eq. (10): the ensemble consisting of pure states ρ_0 and ρ_2, drawn with probabilities $1/(q+1)$ and $q/(q+1)$ respectively, yields the same density operator as the ensemble consisting of pure states ρ_1 and ρ_3, drawn with probabilities $q/(q+1)$ and $1/(q+1)$, respectively:

$$\frac{1}{q+1}\rho_0 + \frac{q}{q+1}\rho_2 = \frac{q}{q+1}\rho_1 + \frac{1}{q+1}\rho_3. \tag{12}$$

This equality is easily verified by considering the geometry of the five states in the Bloch sphere (Fig. 1). The weight of ρ_0 in the mixed state ρ (indicated in Fig. 1) is proportional to the Euclidean distance between ρ and ρ_2, while the weight of ρ_2 is proportional to the Euclidean distance between ρ and ρ_0. The ratio of these two Euclidean distances is $1/q$ where q is the golden ratio, since this is one of the ways in which the golden ratio appears in the geometry of the pentagon. It follows that the ratio of the weight of ρ_0 to the weight of ρ_2 in the mixed state ρ is also $1/q$. Requiring that these weights sum to unity then implies that they must be $1/(q+1)$ and $q/(q+1)$ respectively. The fact that ρ_3 and ρ_1 respectively have the same Euclidean distances to ρ as do ρ_0 and ρ_2 implies that ρ_3 and ρ_1 also appear with weights $1/(q+1)$ and $q/(q+1)$ in a convex decomposition of ρ. Equating the two convex decompositions of ρ yields Eq. (12).

The second operational equivalence relation in Eq. (10) is justified similarly, and the symmetry of these relations under cyclic permutations of the states is also evident from the geometry. A similar analysis holds for the operational equivalence relations in Eq. (11).

Generalized noncontextuality is the assumption that for each operational equivalence among states or among effects, one must posit a corresponding constraint on the ontological representation of these states and effects. Thus, in the concrete example considered here, for each λ, the $P(\lambda|s)$ satisfy linear dependence conditions parallel to those of Eq. (10). Furthermore, as noted above, for each λ, the Bayesian inverses of these conditionals, $P(s|\lambda)$, must satisfy the same linear dependence conditions as the $P(\lambda|s)$, together with normalization. Consequently, we have the constraints:

$$\begin{aligned} &P(s=0|\lambda) - qP(s=1|\lambda) + qP(s=2|\lambda) - P(s=3|\lambda) = 0, \\ &P(s=1|\lambda) - qP(s=2|\lambda) + qP(s=3|\lambda) - P(s=4|\lambda) = 0, \\ &\sum_s P(s|\lambda) = 1. \end{aligned} \tag{13}$$

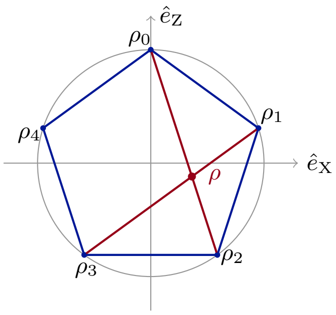

FIG. 1. Understanding the operational equivalences holding among the states in terms of mixtures of states.

For the effects, generalized noncontextuality implies that the $P(y|\lambda)$ must satisfy linear dependence conditions parallel to those of Eq. (11), together with normalization, so that

$$\begin{aligned} &P(y=0|\lambda) - qP(y=1|\lambda) + qP(y=2|\lambda) - P(y=3|\lambda) = 0, \\ &P(y=1|\lambda) - qP(y=2|\lambda) + qP(y=3|\lambda) - P(y=4|\lambda) = 0. \\ &\sum_y P(y|\lambda) = 1. \end{aligned} \tag{14}$$

The constraints in Eq. (13) must hold for every ontic state λ. If a probability distribution on s, i.e., $(P(s=0), P(s=1), P(s=2), P(s=3), P(s=4))$, satisfies these constraints, then it is termed a *noncontextual assignment* to s. In a noncontextual model, every ontic state λ makes such an assignment, which describes what one can retrodict about s given λ.

Similarly, if a probability distribution on y, i.e., $(P(y=0), P(y=1), P(y=2), P(y=3), P(y=4))$, satisfies the constraints in Eq. (14), then it is termed a *noncontextual assignment* to y. Again, every ontic state λ makes such an assignment in a noncontextual model, describing what one can *predict* about y given λ.

It is useful to consider the set of all possible noncontextual assignments to s. By the linearity of the constraints, any convex mixture of a pair of noncontextual assignments is also a valid noncontextual assignment, so the set is convex. In fact, it is a polytope contained within the 5-vertex simplex of all probability assignments over s. Given that the possible noncontextual assignments to y satisfy precisely the same constraints as the noncontextual assignments to s (since Eqs. (13) and (14) have the same form), these describe precisely the same polytope. We turn now to the description of this polytope.

Because of the invariance of the constraints under a cyclic permutation of the five components of the probability distribution, it follows that any cyclic permutation of the components of a noncontextual assignment yields a noncontextual assignment.

In fact, there are precisely five vertices of the polytope of noncontextual assignments, namely, the cyclic permutations of the following assignment:

$$\frac{1}{\sqrt{5}}(1,q^{-1},0,0,q^{-1}). \tag{15}$$

One can easily verify that this probability distribution satisfies Eq. (13) (or Eq. (14)) and satisfies normalization (it suffices to recall that the golden ratio q satisfies $q^{-1}=q-1$ and $1+2q^{-1}=\sqrt{5}$). Extremality can be verified by mathematical software.[1]

Let κ vary over the five extremal noncontextual assignments to y and let κ' vary over the five extremal noncontextual assignments to s. For every ontic state λ, the noncontextual assignment to y must be a convex mixture of the extremal ones: $P(y|\lambda)=\sum_\kappa P(y|\kappa)P(\kappa|\lambda)$ for some $P(\kappa|\lambda)$. Similarly, the noncontextual assignment to s by any ontic state λ must be a convex mixture of the extremal ones: $P(s|\lambda)=\sum_{\kappa'} P(s|\kappa')P(\kappa'|\lambda)$ for some $P(\kappa'|\lambda)$. It follows that Eq. (9) can be rewritten as:

$$P(y,s)=\sum_{\kappa,\kappa'} P(y|\kappa)P(s|\kappa')P(\kappa,\kappa'). \tag{16}$$

where $P(\kappa,\kappa') = \sum_\lambda P(\kappa|\lambda)P(\kappa'|\lambda)P(\lambda)$ describes an arbitrary distribution over κ and κ', and where the distributions $P(y|\kappa)$ and $P(s|\kappa')$ range over the cyclic permutations of the distribution in Eq. (15).

It follows that the distributions $P(y,s)$ with uniform marginal on s that are achievable in a noncontextual model with the specified operational equivalences are all and only those lying in the intersection of the hyperplane defining the uniform marginal condition ($\forall s:\sum_y P(y,s)=1/5$) and the hypercone consisting of the nonnegative hull of the twenty-five product distributions over y and s, $\{P(y|\kappa)P(s|\kappa')\}_{\kappa,\kappa'}$, one for each pair κ,κ'. The product distributions in this set are the extremal rays of the hypercone. Each can be expressed as a 5×5 matrix with rows indexed by s and columns by y. Specifically, using the form of the extremal assignments, Eq. (15), one infers that they are precisely those that can be obtained by cyclic permutations of the rows and columns of the matrix

$$M := \frac{1}{5}\begin{pmatrix} 1 & q^{-1} & 0 & 0 & q^{-1} \\ q^{-1} & q^{-2} & 0 & 0 & q^{-2} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ q^{-1} & q^{-2} & 0 & 0 & q^{-2} \end{pmatrix}. \tag{17}$$

We denote this twenty-five ray unbounded hypercone by $T_{\text{unbounded}}$. The intersection of $T_{\text{unbounded}}$ with the hyperplane defining the uniform marginal condition (i.e., the hyperplane where $\sum_y P(y,s)=1/5$ for all s) defines a bounded polytope which we will denote by T_{bounded}. T_{bounded} is the polytope of noncontextually realizable joint distributions $P(y,s)$ with uniform marginal on s, for the given operational equivalences. One could seek to identify the vertices and facet inequalities of T_{bounded}, but this is computationally nontrivial and it is not necessary for witnessing the failure of noncontextuality. For any distribution $P(y,s)$ with uniform marginal on s, violation of any facet inequality of the hypercone $T_{\text{unbounded}}$ is sufficient to witness the fact that this distribution does not lie within T_{bounded}. Therefore, we here make use of some of the facet inequalities of $T_{\text{unbounded}}$ as our noncontextuality inequalities.

The polytope of noncontextually realizable *conditional* distributions $P(y|s)$, denoted T'_{bounded}, is simply the image of T_{bounded} under the map that rescales every joint distribution $P(y,s)$ with uniform marginal on s by the factor 5, since for such distributions $P(y|s) = P(y,s)/P(s) = 5P(y,s)$. That is, T'_{bounded} is the intersection of $T_{\text{unbounded}}$ with the hyperplane where $\sum_y P(y|s)=1$ for all s. Since the hypercone $T_{\text{unbounded}}$ is invariant under nonnegative rescaling, we find that its facet inequalities constitute both noncontextuality inequalities for conditional distributions $P(y|s)$ or for joint distributions $P(y,s)$ with uniform marginals. [2]

[1] The presence of irrational coefficients complicates the task of identifying the extremal solutions to a set of linear constraints. For very small problems, such as the one here, however, it can be tackled by Mathematica®.

[2] It should be noted that we have conceptualized the prepare side of the prepare-measure scenario as a device wherein which of the five preparations is implemented is determined by the value of an *input* variable s. In other words, s has been conceptualized as a setting variable in a *multi-preparation*. However, one can just as well conceptualize the prepare side of the prepare-measure scenario as consisting of a device that samples a value of s at random, implements the associated preparation, and returns the value of s as an *output*, so that s is the outcome of a *source*. The process by which such a source is obtained from the multi-preparation is simply an instance of flag-convexification. The fact that one can evaluate noncontextual-realizability either in terms of the joint distribution $p(y,s)$ (for uniform marginal on s) rather than the conditional distribution $p(y|s)$ is an instance of the fact that one can always choose to flag-convexify the setting variable of the multi-preparation when assessing noncontextual realizability. While different sources that define the same average state may be compatible or incompatible, the fact that one can always flag-convexify and thereby conceptualize the experiment as involving a *single* source implies that no incompatibility among sources is required for proofs of contextuality. This point is considered in more detail in the companion article [1].

To identify facet inequalities of the hypercone $T_{\text{unbounded}}$ from its extremal rays, one can use linear programming algorithms.[3] Each facet inequality yields a noncontextuality inequality for this scenario. By applying standard such algorithms, we find that one such facet inequality is $\mathcal{C}\geq 0$, where, using the shorthand notation $p_{y|s}:=P(y|s)$, we can write $\mathcal{C}$ as

$$\mathcal{C}:=q(p_{1|0}+p_{1|2})+(q-1)p_{2|0}+p_{0|2}-(q+1)p_{1|1}. \quad (18)$$

Consequently, if a conditional distribution $P(y|s)$ violates the inequality $\mathcal{C}\geq 0$, the possibility of a noncontextual model for the given operational equivalences is ruled out. One can easily verify that the extremal assignments (i.e., permutations of the rows and columns of Eq. (17)) satisfy this inequality.

Note that since the hypercone $T_{\text{unbounded}}$ is invariant under uniform rescaling, by substituting $p_{y|s}\to 5p_{y|s}:=P(y,s)$ we obtain the equally valid inequality $\mathcal{C}'\geq 0$, where

$$\mathcal{C}':=q(p_{1,0}+p_{1,2})+(q-1)p_{2,0}+p_{0,2}-(q+1)p_{1,1} \quad (19)$$

which is satisfied by all $P(y,s)$ with uniform marginal $P(s)$ which admit a noncontextual model for the same operational equivalences.

One can generate forty-nine more facet inequalities of $T_{\text{unbounded}}$ by leveraging some of the symmetries of the problem. Since the set of extremal rays of $T_{\text{unbounded}}$ remains closed under the replacing $p_{y|s}$ with $p_{\pi(y)|\pi'(s)}$ where π and π' are cyclic permutations, or by replacing $p_{y|s}$ with $p_{s|y}$, it follows that any inequality generated by applying the inverse of such replacement to Eq. (12) from the main text would also constitute a facet-defining inequality of $T_{\text{unbounded}}$.

For the prepare-measure scenario described in the main text and depicted in Fig. (2) in the main text, the quantumly realized data table has elements $p_{y|s}=\text{Tr}(\rho_s E_y)$. Computing these overlaps, the conditional distribution is described by the following 5×5 matrix:

$$M^{(q)}:=\frac{1}{10}\begin{pmatrix} 4 & 1+q & 2-q & 2-q & 1+q \\ 1+q & 4 & 1+q & 2-q & 2-q \\ 2-q & 1+q & 4 & 1+q & 2-q \\ 2-q & 2-q & 1+q & 4 & 1+q \\ 1+q & 2-q & 2-q & 1+q & 4 \end{pmatrix}. \quad (20)$$

One easily verifies that if one evaluates $\mathcal{C}$ in Eq. (12) from the main text for $M^{(q)}$, one obtains the value $\frac{q^2-4}{10}\approx -0.138$ reported in Eq. (14) of the main text, which is a violation of the noncontextuality inequality in Eq. (13) from the main text.

* john.h.selby@gmail.com

[1] J. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, "Accessible fragments of generalized probabilistic theories, cone equivalence, and applications to witnessing nonclassicality," (2021), arXiv:2112.04521.

[3] We discover an inequality by minimizing the inner product of a variable vector with some ray outside of the hypercone subject to the restriction that vector have nonnegative inner product with each of the hypercone's extremal rays. The inequality is seen to be facet-defining iff the vector space spanned with the subset of extremal rays which saturate the discovered inequality is exactly one dimension smaller than the vector space spanned by *all* the hypercone's extremal rays.