

Incompatibility in restricted operational theories: connecting contextuality and steering

Martin Plávala* 

Naturwissenschaftlich-Technische Fakultät, Universität Siegen, 57068 Siegen, Germany

E-mail: martin.plavala@uni-siegen.de

Received 21 December 2021, revised 1 March 2022

Accepted for publication 4 March 2022

Published 25 March 2022



Abstract

We investigate the connection between steering and contextuality in general probabilistic theories. We show that for a class of bipartite states the steerability of the state by given set of measurements is equivalent to non-existence of preparation noncontextual hidden variable model for certain restricted theory constructed from the given state and measurements. The connection between steering and contextuality is provided by the concept of incompatibility in restricted theories, which we also investigate.

Keywords: steering, contextuality, general probabilistic theories, incompatibility, quantum foundations

(Some figures may appear in colour only in the online journal)

1. Introduction

One potential solution to the lack of intuitive interpretations of quantum theory is to assume that quantum theory is not fundamental, but that there is a hidden classical theory which can explain any prediction made by quantum theory. Such hidden classical theory is often called hidden variable model. Existence of hidden variable models was investigated in the past: it was shown by John Bell that local hidden variable models are incompatible with predictions of quantum theory [1] and it was shown by Kochen and Specker [2] that noncontextual hidden variable models are also incompatible with the predictions of quantum theory.

* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Here noncontextual means that the probability of observing a measurement outcome in the hidden variable theory does not depend on other possible measurement outcomes. For a recent review of noncontextuality in quantum theory see [3].

One possible operational version of contextuality was recently defined and discussed in [4]. By operational we mean that instead of investigating the existence of noncontextual description of quantum theory, we investigate the existence of noncontextual description of some convex operational theory. The class of convex operational theories we have in mind are called general probabilistic theories, or GPTs for short [5, 6]; they are large class of operational theories including both classical and quantum theories, but also theories that contain PR boxes [7, 8] or postquantum steering [9]. Noncontextuality was investigated in GPTs before [10–16]. It is clear that some GPTs must have noncontextual hidden variable model; trivial example is any classical theory which is itself its own hidden variable model. Using the language of GPTs we can ask which theories apart from classical do allow given form of noncontextual hidden variable model. This is an important question as if it turns out that almost no theories allow for noncontextual hidden variable model, then one can interpret the result of Kochen and Specker as a mathematical statement about a class of non-classical theories, rather than as a result about quantum theory itself.

Incompatibility is one of the fundamental non-classical features of quantum theory; we say that two measurements are incompatible if they cannot be measured jointly. Incompatibility was heavily studied within quantum theory [17, 18] but also in GPTs [19–21]; it is for example known that incompatibility is necessary for steering [22, 23] and for violations of Bell inequalities [24].

In this paper we formulate and investigate different forms of noncontextuality and their connection to the notion of incompatibility in restricted GPTs, we show that there is an inherent connection between incompatibility and contextuality in all GPTs. Our results on compatibility generalize the results obtained in [25] and they are complementary to [26] where it was shown that incompatibility is not necessary for the investigated type of contextuality. Our main result is that we apply the aforementioned connection to steering, where we show that steerability of certain class of states by given set of measurements directly corresponds to non-existence of preparation noncontextual hidden variable model for restricted theory constructed from the given state and measurements.

The paper is organized as follows: in section 2 we review the formalism of GPTs and in section 3 we introduce the three different types of noncontextual hidden variable models in GPTs. In section 4 we introduce two different types of incompatibility in restricted GPTs and we prove that there is a connection between incompatibility and non-existence of preparation noncontextual hidden variable. Finally in section 5 we put all of the concepts together and we prove the connection between steerability of a state by given measurement and non-existence of preparation noncontextual hidden variable model.

2. Restricted GPTs

In this section we will introduce restricted general probabilistic theories (GPTs). This section is intended mainly to set the notation and introduce only the necessary concepts; for an exhaustive review of GPTs see [5, 6].

A state space K is a compact convex subset of a finite-dimensional real vector space V . Given a state space K we define the effect algebra $E(K)$ as the set of affine functions $f : K \rightarrow [0, 1]$. An important element of $E(K)$ is the constant function $1_K \in E(K)$ given as $1_K(\rho) = 1$ for all $\rho \in K$. The interpretation is that $E(K)$ corresponds to all mathematically well-defined

two-outcome measurements on K , i.e., for $f \in E(K)$ and $\rho \in K$ the number $f(\rho)$ is the probability of obtaining the ‘yes’ outcome of the corresponding two-outcome measurement, the probability of ‘no’ outcome is given by the normalization as $1 - f(\rho)$.

Under reasonable assumptions one can show that we can use the tensor product to describe states shared between two parties. For a given state spaces K_A and K_B let K_{AB} be the bipartite state space that contains all possible states that the two parties can share, then

$$K_{AB} \subset \text{span}(\{\rho_A \otimes \rho_B : \rho_A \in K_A, \rho_B \in K_B\}). \quad (1)$$

It is natural to assume that K_{AB} contains all separable states, i.e., that we have $\rho_A \otimes \rho_B \in K_{AB}$ for all $\rho_A \in K_A$ and $\rho_B \in K_B$. This leads to the definition of the minimal tensor product

$$K_A \dot{\otimes} K_B = \text{conv}(\{\rho_A \otimes \rho_B : \rho_A \in K_A, \rho_B \in K_B\}) \quad (2)$$

and we then require $K_A \dot{\otimes} K_B \subset K_{AB}$. We will also assume that K_{AB} allows for all separable measurements, i.e., that for all $f_A \in E(K_A)$ and $f_B \in E(K_B)$ we have $(f_A \otimes f_B)(\rho_{AB}) \geq 0$. This leads to the definition of the maximal tensor product

$$\begin{aligned} K_A \hat{\otimes} K_B = \{ \psi \in \text{span}(K_A \dot{\otimes} K_B) : & (1_{K_A} \otimes 1_{K_B})(\psi) = 1, \\ & (f_A \otimes f_B)(\psi) \geq 0, \forall f_A \in E(K_A), \forall f_B \in E(K_B) \} \end{aligned} \quad (3)$$

and we then require $K_{AB} \subset K_A \hat{\otimes} K_B$. Using linearity we can define the experiment where Alice measures but Bob does not: let $\rho_{AB} \in K_{AB}$, then using (1) we have $\rho_{AB} = \sum_{i=1}^n \alpha_i \sigma_i^A \otimes \sigma_i^B$ where $\alpha_i \in \mathbb{R}$, $\sigma_i^A \in K_A$, $\sigma_i^B \in K_B$ for all $i \in \{1, \dots, n\}$. Let $f_A \in E(K_A)$, we then define $(f_A \otimes \text{id}_B)(\rho_{AB}) = \sum_{i=1}^n \alpha_i f_A(\sigma_i^A) \sigma_i^B$. Here id_B is the identity map on Bob’s system.

Example 1 (Classical theory). A useful example of a state space is the classical theory where the state space is a simplex S_n . Here a simplex is a convex hull of linearly independent pure states, that is, $S_n = \text{conv}(\{s_i\}_{i=1}^n)$ where s_i are linearly independent vectors. The effect algebra $E(S_n)$ is then generated by the dual basis $b_j \in E(S_n)$, which is given by $b_j(s_i) = \delta_{ij}$ where δ_{ij} is the Kronecker symbol. We then have $1_{S_n} = \sum_{j=1}^n b_j$. Note that we will also use the following notation: let I be any index set, e.g., $I = \{1, \dots, n\}$, then we will use S_I to denote the simplex generated by the states indexed by I , i.e., $S_I = \text{conv}(\{s_i\}_{i \in I})$.

Example 2 (Quantum theory). Quantum theory is another example of a GPT. Here the state space is the set of all density matrices, i.e., we have $K = \mathcal{D}(\mathcal{H})$, where $\mathcal{D}(\mathcal{H}) = \{\rho \in \mathcal{B}(\mathcal{H}) : \rho \geq 0, \text{Tr}(\rho) = 1\}$, $\mathcal{B}(\mathcal{H})$ is the vector space of Hermitian operators on the finite-dimensional complex Hilbert space \mathcal{H} and $A \geq 0$ means that A is positive semidefinite. We also have $E(\mathcal{D}(\mathcal{H})) = \mathcal{E}(\mathcal{H}) = \{M \in \mathcal{B}(\mathcal{H}) : 0 \leq M \leq 1\}$ up to an isomorphism; the value of the effect $M \in \mathcal{E}(\mathcal{H})$ on the state $\rho \in \mathcal{D}(\mathcal{H})$ is given by the Born rule $\text{Tr}(\rho M)$. Let $K_A = \mathcal{D}(\mathcal{H}_A)$ and $K_B = \mathcal{D}(\mathcal{H}_B)$, then the bipartite state space is given by the tensor product of the underlying Hilbert spaces, that is $K_{AB} = \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$.

Let K_A and K_B be state spaces, then a channel is an affine map $\Phi : K_A \rightarrow K_B$. A measurement is a channel where the target state space is a simplex, i.e., measurement is a channel $m : K_A \rightarrow S_n$. One can prove that every measurement m corresponds to a tuple of effects $f_i \in E(K)$, $i \in \{1, \dots, n\}$, such that $\sum_{i=1}^n f_i = 1_K$ and for $\rho \in K_A$ we have $m(\rho) = \sum_{i=1}^n f_i(\rho) s_i$. This is exactly

the same as saying that $p(i)$, the probability of observing the outcome i , is given as $p(i) = f_i(\rho)$. Finally, a classical post-processing is a channel $\nu : S_n \rightarrow S_k$ that maps a simplex to another simplex. Every classical post-processing corresponds to a right stochastic matrix ν_{ij} , i.e., $\nu_{ij} \geq 0$ and $\sum_{j=1}^k \nu_{ij} = 1$, the correspondence is given as $\nu(s_i) = \sum_{j=1}^n \nu_{ij} s_j$ where s_i is an extreme point of S_n .

Given a channel $\Phi : K_A \rightarrow K_B$ we can define the adjoint map $\Phi^* : E(K_B) \rightarrow E(K_A)$ as the unique channel such that $(\Phi^*(f))(\rho) = f(\Phi(\rho))$ holds for all $f \in E(K_B)$ and $\rho \in K_A$.

A restricted GPT is a theory in which not all mathematically well-defined states and measurements are allowed. Restricted GPTs arise in experimental scenarios, where we can usually prepare only a certain subset of all possible states and perform only a certain subset of all possible measurements. In order to describe the restricted theory we will use the pair (K, E) where K is a state space containing all preparable states and E is the effect algebra containing all allowed effects, i.e., all allowed two-outcome measurements. Note that the pair (K, E) does not completely describe the restricted theory, because we can also have independent higher-order restrictions on measurements with more than two outcomes [27]. In order to also take into account the restrictions on the set of more than two-outcome measurements, we will denote $\mathcal{M}(E)$ the set of all measurements allowed in the given restricted theory (K, E) . Going the other way, we can also define E as the smallest effect algebra containing effects corresponding to measurements m_x , where x is a some index; in this case we denote $E = \text{effect}(m_x)$. If the set $\{m_x\}$ is closed with respect to post-processing of measurements we get $\mathcal{M}(\text{effect}(m_x)) = \{m_x\}$.

For practical reason we will consider a theory with no restrictions as a special case of a general restricted theory. Moreover we will abuse the notation and we will use $(K, E(K))$ to denote the theory with no restrictions.

Let (K, E) be a restricted theory, then in order to make all probabilities positive we must require that $E \subset E(K)$, i.e., that every $f \in E$ is an affine function $f : K \rightarrow [0, 1]$. Following the consistency results derived in [27] we will also assume that E is convex and $0_K, 1_K \in E$, where $0_K(\rho) = 0$ for all $\rho \in K$. One also has to consider whether E contains enough measurements to distinguish any two states in K . Thus we arrive to the following definition:

Definition 1. Let (K, E) be a restricted theory. We say that (K, E) is tomographically complete if for every $\rho, \sigma \in K$, $\rho \neq \sigma$, there is $f \in E$ such that $f(\rho) \neq f(\sigma)$.

It is reasonable to expect that in most practical applications the restricted theories (K, E) will be tomographically complete. We will not assume topographical completeness because it is not needed for the validity of our results. Given the restricted theory (K, E) it is useful to define the state space of E as the set of all linear functionals ψ on $\text{span}(E)$ such that ψ is positive on E and normalized, $\psi(1_K) = 1$. We get

$$S(E) = \{\psi \in \text{span}(E)^* : \psi(f) \geq 0, \forall f \in E, \psi(1_K) = 1\}, \quad (4)$$

where V^* is the dual of the vector space V . For a theory with no restrictions $(K, E(K))$ we have $S(E(K)) = K$ up to an isomorphism [6]. If (K, E) is tomographically complete theory, then $K \subset S(E)$ up to an isomorphism. Finally we prove the following lemma which is going to be useful for working with tomographically incomplete restricted theories.

Lemma 1. Let (K, E) be a restricted theory, then there is a channel $\Phi_{(K,E)} : K \rightarrow S(E)$ such that $m(\rho) = m(\Phi_{(K,E)}(\rho))$ for all $\rho \in K$ and for all measurements $m \in \mathcal{M}(E)$.

Proof. Let $f \in E$ and $\rho \in K$. By construction of the restricted theory (K, E) there must be a way to evaluate f on ρ , i.e., the expression $f(\rho)$ must be well-defined. We will now employ a standard trick: rather than seeing f as the function and ρ as the input we will exchange their roles and we will interpret ρ as the function and f as the input. This is done using the evaluation map $\xi : \rho \rightarrow \xi_\rho$, where $\xi_\rho : E \rightarrow \mathbb{R}$ is a function defined as $\xi_\rho(f) = f(\rho)$. It is straightforward to prove that $\xi_\rho : E \rightarrow \mathbb{R}$ is linear and so $\xi_\rho \in \text{span}(E)^*$. Similarly we get $\xi_\rho(f) \geq 0$ for all $f \in E$, $\xi_\rho(1_K) = 1$ and that the evaluation map ξ is affine. It thus follows that $\xi_\rho \in S(E)$ and $\xi = \Phi_{(K,E)}$ is the channel we were looking for. \square

The interpretation of lemma 1 is that we can formally interpret every measure-and-prepare experiment in the restricted theory (K, E) as preparation of some state $\rho \in K$, applying the fixed channel $\Phi_{(K,E)}$ and measuring the final state $\Phi_{(K,E)}(\rho) \in S(E)$ using a measurement $m \in \mathcal{M}(E)$. Note that while $\rho \in K$ and $m \in \mathcal{M}(E)$ are given by the choice of the experiment, the channel $\Phi_{(K,E)}$ is always fixed by the structure of the restricted theory (K, E) . If (K, E) is tomographically complete, then $K \subset S(E)$ and one usually gets $\Phi_{(K,E)} = \text{id}$.

3. Hidden variable models

Given a restricted theory (K, E) one can ask whether hidden variable models with certain properties exist; the properties in question are for example locality, superdeterminism [28, 29], or noncontextuality. The hidden variable model is a classical theory that can model any preparation-measurement experiment performable in (K, E) . We will be interested in the following three different versions of noncontextual hidden variable models [10]: preparation noncontextual hidden variable models, measurement noncontextual hidden variable models and preparation and measurement noncontextual hidden variable models.

Consider the following scenario: an experimenter is given a description of the input data, according to the description the experimenter prepares the input state, measures and obtains statistics of the experiment. Let S_I be the simplex that describes the classical input and S_O the simplex that describes the classical output. The scenario in question can be described using the channels $P : S_I \rightarrow K$ and $m : K \rightarrow S_O$, where P is the preparation map and m is the measurement. Thus the whole experiment can be pictorially represented as

$$\begin{array}{c} S_I \\ \downarrow P \\ K \\ \downarrow m \\ S_O \end{array} \quad (5)$$

The upside of including the input data and S_I in our model is that the whole formulation is closer to being time-symmetric [30]. Note that S_I contains the full information about the input of the state preparation which allows us to distinguish scenarios that are later indistinguishable. For example, assume that $K = \mathcal{D}(\mathcal{H})$, $\dim(\mathcal{H}) = 2$, is a two-level quantum system and let $S_I = \text{conv}(\{s_0, s_1, s_+, s_-\})$. Let $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ and define the preparation map $P : S_I \rightarrow K$ as $P(s_0) = |0\rangle\langle 0|$, $P(s_1) = |1\rangle\langle 1|$, $P(s_+) = |+\rangle\langle +|$, $P(s_-) = |-\rangle\langle -|$. It then follows that $P\left(\frac{1}{2}s_0 + \frac{1}{2}s_1\right) = P\left(\frac{1}{2}s_+ + \frac{1}{2}s_-\right)$ and so $\frac{1}{2}s_0 + \frac{1}{2}s_1$ and $\frac{1}{2}s_+ + \frac{1}{2}s_-$ lead to preparing

indistinguishable states. But despite that $\frac{1}{2}s_0 + \frac{1}{2}s_1$ and $\frac{1}{2}s_+ + \frac{1}{2}s_-$ are perfectly distinguishable states of S_I . Hence at the level of S_I one can distinguish preparation inputs that are indistinguishable at the level of K .

A hidden variable model for the scenario described by P and m is given by a simplex S_Λ , preparation $\eta_P : S_I \rightarrow S_\Lambda$ and measurement $\mu_m : S_\Lambda \rightarrow S_O$ such that the statistics of both experiments are the same, i.e., for all $s_i \in S_I$ we have $(m \circ P)(s_i) = (\mu_m \circ \eta_P)(s_i)$, which can be written in terms of probability distributions as

$$p(a|P, m, i) = \sum_{\lambda \in \Lambda} p(a|m, \lambda)p(\lambda|P, i), \quad (6)$$

where $p(a|P, m, i)$ is the probability of obtaining an outcome a of measurement m when the input data is i and preparation P is used, $p(\lambda|P, i)$ is the probability of obtaining the value of the hidden variable λ when modeling preparation P with input data i and $p(a|m, \lambda)$ is the probability of obtaining the outcome a when the value of the hidden variable is λ and when modeling the measurement m . Here η_P and μ_m are essentially the maps $\eta_P : i \rightarrow p(\lambda|P, i)$ and $\mu_m : \lambda \rightarrow p(a|m, \lambda)$. Note that the preparation η_P can depend on P and the measurement μ_m can depend on m , one can capture these dependencies by defining the maps $\eta : P \mapsto \eta_P$ and $\mu : m \mapsto \mu_m$. Pictorially we get



where the double arrows denote supermaps, that is maps that map maps. The meaning of (7) is that both paths from S_I to S_O lead to the same result. It is straightforward to see that the hidden variable model corresponding to (7) always exists, for example one may construct such hidden variable model by calculating the outcomes of the experiment by hand, the hidden variable Λ then corresponds to all possible (finite) calculations one can write down. This is not surprising as so far we did not require that the hidden variable model satisfies any properties except that i and a do not depend on λ and it was already observed in [2] that the hidden variable model should be required to have some additional properties in order to not be trivial.

One can require the hidden variable model to be preparation noncontextual, meaning that for all preparations $P : S_I \rightarrow K$ and all input data $s_i \in S_I$ the prepared distribution of hidden variables depends only on $P(s_i)$. This means that for any $s, s' \in S_I$ and $P, P' : S_I \rightarrow K$ such that $P(s) = P'(s')$ we must have $\eta_P(s) = \eta_{P'}(s')$. It follows that one can define a map $\iota : K \rightarrow S_\Lambda$ as follows: let P be a preparation $P : S_I \rightarrow K$, let $s \in S_I$ and denote $P(s_i) = \rho$, then we define $\iota(\rho) = \eta_P(s)$. This is well-defined because if there is a different preparation $P' : S_I \rightarrow K$ and input data $s' \in S_I$ such that $P'(s') = \rho$, then $P'(s') = \rho = P(s)$ and so $\iota(\rho) = \eta_P(s) = \eta_{P'}(s')$. Moreover ι is affine map: let $\rho, \sigma \in K$ and let P be a preparation such that $P(s) = \rho$ and $P(s') = \sigma$, we can always construct such preparation as a preprocessing of the preparations of ρ and σ . Let $\lambda \in [0, 1]$, we then have $P(\lambda s + (1 - \lambda)s') = \lambda\rho + (1 - \lambda)\sigma$ and

$$\begin{aligned} \iota(\lambda\rho + (1 - \lambda)\sigma) &= \eta_P(\lambda s + (1 - \lambda)s') = \lambda\eta_P(s) + (1 - \lambda)\eta_P(s') \\ &= \lambda\iota(\rho) + (1 - \lambda)\iota(\sigma). \end{aligned} \quad (8)$$

Thus ι is a channel and we get the following definition:

Definition 2 (Preparation noncontextuality). A restricted theory (K, E) is preparation noncontextual if there is a map $\iota : K \rightarrow S_\Lambda$ such that for any experiment given by preparation $P : S_I \rightarrow K$ and measurement $m : K \rightarrow S_O$ we have

$$p(a|P, m, i) = \sum_{\lambda \in \Lambda} p(a|m, \lambda) b_\lambda((\iota \circ P)(s_i)), \quad (9)$$

where $b_\lambda \in E(S_\Lambda)$ is such that $b_\lambda(s_{\lambda'}) = \delta_{\lambda\lambda'}$. One can also denote $b_\lambda((\iota \circ P)(s_i)) = p(\lambda|P(s_i))$, explicitly showing that the distribution of hidden variable λ depends only on $P(s_i)$. (9) is the same as

$$\begin{array}{ccc} S_I & & \\ \downarrow P & & \\ K & \xrightarrow{\iota} & S_\Lambda \\ \downarrow m & \searrow \mu & \\ S_O & \swarrow \mu_m & \end{array} \quad (10)$$

Analogically one can require that the hidden variable model is measurement noncontextual, meaning that for all measurements $m : K \rightarrow S_O$ and all outcome labels $s_a \in S_O$ the distribution of the outcome data depends only on $f_a = m^*(b_a)$ and λ . This means that for any $b, b' \in E(S_O)$ and $m, m' : K \rightarrow S_O$ such that $m^*(b) = (m')^*(b')$ we must have $\mu_m^*(b) = \mu_{m'}^*(b')$. It follows that there must exist a map $\kappa : E \rightarrow E(S_\Lambda)$ such that if $f = m^*(b)$, then $\kappa(f) = \mu_m^*(b)$, moreover one can show that κ is affine in the same way as for ι . We thus get the definition:

Definition 3 (Measurement noncontextuality). A restricted theory (K, E) is measurement noncontextual if there is a map $\kappa^* : S_\Lambda \rightarrow S(E)$ such that for any experiment given by preparation $P : S_I \rightarrow K$ and measurement $m : K \rightarrow S_O$ we have

$$p(a|P, m, i) = \sum_{\lambda \in \Lambda} b_a((m \circ \kappa^*)(s_\lambda)) p(\lambda|P, i), \quad (11)$$

where one can denote $b_a((m \circ \kappa^*)(s_\lambda)) = p(a|m^*(b_a))$. Diagrammatic equivalent of (11) is

$$\begin{array}{ccc} S_I & & \\ \downarrow P & \searrow \eta & \\ K & & S_\Lambda \xrightarrow{\kappa^*} S(E) \\ \downarrow m & & \\ S_O & \swarrow m & \end{array} \quad (12)$$

The diagram (12) is quite different from (10), but one can make the diagrams similar by explicitly including the channel $\Phi_{(K,E)}$ given by lemma 1 and treating m as only measurement on $S(E)$.

One can also require both at the same time, i.e., one can require the hidden variable model to be both preparation noncontextual and measurement noncontextual. In this case we will say that (K, E) is simplex-embeddable, as already defined in [4].

Definition 4 (Simplex-embeddability). A restricted theory (K, E) is simplex-embeddable if there is a simplex S_Λ and channels $\iota : K \rightarrow S_\Lambda$ and $\kappa^* : S_\Lambda \rightarrow S(E)$ such that for all $\rho \in K$ and $f \in E$ we have $(\kappa(f))(\iota(\rho)) = f(\rho)$, which is the same as

$$\begin{array}{ccccc}
 & & S_I & & \\
 & & \downarrow P & & \\
 & K & \xrightarrow{\iota} & S_\Lambda & \xrightarrow{\kappa^*} S(E) \\
 & \downarrow m & & \nearrow m & \\
 & S_O & & &
 \end{array} \tag{13}$$

4. Two definitions of incompatibility in restricted operational theories

Before defining incompatibility of measurements in restricted theories, let's remind ourselves of the definition used in theories with no restrictions. Although we will define compatibility of only two measurements, it is straightforward to extend the notion to finite or infinite amount of measurements. Let $(K, E(K))$ be a theory satisfying no-restriction hypothesis and let m_1 and m_2 be two measurements corresponding to the sets of effects $f_i \in E(K)$ and $g_j \in E(K)$ respectively, where $i \in \{1, \dots, n_1\}$, $j \in \{1, \dots, n_2\}$ and $\sum_{i=1}^{n_1} f_i = \sum_{j=1}^{n_2} g_j = 1_K$. Then we say that m_1 and m_2 are compatible if there is a measurement m given by the set of effects $h_{ij} \in E(K)$, $i \in \{1, \dots, n_1\}$, $j \in \{1, \dots, n_2\}$ such that

$$f_i = \sum_{j=1}^{n_2} h_{ij}, \tag{14}$$

$$g_j = \sum_{i=1}^{n_1} h_{ij}. \tag{15}$$

Measurement m satisfying (14) and (15) is called the joint measurement. The definition of compatibility is equivalent to requiring that there is a measurement m given by the set of h_λ , $\lambda \in \Lambda$, where Λ is some arbitrary index set, and conditional probability distributions $p(i|1, \lambda)$ and $p(j|2, \lambda)$ such that $f_i = \sum_{\lambda \in \Lambda} p(i|1, \lambda)h_\lambda$, $g_j = \sum_{\lambda \in \Lambda} p(j|2, \lambda)h_\lambda$, see [19] for a proof.

Given a restricted theory (K, E) we have two options when defining compatibility of measurements: we either require that $h_{ij} \in E$ or we relax this to $h_{ij} \in E(K)$. One may be inclined to argue that $h_{ij} \in E$ is the correct option, as the joint measurement m given by the effects h_{ij} should be an allowed measurement, but, as we will see, the other option is also useful and quite interesting, since it implies that even though the joint measurement is not an allowed measurement, the available states from K cannot witness [31] the incompatibility of m_1 and m_2 . Thus we obtain the following two definitions:

Definition 5 (E-compatibility). Let (K, E) be a restricted theory and let m_1 and m_2 be measurements given by $f_i \in E$ and $g_j \in E$ respectively, where $i \in \{1, \dots, n_1\}$ and $j \in \{1, \dots, n_2\}$.

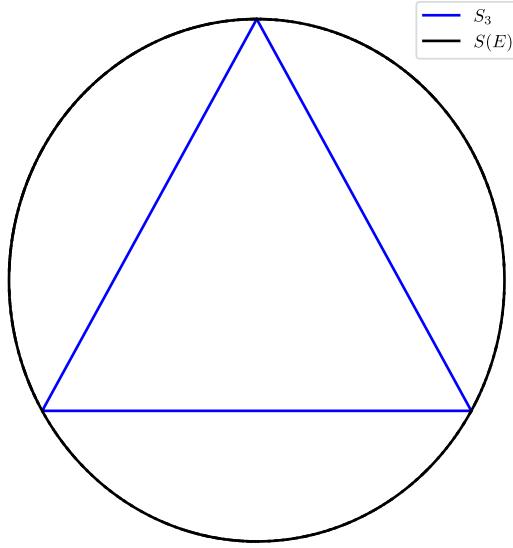


Figure 1. Embedding of the classical state space S_3 inside the circle $S(E)$ used in example 3.

We say that m_1 and m_2 are E -compatible if there are $h_{ij} \in E$, $i \in \{1, \dots, n_1\}$, $j \in \{1, \dots, n_2\}$, such that (14) and (15) hold. The measurement m will be called E -joint measurement of m_1 and m_2 .

Definition 6 ($E(K)$ -compatibility). Let (K, E) be a restricted theory and let m_1 and m_2 be measurements given by $f_i \in E$ and $g_j \in E$ respectively, where $i \in \{1, \dots, n_1\}$ and $j \in \{1, \dots, n_2\}$. We say that m_1 and m_2 are $E(K)$ -compatible if there are $h_{ij} \in E(K)$, $i \in \{1, \dots, n_1\}$, $j \in \{1, \dots, n_2\}$, such that (14) and (15) hold. The measurement m will be called $E(K)$ -joint measurement of m_1 and m_2 .

The following result is immediate:

Lemma 2. Let (K, E) be a restricted theory and let m_1 and m_2 be measurements. If m_1 and m_2 are E -compatible, then they are also $E(K)$ -compatible.

Proof. The proof follows from $E \subset E(K)$. □

It is straightforward to construct a restricted theory (K, E) and measurements m_1 and m_2 that are $E(K)$ -compatible, but E -incompatible.

Example 3 (Triangle-in-circle). Let $K = S_3$ be a triangle and let $S(E)$ be a circumscribed circle of K , see figure 1. Then $S(E)$ coincides with the state space of the real quantum theory and so there clearly are E -incompatible measurements m_1 and m_2 . But since $K = S_3$ is a simplex, it follows that all measurements are $E(K)$ -compatible.

In the following we will investigate what happens if all measurements in a restricted theory (K, E) are either E -compatible or $E(K)$ -compatible and we will compare these concepts to simplex-embeddability. We will take this path because, as we will see later, exactly these concepts become crucial when investigating the steerability of a bipartite state. Also note that we are not loosing too much generality by taking this path as when investigating the compatibility

of measurement m_1, \dots, m_N one can always take E to be the smallest effect algebra including all effects corresponding to the measurements m_1, \dots, m_N . The first result is an immediate application of known results on theories with no-restrictions.

Proposition 1. *Let (K, E) be a restricted theory. All measurements in (K, E) are E -compatible if and only if $S(E)$ is a simplex.*

Proof. If $S(E)$ is a simplex, then all measurements containing effects from E are compatible and so all measurements in (K, E) are E -compatible.

If all measurements are E -compatible, then it follows that all pairs of two-outcome measurements are E -compatible. Let m_1 and m_2 be a pair of two-outcome measurements, then their E -joint measurement m is a well-defined measurement on $S(E)$. Then $(S(E), E)$ is a theory with no restrictions in which all pairs of two-outcome measurements are compatible. It follows that $S(E)$ is a simplex [32, 33]. \square

The following result is similar to [25, theorem 1].

Proposition 2. *Let (K, E) be a restricted theory. All measurements in (K, E) are $E(K)$ -compatible if and only if there exists a preparation noncontextual hidden variable model for (K, E) .*

Proof. Let $\{m_x\} = \mathcal{M}(E)$, i.e., we index all measurements in $\mathcal{M}(E)$ by the index x , and denote $f_{a|x}$ the effects corresponding to m_x . Since m_x are $E(K)$ -compatible, it follows that there are $h_\lambda \in E(K)$ such that $f_{a|x} = \sum_{\lambda \in \Lambda} p(a|x, \lambda)h_\lambda$. Let us define $\iota : K \rightarrow S_\Lambda$ by $\iota(\rho) = \sum_{\lambda \in \Lambda} h_\lambda(\rho)s_\lambda$ and $\mu_{m_x} : S_\Lambda \rightarrow S_O$ by $b_a(\mu_{m_x}(s_\lambda)) = p(a|x, \lambda)$. We then have

$$p(a|P, m_x, i) = f_{a|x}(P(s_i)) = \sum_{\lambda \in \Lambda} p(a|x, \lambda)h_\lambda(P(s_i)) \quad (16)$$

$$= b_a((\mu_{m_x} \circ \iota)(P(s_i))). \quad (17)$$

Thus ι and $\mu : m_x \mapsto \mu_{m_x}$ constitute the preparation noncontextual hidden variable model.

Now assume that there is a preparation noncontextual hidden variable model for (K, E) . We then have that for every measurement $m : K \rightarrow S_O$ there are maps $\iota : K \rightarrow S_\Lambda$ and $\mu_m : S_\Lambda \rightarrow S_O$ such that $m = \mu_m \circ \iota$. Observe that ι is a measurement on K corresponding to effects $h_\lambda \in E(K)$ and μ_m is just a classical post-processing, so any measurement m in (K, E) can be obtained as a classical post-processing μ_m of the measurement ι . It follows that all measurements in (K, E) are $E(K)$ -compatible [17, subsection 5.2]. \square

Finally, we get to describe the hierarchy between E -compatibility, simplex-embeddability and $E(K)$ -compatibility.

Theorem 1. *Let (K, E) be a restricted theory. Then the following two implications hold:*

- (a) *If all measurements on (K, E) are E -compatible, then (K, E) is simplex-embeddable.*
- (b) *If (K, E) is simplex-embeddable, then all measurements on (K, E) are $E(K)$ -compatible.*

Proof. If all measurements in (K, E) are E -compatible then $S(E)$ is a simplex, see proposition 1. The main idea used to prove (a) is to take $S_\Lambda = S(E)$, $\kappa^* = \text{id}$ and $\iota = \Phi_{(K, E)}$, where $\Phi_{(K, E)}$ is the channel given by lemma 1. Let $\rho \in K$ and $f \in E$, we then have $\kappa(f) = f$ and $f(\Phi_{(K, E)}(\rho)) = f(\rho)$ follows from lemma 1.

Assume that (K, E) is simplex-embeddable and let m_x be a collection of measurements in (K, E) . Let $\rho \in K$, we then have

$$m(\rho) = (m \circ \kappa^* \circ \iota)(\rho) = (m \circ \kappa^*)(\iota(\rho)), \quad (18)$$

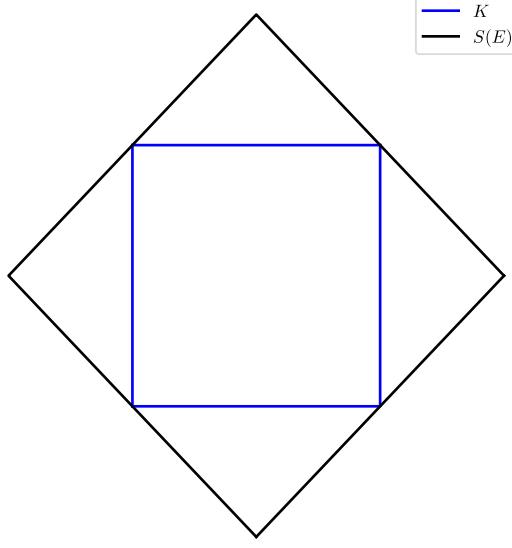


Figure 2. Embedding of the square state space K inside the larger square $S(E)$ used in example 4.

where $\iota : K \rightarrow S_\Lambda$ can be seen as a measurement with effects from $E(K)$ and $m \circ \kappa^* : S_\Lambda \rightarrow S_O$ can be seen as classical post-processing. Thus the measurements m_x are $E(K)$ -compatible [17, subsection 5.2]. \square

Another way to prove (b) is to observe that if (K, E) is simplex embeddable, then there also exists a preparation noncontextual model for (K, E) and we get the wanted result using proposition 2.

The following example proves that the inclusion in theorem 1(a) is strict.

Example 4 (Square-in-square). Example of a theory that is simplex-embeddable, but not E -compatible is the stabilizer rebit theory used in [4]. Let both K and $S(E)$ be squares, such that K is the convex hull of the midpoints of edges of $S(E)$, see figure 2. It is well-known that there are E -incompatible two-outcome measurements [34, 35], since $S(E)$ is a square and $E(S(E)) = E$. It was shown in [4] that (K, E) is simplex-embeddable.

Using the result of proposition 2 and theorem 1 we get the following corollaries about theories with no restrictions.

Corollary 1. *Let $(K, E(K))$ be a theory satisfying the no-restriction hypothesis and assume that $(K, E(K))$ is simplex-embeddable. Then K is a simplex.*

Proof. If $(K, E(K))$ is simplex-embeddable, then according to theorem 1 all measurements in $(K, E(K))$ are $E(K)$ -compatible and so K must be a simplex [32, 33]. \square

Corollary 2. *Let $(K, E(K))$ be a theory satisfying the no-restriction hypothesis and assume that there is preparation noncontextual hidden variable model for $(K, E(K))$. Then K is a simplex.*

Proof. The result immediately follows from proposition 2: the existence of preparation noncontextual hidden variable model for $(K, E(K))$ implies that all measurements are $E(K)$ -compatible, which implies that K is a simplex [32, 33]. \square

5. Connecting steering and noncontextuality

In this section we will apply the concepts of $E(K)$ -compatibility to the problem of a steerability of a bipartite state $\rho_{AB} \in K_{AB}$. In order to do so, we first provide a short review on steering in GPTs and then we show how to construct the restricted theory (K, E) corresponding to a given steering scenario. For an in-depth review on steering in quantum theory see [36].

Let $\rho_{AB} \in K_{AB}$ be a bipartite state shared between Alice and Bob and let m_x be a set of measurements that Alice can choose from, each given by a set of effects $f_{a|x} \in E(K_A)$, such that $\sum_a f_{a|x} = 1_{K_A}$. After measuring m_x and obtaining the outcome a , Bob is left with the post-measurement state $\sigma_{a|x}$ given as $\sigma_{a|x} = (f_{a|x} \otimes \text{id}_B)(\rho_{AB})$. Note that if $K_A = \mathcal{D}(\mathcal{H}_A)$, $K_B = \mathcal{D}(\mathcal{H}_B)$ and $K_{AB} = \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, then the measurements m_x are given by quantum effects $M_{a|x} \in \mathcal{E}(\mathcal{H}_A)$ and $\sigma_{a|x} = \text{Tr}_A((M_{a|x} \otimes 1)\rho_{AB})$. In general $\sigma_{a|x} \notin K_B$ since they are subnormalized as $1_K(\sigma_{a|x})$ is the probability of Alice obtaining the outcome a when measuring x . The set of states $\{\sigma_{a|x}\}$ is called an assemblage and we say that the assemblage $\{\sigma_{a|x}\}$ is not steerable if there exists a local hidden state model [37], that is if there is set of states $\{\sigma_\lambda\} \subset K_B$, $\lambda \in \Lambda$, and probability distributions $p(\lambda)$ and $p(a|x, \lambda)$ such that $\sigma_{a|x} = \sum_{\lambda \in \Lambda} p(\lambda)p(a|x, \lambda)\sigma_\lambda$. It is known that a necessary conditions for steering are that the state ρ_{AB} is not separable and that the measurements m_x are incompatible. It is of special interest to decide whether a given state $\rho_{AB} \in K_{AB}$ is steerable by some set of measurements m_x , or whether ρ_{AB} is steerable by all possible measurements. We will not restrict the cardinality of the set m_x , hence one can take m_x to include all possible measurements.

For the purposes of the following proof we will introduce the notion of full-dimensionality of ρ_{AB} .

Definition 7. Let $\rho_{AB} \in K_{AB}$. We say that ρ_{AB} is K_B -full-dimensional if

$$\text{span}(K_B) = \text{span}(\{(f \otimes \text{id}_B)(\rho_{AB}) : f \in E(K_\rho)\}), \quad (19)$$

where K_ρ is given as in (21).

The following is an immediate result.

Lemma 3. Let $\rho_{AB} \in K_{AB}$. ρ_{AB} is K_B -full-dimensional if

$$\text{span}(K_B) = \text{span}(\{(f \otimes \text{id}_B)(\rho_{AB}) : f \in E(K_A)\}). \quad (20)$$

Proof. The result follows from $E(K_A) \subset E(K_\rho)$, which follows from $K_\rho \subset K_A$. \square

It is easy to see that in quantum theory if the dimensions of the local Hilbert spaces are the same, then every state with full Schmidt rank is both K_A -full-dimensional and K_B -full-dimensional. Let us now build the restricted theory (K, E) corresponding to the steering scenario. Let $\rho_{AB} \in K_{AB}$ and let $g \in E(K_B)$, then $\frac{(\text{id}_A \otimes g)(\rho_{AB})}{(1_{K_A} \otimes g)(\rho_{AB})} \in K_A$ and so we define

$$K_\rho = \left\{ \frac{(\text{id}_A \otimes g)(\rho_{AB})}{(1_{K_A} \otimes g)(\rho_{AB})} : g \in E(K_B) \right\}. \quad (21)$$

Let $\{m_x\}$ be the set of measurements that we want to use for steering. Note that we can without loss of generality assume that $\{m_x\}$ is closed with respect to post-processing of measurements, as if the set $\{m_x\}$ does not steer a state ρ_{AB} , then we can also construct the local hidden state model for every measurement that is a post-processing of m_x . We will define E_m to be the smallest effect algebra containing all of the effects corresponding to m_x , i.e.,

$$E_m = \text{effect}(m_x). \quad (22)$$

It follows that if m_x includes all possible measurements, then $E_m = E(K_A)$. The theory (K_ρ, E_m) does not have to be tomographically complete, take for example $\rho_{AB} = \rho_A \otimes \rho_B$, then K_ρ contains only a single state. In some cases one can get rid of the tomographical incompleteness but there are also pathological cases, both of these cases will be addressed later on.

We can now introduce the result that connects noncontextuality and steering.

Theorem 2. *Let $\rho_{AB} \in K_{AB}$ be K_B -full-dimensional state and let (K_ρ, E_m) be the restricted theory as given by (21) and (22). Then ρ_{AB} is not steerable by m_x if and only if there exists a preparation noncontextual hidden variable model for (K_ρ, E_m) .*

Proof. Assume that a preparation noncontextual hidden variable model for (K_ρ, E_m) exists. According to theorem 1 the measurements given by $f_{a|x}$ are $E(K_\rho)$ -compatible, and so there is a measurement given by effects $h_\lambda \in E(K_\rho)$, $\lambda \in \Lambda$, such that $f_{a|x} = \sum_{\lambda \in \Lambda} p(a|x, \lambda)h_\lambda$. Let $\omega \in K_\rho$, then there is $g \in E(K_B)$ such that $\omega = \frac{(\text{id}_A \otimes g)(\rho_{AB})}{(\text{I}_{K_A} \otimes g)(\rho_{AB})}$ and we get

$$f_{a|x}(\omega) = \sum_{\lambda \in \Lambda} p(a|x, \lambda)h_\lambda(\omega), \quad (23)$$

$$(f_{a|x} \otimes g)(\rho_{AB}) = \sum_{\lambda \in \Lambda} p(a|x, \lambda)(h_\lambda \otimes g)(\rho_{AB}), \quad (24)$$

for all $g \in E(K_B)$, which implies that

$$(f_{a|x} \otimes \text{id}_B)(\rho_{AB}) = \sum_{\lambda \in \Lambda} p(a|x, \lambda)(h_\lambda \otimes \text{id}_B)(\rho_{AB}). \quad (25)$$

Denoting $(h_\lambda \otimes \text{id}_B)(\rho_{AB}) = p(\lambda)\sigma_\lambda$ we get

$$(f_{a|x} \otimes \text{id}_B)(\rho_{AB}) = \sum_{\lambda \in \Lambda} p(\lambda)p(a|x, \lambda)\sigma_\lambda \quad (26)$$

which is the local hidden state model for ρ_{AB} and $f_{a|x}$. Hence ρ_{AB} is not steerable by $f_{a|x}$.

Now assume that ρ_{AB} is not steerable by $f_{a|x}$ and so there always is a local hidden state model given as

$$(f_{a|x} \otimes \text{id}_B)(\rho_{AB}) = \sum_{\lambda \in \Lambda} p(\lambda)p(a|x, \lambda)\sigma_\lambda, \quad (27)$$

where $\sigma_\lambda \in K_B$ for all $\lambda \in \Lambda$. Since ρ_{AB} is K_B -full-dimensional it follows that for every $\sigma_\lambda \in K_B$ there is some $h_\lambda \in \text{span}(E(K_\rho))$ such that

$$p(\lambda)\sigma_\lambda = (h_\lambda \otimes \text{id}_B)(\rho_{AB}) \quad (28)$$

and so we have

$$(f_{a|x} \otimes \text{id}_B)(\rho_{AB}) = \sum_{\lambda \in \Lambda} p(a|x, \lambda)(h_\lambda \otimes \text{id}_B)(\rho_{AB}). \quad (29)$$

Let $g \in E(K_B)$, then we have

$$(f_{a|x} \otimes g)(\rho_{AB}) = \sum_{\lambda \in \Lambda} p(a|x, \lambda)(h_\lambda \otimes g)(\rho_{AB}) \quad (30)$$

which implies

$$f_{a|x}(\omega) = \sum_{\lambda \in \Lambda} p(a|x, \lambda) h_\lambda(\omega) \quad (31)$$

for all $\omega \in K_\rho$. Thus we have

$$f_{a|x} = \sum_{\lambda \in \Lambda} p(a|x, \lambda) h_\lambda. \quad (32)$$

This already is close to the wanted result, we only need to prove that h_λ are positive functions and that $\sum_{\lambda \in \Lambda} h_\lambda = 1_{K_\rho}$.

So assume that there is some $\omega \in K_\rho$ and $\lambda \in \Lambda$ such that $h_\lambda(\omega) < 0$. Then there is some $g \in E(K_B)$ such that $\omega = \frac{(\text{id}_A \otimes g)(\rho_{AB})}{(1_{K_A} \otimes g)(\rho_{AB})}$ and we get

$$0 > h_\lambda(\omega) = \frac{(h_\lambda \otimes g)(\rho_{AB})}{(1_{K_A} \otimes g)(\rho_{AB})} = \frac{p(\lambda)}{(1_{K_A} \otimes g)(\rho_{AB})} g(\sigma_\lambda). \quad (33)$$

Since $\frac{p(\lambda)}{(1_{K_A} \otimes g)(\rho_{AB})} \geq 0$ it follows that $g(\sigma_\lambda) < 0$, but this is a contradiction because $g \in E(K_B)$ and $\sigma_\lambda \in K_B$. To show that $\sum_{\lambda \in \Lambda} h_\lambda = 1_{K_\rho}$ observe that

$$(1_{K_A} \otimes \text{id}_B)(\rho_{AB}) = \sum_a (f_{a|x} \otimes \text{id}_B)(\rho_{AB}) = \sum_{\lambda \in \Lambda} \sum_a p(\lambda) p(a|x, \lambda) \sigma_\lambda \quad (34)$$

$$= \sum_{\lambda \in \Lambda} p(\lambda) \sigma_\lambda = \sum_{\lambda \in \Lambda} (h_\lambda \otimes \text{id}_B)(\rho_{AB}). \quad (35)$$

Let $g \in E(K_B)$, then we have $(1_{K_A} \otimes g)(\rho_{AB}) = \sum_{\lambda \in \Lambda} (h_\lambda \otimes g)(\rho_{AB})$ which implies that for all $\omega \in K_\rho$ we get $1_{K_A}(\omega) = \sum_{\lambda \in \Lambda} h_\lambda(\omega)$ and the result follows. \square

Let us first discuss the requirement that ρ_{AB} is K_B -full-dimensional. If ρ_{AB} is not K_B -full-dimensional, then not all is lost, as in some cases one can restrict K_B to $K_B \cap \text{span}(\{(f \otimes \text{id}_B)(\rho_{AB}) : f \in E(K_\rho)\})$. Denote $J = \text{span}(\{(f \otimes \text{id}_B)(\rho_{AB}) : f \in E(K_\rho)\})$ and assume that there is a channel $\Phi : K_B \rightarrow K_B$ such that for all $x \in K_B$ we have $\Phi(x) \in J$ and $\Phi(v) = v$ for all $v \in J$. Assume there is a local hidden state model $(f_{a|x} \otimes \text{id}_B)(\rho_{AB}) = \sum_{\lambda \in \Lambda} p(\lambda) p(a|x, \lambda) \sigma_\lambda$. Since $(f_{a|x} \otimes \text{id}_B)(\rho_{AB}) \in J$ by definition, by applying Φ to both sides we get $(f_{a|x} \otimes \text{id}_B)(\rho_{AB}) = \sum_{\lambda \in \Lambda} p(\lambda) p(a|x, \lambda) \Phi(\sigma_\lambda)$. Now since $\Phi(\sigma_\lambda) \in K_B \cap J$ we can without loss of generality replace K_B with $K_B \cap J$, simply because from any local hidden state model given by $\sigma_\lambda \in K_B$ we can construct a new local hidden state model given by $\Phi(\sigma_\lambda) \in K_B \cap J$. Note that it is sufficient that there is a channel $\Phi' : K_B \rightarrow K_B$ such that $\Phi'(v) = v$ for all $v \in J$, then one can use [38, lemma 1] to construct $\Phi : K_B \rightarrow K_B$ with the desired properties.

One may now hope to use similar ideas to extend theorem 2 to all states $\rho_{AB} \in K_{AB}$, but the following example shows that this is not possible.

Example 5 (Square-in-tetrahedron). Let $s_{00}, s_{10}, s_{01}, s_{11}$ be four points such that $\frac{1}{2}(s_{00} + s_{11}) = \frac{1}{2}(s_{10} + s_{01})$, then $S = \text{conv}(\{s_{00}, s_{10}, s_{01}, s_{11}\})$ is, up to an isomorphism, a

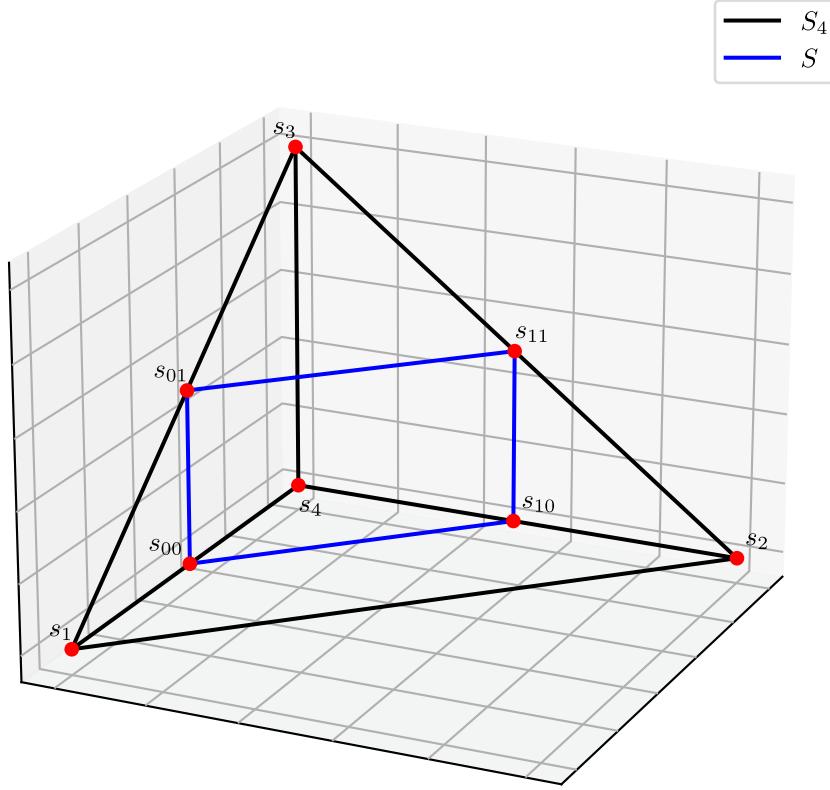


Figure 3. Embedding of the square state space S inside the simplex S_4 used in example 5.

square. Moreover let S_4 be the tetrahedron, i.e., a simplex with four vertices, then there is a subspace J such that $S = S_4 \cap J$, see figure 3. Note that $J \subset \text{span}(S_4)$, where $\dim(\text{span}(S_4)) = 4$, while figure 3 represents only $\text{aff}(S_4)$ as $\dim(\text{aff}(S_4)) = 3$. Let $K_A = S$ and $K_B = S_4$ then we have $S \dot{\otimes} S_4 = S \hat{\otimes} S_4$ [39] and so all bipartite states are separable. This trivially implies that for any measurements $m_x \in \mathcal{M}(E(K_A)) = \mathcal{M}(E(S))$ there is a local hidden state model and so no steering occurs. Let $V = \text{span}(S)$, then it follows from [21, proposition 4.5] that $(S \dot{\otimes} S_4) \cap (V \otimes J) = S \hat{\otimes} S$. It is known that there are bipartite states $\rho_{AB} \in S \hat{\otimes} S$ that maximally violate the CHSH inequality [7, 8, 40–44] which necessarily implies that there are measurements given by $m_x \in \mathcal{M}(E(K_A)) = \mathcal{M}(E(S))$ that steer $\rho_{AB} \in S \hat{\otimes} S$, because steering is necessary for violations of Bell inequalities [45]. Since $\rho_{AB} \in S \hat{\otimes} S$ and $S \subset S_4$, we also have $\rho_{AB} \in S \dot{\otimes} S_4$. It follows that in this case we cannot restrict S_4 to $S_4 \cap J$, because for $\rho_{AB} \in S \dot{\otimes} S_4$ a local hidden variable model always exists, while for the same $\rho_{AB} \in S \hat{\otimes} S$ it does not.

Note that in the last example we said that $\rho_{AB} \in S \dot{\otimes} S_4$ is separable while $\rho_{AB} \in S \hat{\otimes} S$ violates Bell inequality. This is because the measurements that Bob has to measure on S in order to violate the CHSH inequality are not valid measurements on S_4 . This is simply because $S \subset S_4$ implies $E(S_4) \subset E(S)$ where the inclusion is strict and the missing effects correspond to those measurements that are needed in order to violate the CHSH inequality.

Finally we present an example of steering of the isotropic state. Similar result was also obtained in [46] using incompatibility breaking channels.

Example 6 (Steering of the isotropic state). Let \mathcal{H} be a complex Hilbert space, $\dim(\mathcal{H}) = 2$ and let $K_A = K_B = \mathcal{D}(\mathcal{H})$. Let $\rho_{AB}^\gamma = \gamma|\phi^+\rangle\langle\phi^+|_{AB} + \frac{1-\gamma}{4}1_A \otimes 1_B$ be the isotropic state, where $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is the maximally entangled state and $|0\rangle, |1\rangle$ is the computational basis. We will assume that Alice and Bob share the state ρ_{AB} and that Alice has access to all possible measurements. It is known that the isotropic state is not steerable for $\gamma \leq \frac{5}{12}$ [36, 47]. In order to apply the result of theorem 2 we must show that ρ_{AB}^γ is K_B -full-dimensional. Let $M \in \mathcal{E}(\mathcal{H})$, which is equivalent to $0 \leq M \leq I$, then we have

$$\mathrm{Tr}_A((M \otimes 1)\rho_{AB}^\gamma) = \gamma M^\top + (1 - \gamma) \mathrm{Tr}(M)1. \quad (36)$$

It is now straightforward to check that for $\gamma > 0$ we have

$$\mathrm{span}(\gamma M^\top + (1 - \gamma) \mathrm{Tr}(M)1 : 0 \leq M \leq I) = \mathcal{B}(\mathcal{H}), \quad (37)$$

thus according to lemma 3 ρ_{AB}^γ is K_B -full-dimensional for $\gamma > 0$.

Let $0 < \gamma \leq \frac{5}{12}$ then according to theorem 2 $(K_{\rho_{AB}^\gamma}, \mathcal{E}(\mathcal{H}))$ is a restricted theory with preparation noncontextual hidden variable model. According to proposition 2 this is equivalent to all measurements in $(K_{\rho_{AB}^\gamma}, \mathcal{E}(\mathcal{H}))$ being $E(K_{\rho_{AB}^\gamma})$ -compatible. This is in fact a dual form of the known result that for $0 \leq \gamma \leq \frac{5}{12}$ the depolarizing channel is incompatibility-breaking [46], the connection here is that the state ρ_{AB}^γ is the Choi matrix of the corresponding depolarizing channel.

Let us now change the set of measurements available to Alice: we will assume that Alice can perform only projective measurements and their post-processings. The corresponding effect algebra is still $\mathcal{E}(\mathcal{H})$, but the set of all measurements available to Alice is not $\mathcal{M}(\mathcal{E}(\mathcal{H}))$ since we have just introduced higher-order restrictions. It is known that the state ρ_{AB}^γ is steerable by projective measurements if and only if $\gamma > \frac{1}{2}$ [36, 37], thus we get that the induced restricted theory has preparation noncontextual hidden variable model for $\gamma \leq \frac{1}{2}$. This is clearly different from our previous results and demonstrates that also introducing higher-order restrictions can in principle affect the existence of a hidden variable model.

It is not known whether the state ρ_{AB}^γ is steerable by all measurements for $\frac{5}{12} < \gamma < \frac{1}{2}$, which means that it is not known whether for such γ there is a preparation noncontextual hidden variable model for $(K_{\rho_{AB}^\gamma}, \mathcal{E}(\mathcal{H}))$ where Alice is allowed to perform any measurement. In the language of steering this is usually framed as an open problem of whether POVMs provide an advantage over projective measurements, in the language of contextuality this leads to an open problem of whether introducing higher-order restrictions can affect the existence of the respective hidden variable model for the induced restricted theory.

6. Conclusions

We investigated the connection between existence of noncontextual hidden variable models, incompatibility in restricted GPTs and steering. Our main results are theorem 1 which connects simplex-embeddability and incompatibility in restricted GPTs and theorem 2 which connects steerability of a state with existence of preparation noncontextual hidden variable model for the induced restricted theory. The proof of theorem 2 is build on proposition 2 which connects compatibility and existence of preparation noncontextual hidden variable model.

There are several open questions to be addressed in the future, the main is to search for operational equivalents of measurement noncontextual hidden variable models similar to proposition 2 and theorem 2. One can also ask for which GPTs satisfying no-restriction hypothesis measurement noncontextual hidden variable model exists; it is an open question whether

this is the case only if the state space of the GPT in question is a simplex, i.e., it is an open question whether equivalent of corollaries 1 and 2 also holds for measurement noncontextual hidden variable models.

Acknowledgments

The author is thankful to Chau Nguyen for discussions. This research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, Project Numbers 447948357 and 440958198), by the Sino-German Center for Research Promotion (Project M-0294), by the ERC (Consolidator Grant 683107/TempoQ), and by the Alexander von Humboldt Foundation.

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

Martin Plávala  <https://orcid.org/0000-0002-3597-2702>

References

- [1] Bell J S 1964 *Physics* **1** 195–200
- [2] Kochen S and Specker E 1967 *Indiana Univ. Math. J.* **17** 59–87
- [3] Budroni C, Cabello A, Gühne O, Kleinmann M and Larsson J Å 2021 Quantum contextuality (arXiv:[2102.13036](https://arxiv.org/abs/2102.13036))
- [4] Schmid D, Selby J H, Wolfe E, Kunjwal R and Spekkens R W 2021 *PRX Quantum* **2** 010331
- [5] Müller M 2021 *SciPost Phys. Lect. Notes* **28**
- [6] Plávala M 2021 General probabilistic theories: an introduction (arXiv:[2103.07469](https://arxiv.org/abs/2103.07469))
- [7] Popescu S and Rohrlich D 1994 *Found. Phys.* **24** 379–85
- [8] Barrett J 2007 *Phys. Rev. A* **75** 032304
- [9] Cavalcanti P J, Selby J H, Sikora J, Galley T D and Sainz A B 2021 Witworld: a generalised probabilistic theory featuring post-quantum steering (arXiv:[2102.06581](https://arxiv.org/abs/2102.06581))
- [10] Spekkens R W 2005 *Phys. Rev. A* **71** 052108
- [11] Spekkens R W 2007 *Phys. Rev. A* **75** 032110
- [12] Chiribella G and Yuan X 2014 Measurement sharpness cuts nonlocality and contextuality in every physical theory (arXiv:[1404.3348](https://arxiv.org/abs/1404.3348))
- [13] Schmid D, Spekkens R W and Wolfe E 2018 *Phys. Rev. A* **97** 062103
- [14] Schmid D, Selby J H, Pusey M F and Spekkens R W 2020 A structure theorem for generalized-noncontextual ontological models (arXiv:[2005.07161](https://arxiv.org/abs/2005.07161))
- [15] Schmid D, Selby J H and Spekkens R W 2020 Unscrambling the omelette of causation and inference: the framework of causal-inferential theories (arXiv:[2009.03297](https://arxiv.org/abs/2009.03297))
- [16] Gitton V and Woods M P 2020 Solvable criterion for the contextuality of any prepare-and-measure scenario (arXiv:[2003.06426](https://arxiv.org/abs/2003.06426))
- [17] Heinosaari T, Miyadera T and Ziman M 2015 *J. Phys. A: Math. Theor.* **49** 123001
- [18] Gühne O, Haapasalo E, Kraft T, Pellonpää J P and Uola R 2021 Incompatible measurements in quantum information science (arXiv:[2112.06784](https://arxiv.org/abs/2112.06784))
- [19] Filippov S N, Heinosaari T and Leppäjärvi L 2017 *Phys. Rev. A* **95** 032127
- [20] Jenčová A 2018 *Phys. Rev. A* **98** 012133

[21] Bluhm A, Jenčová A and Nechita I 2020 Incompatibility in general probabilistic theories, generalized spectrahedra, and tensor norms (arXiv:2011.06497)

[22] Uola R, Moroder T and Gühne O 2014 *Phys. Rev. Lett.* **113** 160403

[23] Quintino M T, Vértesi T and Brunner N 2014 *Phys. Rev. Lett.* **113** 160402

[24] Brunner N, Cavalcanti D, Pironio S, Scarani V and Wehner S 2014 *Rev. Mod. Phys.* **86** 419–78

[25] Tavakoli A and Uola R 2020 *Phys. Rev. Res.* **2** 013011

[26] Selby J H, Schmid D, Wolfe E, Sainz A B, Kunjwal R and Spekkens R W 2021 Contextuality without incompatibility (arXiv:2106.09045)

[27] Filippov S N, Gudder S, Heinosaari T and Leppäjärvi L 2020 *Found. Phys.* **50** 850–76

[28] Donadi S and Hossenfelder S 2020 A toy model for local and deterministic wave-function collapse (arXiv:2010.01327)

[29] Hossenfelder S and Palmer T 2020 *Front. Phys.* **8** 139

[30] Hardy L 2021 Time symmetry in operational theories (arXiv:2104.00071)

[31] Carmeli C, Heinosaari T and Toigo A 2019 *Phys. Rev. Lett.* **122** 130402

[32] Plávala M 2016 *Phys. Rev. A* **94** 042108

[33] Kuramochi Y 2020 *Positivity* **24** 1479–86

[34] Busch P, Heinosaari T, Schultz J and Stevens N 2013 *Europhys. Lett.* **103** 10002

[35] Jenčová A and Plávala M 2017 *Phys. Rev. A* **96** 022113

[36] Uola R, Costa A C S, Nguyen H C and Gühne O 2020 *Rev. Mod. Phys.* **92** 015001

[37] Wiseman H M, Jones S J and Doherty A C 2007 *Phys. Rev. Lett.* **98** 140402

[38] Barnum H, Barrett J, Leifer M and Wilce A 2007 *Phys. Rev. Lett.* **99** 240501

[39] Aubrun G, Lami L, Palazuelos C and Plávala M 2021 *Geom. Funct. Anal.* **31** 181–205

[40] Beckman D, Gottesman D, Nielsen M A and Preskill J 2001 *Phys. Rev. A* **64** 052309

[41] Hoban M J and Sainz A B 2018 *New J. Phys.* **20** 053048

[42] Crépeau C, Li J and Yang N 2017 private communication

[43] Plávala M and Ziman M 2020 *Phys. Lett. A* **384** 126323

[44] Jenčová A and Plávala M 2020 *Phys. Rev. A* **102** 042208

[45] Plávala M 2017 *Phys. Rev. A* **96** 052127

[46] Heinosaari T, Kiukas J, Reitzner D and Schultz J 2015 *J. Phys. A: Math. Theor.* **48** 435301

[47] Barrett J 2002 *Phys. Rev. A* **65** 042302