Bridgmannian operational theories

*Gábor Hofer-Szabó**

Abstract

In the framework of operational theories, introduced by Rob Spekkens (2005) as a generalization of quantum mechanics, one standardly assumes operational equivalence, the identification of measurement having the same distribution of outcomes in every preparation. From the Bridgmannian perspective of operationalism, this identification is unwarranted. Strict operationalism identifies observables not with operators but with measurement procedures. In this paper, we explore how operational theories without operational equivalence look like. We analyze these Bridgmannian theories with respect to the possible ontological models, contextuality, signaling, causal structure, fine-tuning and many other features.

Keywords: operational theory, operational equivalence, contextuality, causal models, faithfulness

1 Introduction

On strict operationalism, concepts should be defined by empirical operations. In this tradition, going back to Percy Bridgman (1927) and the Vienna Circle (Schlick, 1930), two concepts which are defined by different operational procedures cannot be the same. Using Bridgman's example, length measured by a ruler and length measured by light signals are different concepts, and true science should use different names to discern them. As times passed, philosophy of science (and also Bridgman himself) has gradually drifted away from strict operationalism and revealed various theoretical, semantic, pragmatic and common sense criteria for identifying concepts with different operational basis (Chang, 2019). In the case of physical magnitudes or observables, the standard way was to check whether the two measurements defining the two observables have the same outcome in their common domain. If the length of medium sized objects agree when measured by ruler or measured by light signals, then—at least in this common domain—one is justified in using one length concept instead of two. But for this comparison at least one of the two conditions needs to hold: (i) either the two measurements should be performed simultaneously on each system; (ii) or we need a precise enough preparation procedure which guarantees that the same system at different times is indeed in the same state and hence the measurements can be performed one after another. This second happens in classical physics where systems can be prepared in eigenstates providing definite results for every measurement. We prepare a system in a given eigenstate, perform the one measurement, prepare the system again in the previous eigenstate (or take a copy of it), perform the other measurement and compare the outcomes.

*Research Center for the Humanities, Budapest, email: szabo.gabor@btk.mta.hu

If the preparation procedure, however, is not as fine-grained as to yield eigenstates for every measurement, as is the case in quantum mechanics, it remains only the first option to identify observables lying behind different measurement procedures: to measure them simultaneously and check that the outcomes match in every single run.

But what if the two measurement procedures cannot be performed at the same time? Strictly speaking, in this case we are not entitled to identify the two observables. Still, in quantum mechanics this is what happens. Observables are identified not by measurement processes but by self-adjoint operators. An operator, however, specifies only the distribution of outcomes in a given state and not the outcome itself. Consequently, two different measurements which are represented by the same self-adjoint operator that is which provide the same outcome statistics in every quantum state are taken to measure the same observable even if the two measurements cannot be simultaneously performed in the above experimental-operational sense. From the Bridgmannian perspective, this identification is physically unjustified. The mere statistical match of outcomes of two measurements which cannot be performed at the same time on the same system does not guarantee that the two measurement would give the same outcome run-by-run and hence that they measure the same observable.

The identification of measurements represented by the same self-adjoint operator in quantum mechanics is often referred to as operational equivalence. It is introduced inductively and successively into an operational theory: one starts with a set of measurements and preparations and identifies measurements which provide the same outcome statistics in all preparations. Note, that this identification is relative to the set of available preparations; a new preparation procedure can break down operational equivalence if it discerns the measurement with respect to their outcome statistics.

Operational equivalence is sometimes expressed in the form that observables are not associated with measurement procedures, as in Bridgman, but with operators. So instead of *one measurement–one observable* we have *one operator–one observable*. Let me refer to the first identification of observables as *Bridgmannian* and to the second as *standard* (standard in quantum mechanics). Schematically:

$$
\begin{array}{rccccccc}
 & & \textit{Bridgmannian} & & \qquad & & \textit{Standard} & \\
\textit{Operator:} & & \mathbf{O} & & & & \mathbf{O} & \\
 & \swarrow & & \searrow & & & \downarrow & \\
\textit{Observable:} & \mathcal{O}_1 & & \mathcal{O}_2 & & & \mathcal{O} & \\
 & \downarrow & & \downarrow & & \swarrow & & \searrow \\
\textit{Measurement:} & M_1 & & M_2 & & M_1 & & M_2
\end{array}
$$

Although from the Bridgemannian perspective, the standard position is unsatisfactory, it has its own arguments. If all that quantum mechanics can predict, is the distribution of outcomes, and if there are no preparations which would discern two measurements with respect to the outcome distribution, then why would one like to discern the two measurements? They measure the same observables, like gas thermometer and alcohol thermometer measure the same temperature, and any difference in the concrete realization of the measurements is just of secondary importance.

The standard position remains consistent as long as we remain at the level of quantum theory. But at the moment when we try to extend the ontology by hidden states, the identification of different measurements represented by the same operator becomes problematic. The Kochen-Specker theorems highlight just this fact. It is instructive to see how Kochen-Specker theorems are interpreted on the standard approach. On this account, the lesson of the Kochen-Specker theorems is that the value of certain observables associated with operators depends on the measurement with which it is measured or co-measured. This fact is commonly referred to as contextuality or "ontological contextuality" (Redhead, 1989) if not only the value but also the observables themselves depend on which measurement they are measured by. But note, that from the Bridgmannian perspective, there is nothing contextual in this fact; it simply shows that we were too quick to identify observables measured by different measurements when we relied simply on the match of the outcome statistics.

In this paper, I will revisit the Bridgmannian view of operationalism and investigate how far we get when we do *not* identify operationally equivalent measurements. To this goal, I will use the framework of operational theories and ontological models introduced by Rob Spekkens (2005). This framework is general enough to embrace classical, quantum, super-quantum theories, and to analyze contextuality, causal structure and many other important features across the different theories. Operational theories come together with measurements and sets of simultaneous measurements. In the Bridgmannian spirit, I will identify measurements by sets of laboratory instructions and simultaneous measurements by the *conjunction* of such sets of instructions. In quantum mechanics, a set of simultaneous measurements is often replaced by one single operationally equivalent ("global") measurement which is represented by the same operator as the simultaneous measurements. In our Bridgmannian approach, however, this replacement leads to another operational theory with different measurements and different sets of simultaneous measurements. In the paper, I will construct to every operational theory another theory where no two measurements can be measured simultaneously, still the marginalizations of some of these new measurements are operationally equivalent. I will call this construction *trivialization.* With this construction in hand, I show the following:

1. The trivialization of an operational theory can be nicely represented graph theoretically as taking the line graph of the graph representing the original theory.

2. On the example of three non-disturbing (no-signaling) operational theories—a classical theory, the EPR-Bell scenario, and the Popescu-Rorhlich box, I will show how the most important features (such as contextuality, causal structure, etc.) of the ontological model change when we replace an operational theory with a new, trivialized theory.

3. We will discern two different and logically independent concepts of contextuality, simultaneous contextuality and measurement contextuality, and show that the trivialization can alter the ontological models with respect to the former but not to the latter.

4. We argue that the physical realization of the very same set of operators in quantum mechanics by different measurements can give rise to completely different ontological models with regard to simultaneous contextuality, causal structure and fine-tuning.

2 Operational theories

An *operational theory* is a theory which specifies the probability of the outcomes of certain measurements performed on a physical system which was previously prepared in certain states. Let $\mathcal{P} = \{P_1, P_2, \dots\}$ be set of *preparations* of the system, $\mathcal{M} = \{M_1, M_2, \dots\}$ the set of *measurements* which can be performed on the system, and let $\mathcal{X} = \{X_1, X_2, \dots\}$ be the set of *outcomes* of the measurement.[1] Let P, M, and X be random variables running over the preparations, measurements and outcomes, respectively, assigning to every event its index. (Sometimes, we will refer with "$P = 1$" to the preparation P_1, with "$M = 2$" to the measurement M_2, etc.) Using these random variables, an operational theory is simply a set of *conditional probabilities* of the outcomes given the various measurements and preparations, that is

$$p(X|M, P) \tag{1}$$

where P, M, and X run over the whole set $\mathcal{P}$, $\mathcal{M}$, and $\mathcal{X}$, respectively.

Two measurements M_1 and M_2 are *simultaneously measurable*, if they can be performed on the same system at the same time. Simultaneous measurability is an empirical question. Operationally, one identifies measurements by sets of laboratory instructions. The spin measurement of an electron, for example, is given by the detailed description of the path of the electron, the position of the Stern-Gerlach magnets and detectors, etc. As a consequence of this characterization of measurements by sets of laboratory instructions, two measurements M_1 and M_2 will be simultaneously measurable if and only if there is a measurement which can be identified by the *conjunction* of the sets of instructions characterizing M_1 and M_2. We call this measurement the *simultaneous measurement* of M_1 and M_2 and denote it by $M_1 \wedge M_2$ (which is again a measurement in $\mathcal{M}$). If M_1 and M_2 are not simultaneously measurable, we write $M_1 \wedge M_2 = \emptyset$. If a measurement in an operational theory is not a simultaneous measurement of two or more other measurements, then we call it a *basic measurement.*

An important consequence of defining measurements by sets of instructions is that we do *not* identify two measurements if they are operationally equivalent. Two measurements M_1 and M_2 are called *operationally equivalent* and denoted by $M_1 \sim M_2$ if they yield the same probability distribution of outcomes in every preparation[2] of the system, that is if

$$p(X|M_1, P) \quad = \quad p(X|M_2, P) \tag{2}$$

Being operationally equivalent does not mean that the two measurements are defined by the same set of instructions. Consequently, the simultaneous measurement of M_1 and M_2 will *not* be any measurement which is operationally equivalent to $M_1 \wedge M_2$. If it exists, $M_1 \wedge M_2$ will be the measurement defined by the conjunction of the instructions defining M_1 and the instructions defining M_2.

A maximal set of measurements which can be performed simultaneously on a system in an operational theory is called a *context.* M_1 and M_2 are in the same context if and only if

[1] Without loss of generality, we can assume that all measurements have the same number of outcomes. If not, we just add null-outcomes to the outcome set of some measurements.

[2] Which are again identified by sets of laboratory instructions.

$M_1 \wedge M_2 \in \mathcal{M}$. The set of all contexts is a *compatibility structure* of the theory. If in an operational theory there are no two measurements which can be simultaneously measured, then the compatibility structure is the empty set. We also refer to such operational theories as *trivial.*

We call an operational theory *non-disturbing*[3] if no conditional probability depends on whether the measurements are performed alone or along with simultaneous measurements, that is:

$$p(X|M, P) = p(X|M, M', P) \tag{3}$$

for any value pairs of the random variables M and M' for which the preimages are in the same context. Otherwise, the operational theory is called *disturbing.* Obviously, trivial operational theories are non-disturbing. If you want, you can always redefine measurements in a disturbing operational theory to make the theory non-disturbing. If two measurements M_1 and M_2 in $\mathcal{M}$ disturb one another, one can simply drop them from the theory and keep only $M_1 \wedge M_2$ as a kind of fine-grained measurement.

Next, we introduce a graph theoretical representation of operational theories borrowed from the literature on the Kochen-Specker theorems (Kochen and Specker, 1967). In Figure 1, we depicted

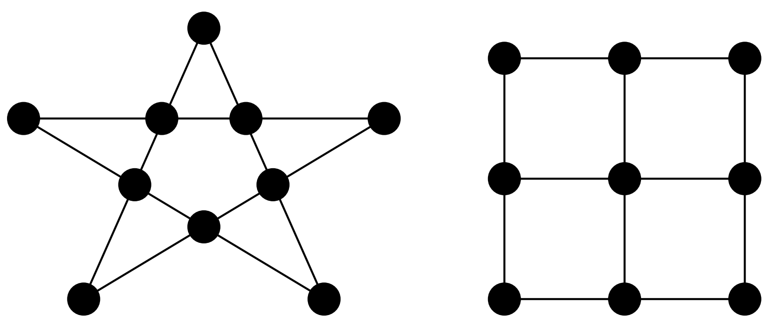

Figure 1: The GHZ graph and Peres-Mermin graph

the graph of two Kochen-Specker theorems, the GHZ theorem (Greenberger et al., 1990) on the left and the Peres-Mermin square (Peres, 1990; Mermin, 1993) on the right. The vertices of the graph represent self-adjoint operators and a subset of vertices is connected by a (hyper)edge[4] if and only if the corresponding operators are pairwise commutating. In the GHZ graph one has 10 operators and 5 commuting subsets; in the Peres-Mermin graph one has 9 operators and 6 commuting subsets. The (hyper)graph of most of the Kochen-Specker theorems is *linear*: each pair of hyperedges intersects in at most one vertex.

In this paper, we take over this graphic representation and use it in the framework of the operational theories but with a different meaning. Vertices will represent here basic measurements and (hyper)edges will represent sets of simultaneous measurements, that is contexts. In this interpretation, the above GHZ graph represents a non-trivial operational theory with 10 basic measurement arranged in 5 contexts and the Peres-Mermin graph represents a non-trivial theory

[3]Or *no-signaling*, if the measurements are spacelike separated.

[4]A hyperedge can connect more than two vertices.

with 9 basic measurement and 6 contexts. The (hyper)graph of both theory is linear: each basic measurement is featuring in exactly two contexts.

3 Ontological models

The role of an *ontological model* (hidden variable model) is to account for the conditional probabilities of an operational theory in terms of underlying realistic entities of the measured system called *ontic states* (hidden variables, elements of reality, beables). Let the set of ontic states be $\mathcal{L}$=$\{\Lambda_1, \Lambda_2, \dots\}$ and the random variable $\mathcal{L}$ over is Λ. An ontological model specifies a *probability distribution* over the ontic states associated with each preparation:

$$p(\Lambda|P) \tag{4}$$

and a set of *response functions* that is a set of conditional probabilities associated with every measurement and every ontic state:

$$p(X|M, \Lambda) \tag{5}$$

again with the obvious normalizations. Assuming the independence of the probability distributions from the measurements, called *no-conspiracy*:

$$p(\Lambda|M, P) = p(\Lambda|P) \tag{6}$$

and the independence of the response functions from the preparations in which the ontic states are featuring, called λ-*sufficiency*:

$$p(X|M, P, \Lambda) = p(X|M, \Lambda) \tag{7}$$

and using the theorem of total probability, one can recover the operational theory from the ontological model in terms of the probability distributions and response functions:

$$p(X|M, P) \quad = \quad \sum_{\Lambda} p(X|M, \Lambda)\, p(\Lambda|P) \tag{8}$$

An ontological model is called *outcome-deterministic (value-definite)* if

$$p(X|M, \Lambda) \in \{0, 1\} \tag{9}$$

otherwise it is called *outcome-indeterministic*.

Next, we define two different and logically independent concepts of noncontextuality (see Hofer-Szabó, 2021a, b, c). First, an ontological model is called *simultaneous noncontextual* if every ontic state determines the probability of the outcomes of every measurement independently of what other measurements are simultaneously performed, that is

$$p(X|M_i, \Lambda) = p(X|M_i \wedge M_j, \Lambda) \tag{10}$$

for any value pairs of measurements $M_i, M_j \in M$ which are in the same context; otherwise the model is called *simultaneous contextual.* Simultaneous noncontextuality is a kind of inference to the best explanation for why an operational theory is non-disturbing: if the ontological model for an operational theory is noncontextual in the sense of (10), then—assuming no-conspiracy (6) and λ-sufficiency (7)—one can show that the operational theory is non-disturbing (3).

Second, an ontological model is called *measurement noncontextual* if any two *operationally equivalent* measurements, that is $M_i, M_j \in M$ which have the same probability distribution of outcomes in every preparation

$$p(X|M_i, P) = p(X|M_j, P) \tag{11}$$

also have the same probability distribution of outcomes in every ontic state

$$p(X|M_i, \Lambda) = p(X|M_j, \Lambda) \tag{12}$$

Otherwise the model is called *measurement contextual.* Measurement noncontextuality is again a kind of inference to the best explanation; in this case the explanation of operational equivalence: (12)—together with no-conspiracy (6) and λ-sufficiency (7)—implies (11).

In quantum mechanics where operationally equivalent measurements $M_1 \sim M_2$ are represented by the same operator $\mathbf{O}$, measurement noncontextuality is just the requirement that the response functions of an ontological model should depend only on the operator and not on which specific measurement is realizing the operator, that is

$$p(X|M_1, \Lambda) = p(X|M_2, \Lambda) = p(X|\mathbf{O}, \Lambda)$$

Note, that trivial operational theories are trivially simultaneously noncontextual (since there are no simultaneous measurements) but they still can be measurement contextual. Also note that although simultaneous noncontextuality and measurement noncontextuality are different and logically independent notions, in case of non-disturbing theories measurement noncontextuality implies simultaneous noncontextuality: if M_j does not disturb M_i, then (11) holds for M_i and $M_i \wedge M_j$, but then, due to measurement noncontextuality, also (12), which is just simultaneous noncontextuality (10).

4 Trivialization

In this paper, I will investigate operational theories in pairs where the one theory has a non-trivial and the other theory a trivial compatibility structure. More precisely, I will construct from a non-trivial operational theory a trivial one by simply replacing each context of simultaneous measurements by a basic measurement (while preserving the preparations). This replacement goes along the following lines: Take a context in the non-trivial theory, that is a maximal set of simultaneous measurements $\{M_1, M_2, \ldots, M_k\}$ and replace it by one single measurement, denoted by $M_{12\ldots k}$ with outcome set of the form $\mathcal{X}^{(1)} \times \mathcal{X}^{(2)} \times \ldots \times \mathcal{X}^{(k)}$. Schematically,

$$\text{Context:}\, \{M_1, M_2, \ldots, M_k\} \quad \longrightarrow \quad \text{Basic measurement:}\, M_{12\ldots k} \tag{13}$$

Furthermore, require that $M_{12\ldots k}$ be operationally equivalent to the simultaneous measurement $M_1 \wedge M_2 \wedge \cdots \wedge M_k$. If the non-trivial operational theory is non-disturbing, then it also follows that for all preparations the outcome distribution for the simultaneous measurement of every *subset* of the context is recovered as a marginal of the outcome distribution of $M_{12\ldots k}$.

An example might help. Suppose the measurements of a non-trivial theory are

$$\mathcal{M} = \{M_1, M_2, M_3, M_1 \wedge M_2, M_1 \wedge M_3\}$$

that is the compatibility structure is $\big\{\{M_1, M_2\}, \{M_1, M_3\}\big\}$. Then the measurements of the new, trivial theory will be $\mathcal{M}' = \{M_{12}, M_{13}\}$ with the following operational equivalences:

$$M_{12} \sim M_1 \wedge M_2, \quad M_{13} \sim M_1 \wedge M_3, \quad \sum_2 M_{12} \sim \sum_3 M_{13} \sim M_1$$

where $\sum_2 M_{12}$ denotes that we cluster the outcomes of M_{12} only according to the first index.

Note that even though M_{12} is indexed by two indices, it is just as a basic measurement in the new operational theory as M_1 and M_2 was in the old theory. The only reason why we use these multiple indices is to be able to relate M_{12} to M_1 and M_2 simply by marginalization and operational equivalence. We could have indexed M_{12} also using only one single index and then relate it to M_1 and M_2 by different functions: These functions, however, can be quite complicated. Again, M_{12} is not necessarily a composite measurement such as the simultaneous measurement $M_1 \wedge M_2$. it well can be a simple measurement. One need not think of M_{12} as being two measurements performed on two different subsystems, as in the usual spin measurement scenarios in quantum mechanics. M_{12} can be also be a simple measurement on a localized system.

Now, having replaced each set of simultaneous measurements with a basic measurement, the new operational theory will be trivial: no two measurements can be simultaneously measured. If we represent this new operational theory by a graph, this graph will have only vertices but no edges. The first two graphs in Figure 2 show the graph of our above mini operational theory and the trivialized new theory.

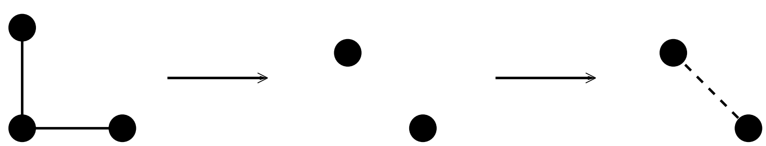

Figure 2: The graph and line graph of our mini operational theory

We could stop at this point but then some information would be lost in the new operational theory, namely, that certain marginalizations of the new measurements are operationally equivalent. To preserve this information, we add (hyper)edges to the graph of the new theory *with the following meaning*: we draw an (hyper)edge between a set of vertices in the trivial theory, if the contexts they represented in the non-trivial theory had at least one common basic measurement, or equivalently, if the corresponding basic measurements have operationally equivalent marginalizations in the trivial theory. For example, the graph of our mini operational will get an edge (see the third graph in Figure 2) because the marginalization of M_{12} and M_{13} are operationally equivalent to one another (both being operationally equivalent to M_1). Note that the (hyper)edges in the graph of the non-trivial and trivial theories mean different things: in the non-trivial operational theory they meant simultaneous measurability, while in the trivial operational theory they mean having operationally equivalent marginalizations. To express this different interpretations of the edges, we use continuous line in the non-trivial operational theories and broken line in the trivial ones.

This construction can be nicely represented graph theoretically by simply taking the line graphs of the (hyper)graph of the non-trivial operational theory. A *line graph* $L(G)$ is constructed from a graph G such that for each (hyper)edge in G we make a vertex in $L(G)$ and for every two (hyper)edges in G that have a vertex in common, we make an edge between their corresponding vertices in $L(G)$. The line graphs of the GHZ graph and Peres-Mermin graph, for example, are depicted in Figure 3. The number of the vertices and edges flip in both line graphs: the line

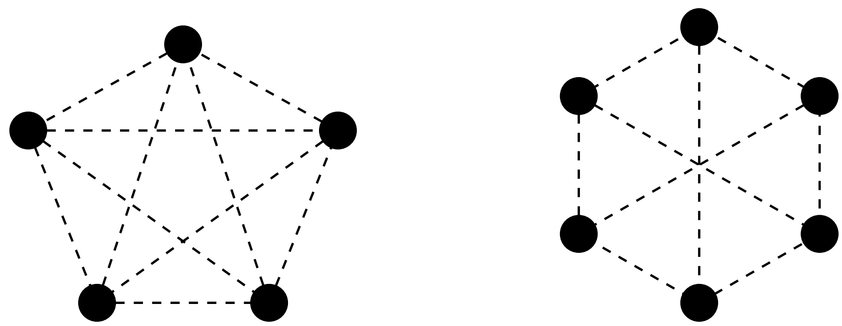

Figure 3: The line graphs of the GHZ graph and Peres-Mermin graph

graph of the GHZ graph contains 5 vertices and 10 edges, the line graph of the Peres-Mermin graph contains 6 vertices and 9 edges. Since both the GHZ graph and Peres-Mermin graph are linear, their line graphs contain only edges but no hyperedges.

To sum up, in the graph G of the non-trivial operational theory, vertices represent the old basic measurements and (hyper)edges represented contexts that is sets of simultaneous measurements. In the line graph $L(G)$ of the trivialized theory, vertices represent the new basic measurements but—since there are no simultaneous measurements— the (hyper)edges mean something else: they connect vertices representing measurements which have operationally equivalent marginalizations.

5 Three operational theories with non-trivial compatibility structure

In this Section, we consider three non-trivial operational theories each with the same four basic measurements A_0, A_1, B_0, B_1 such that A_0, A_1 have binary outcomes X_0, X_1 and B_0, B_1 have binary outcomes Y_0, Y_1. The compatibility structure of all three theories is

$$\big\{\{A_0, B_0\}, \{A_0, B_1\}, \{A_1, B_0\}, \{A_1, B_1\}\big\}$$

Consequently, the graph (and line graph, see next section) depicted in Figure 4 is the same for all three operational theories. Since the graph is linear, the line graph contains only edges and no hyperedges.

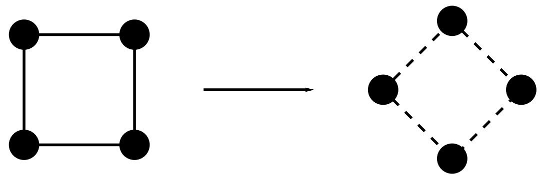

Figure 4: The graph and line graph of the three operational theories

Let A be a random variable over the measurements $\{A_0, A_1\}$ and B a random variable over the measurements over the measurements $\{B_0, B_1\}$. Similarly, let X and Y be random variables over the outcomes $\{X_0, X_1\}$ and $\{Y_0, Y_1\}$, respectively, such that $A, B, X, Y = 0, 1$. The operational theories differ in the preparations. Each theory has only one preparation: the first one P_{CL}, the second P_{EPR}, and the third P_{PR}. We refer to the operational theories as a *classical operational theory*, the *EPR-Bell situation*, and the *Popescu-Rorhlich (PR) box* (Popescu and Rohrlich, 1994), respectively.

The three operational theories can be characterized by the following conditional probabilities:

$$p(X|A,P) = p(Y|B,P) \;=\; \frac{1}{2} \tag{14}$$

$$p(X,Y|A,B,P_{\mathrm{CL}}) \;=\; \begin{cases} \frac{1}{2} & \text{if } X \oplus Y = 0 \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

$$p(X,Y|A,B,P_{\mathrm{EPR}}) \;=\; \begin{cases} \frac{3}{8} & \text{if } X \oplus Y = 0 \text{ and } A \cdot B = 0 \\ \frac{1}{8} & \text{if } X \oplus Y = 1 \text{ and } A \cdot B = 0 \\ \frac{1}{2} & \text{if } X \oplus Y = 0 \text{ and } A \cdot B = 1 \\ 0 & \text{if } X \oplus Y = 1 \text{ and } A \cdot B = 1 \end{cases} \tag{16}$$

$$p(X,Y|A,B,P_{\mathrm{PR}}) \;=\; \begin{cases} \frac{1}{2} & \text{if } X \oplus Y = A \cdot B \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

where P is a variable over $\mathcal{P} = \{P_{\mathrm{CL}}, P_{\mathrm{EPR}}, P_{\mathrm{PR}}\}$ and $\oplus$ is the sum modulo 2. We come back to the quantum mechanical representation of the EPR-Bell situation in Section 7.

All three operational theories are non-disturbing:

$$p(X|A,P) \;=\; p(X|A,B,P) = \frac{1}{2} \tag{18}$$

$$p(Y|B,P) \;=\; p(Y|A,B,P) = \frac{1}{2} \tag{19}$$

The CHSH expression (Clauser, Horne, Shimony, and Holt, 1969)

$$\mathrm{CHSH}_P = \;\; \langle A_0, B_0\rangle_P + \langle A_0, B_1\rangle_P + \langle A_1, B_0\rangle_P - \langle A_1, B_1\rangle_P \tag{20}$$

where

$$\langle A, B\rangle_P \quad = \quad p(X \oplus Y = 0|A, B, P) - p(X \oplus Y = 1|A, B, P)$$

is 2 for the classical theory, satisfying the CHSH inequality, $|\mathrm{CHSH}_P| \leqslant 2$; it is 2.5 for the EPR-Bell situation, violating (not maximally) the CHSH inequality; and 4 for the PR box which is beyond the Tsirelson bound $2\sqrt{2}$.

Next, we construct an ontological model for each operational theory. The exact probabilistic specification of the models in terms of distributions and response functions is given in the Appendix. From our perspective, however, it will be more instructive to look at the *bundle diagrams* (see Abramsky et al., 2017; Abramsky & Brandenburger, 2011) of the models depicted in Figure 5.

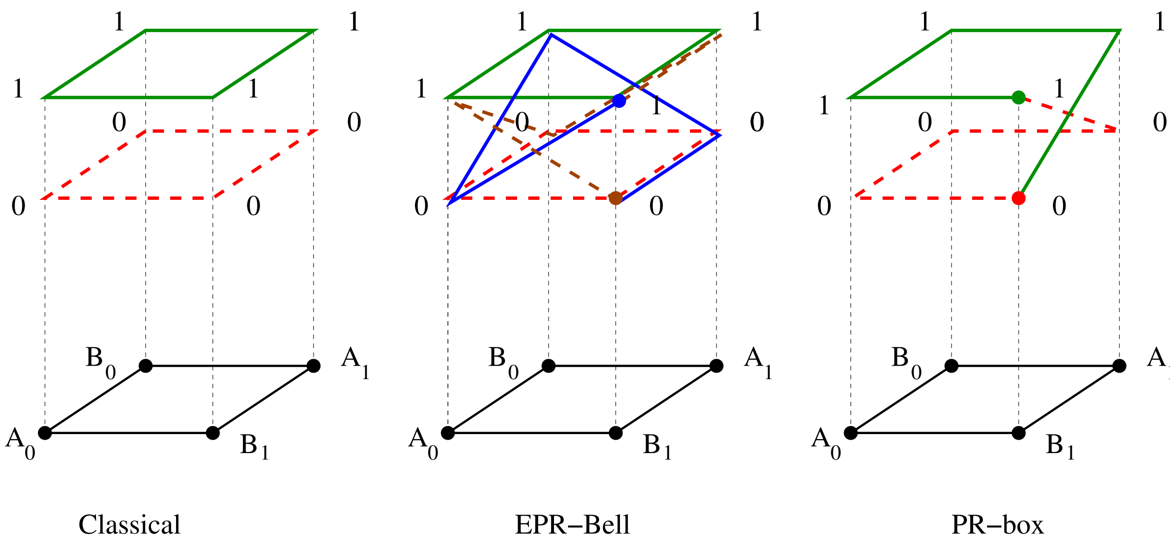

Figure 5: Bundle diagrams of the ontological models for the operational theories with non-trivial compatibility structure

First, look at the "cuboid" of the classical model on the left. The quadrangle at the bottom is the base space of the bundle, actually the graph of the operational theory "laid down". It consists of four vertices representing the four measurements A_0, A_1, B_0, B_1 such that two measurements are connected if and only if they are in the same context. The vertical broken lines are the fibers of the bundle. The two vertices on a given fiber at different heights denoted by 0 and 1 represent the outcomes of the corresponding measurements: $X = 0, 1$ for $A = 0, 1$ and $Y = 0, 1$ for $B = 0, 1$. Now, there are two quadrangles in the figure, one connecting the upper vertices of the adjacent fibers and one connecting the lower vertices. Each quadrangle represents the response functions of the model for a given ontic state. The green and continuous upper quadrangle belongs to the ontic state Λ_1. In this ontic state the outcome of each measurement is 1. (The outcome of A_0, A_1 is X_1 and the outcome of B_0, B_1 is Y_1.) The red and broken lower quadrangle belongs to the ontic state Λ_0 for which each outcome is 0. The model is outcome-deterministic. It also fixes the outcomes of the simultaneous measurements in the different contexts such that no outcome of any measurement in any ontic state depends on whether a simultaneous measurement is also performed. Thus, the model is simultaneous noncontextual. Moreover, the model is also measurement noncontextual: in both ontic state the outcome of any two operationally equivalent measurements is the same. Setting the probability of both ontic states to $\frac{1}{2}$, the operational theory can be recovered.

Let us now go over to the bundle diagram of the PR-box on the right of Figure 5. Again, we have two ontic states Λ_1 and Λ_0 but the green and red lines do not close now. They are discontinuous at the fibre of B_1. To avoid ambiguity with respect to the outcome of B_1, we put a dot at the one end of both discontinuous lines. This dot indicates the outcome of B_1 if measured *alone* and not together with A_0 or A_1 (when the outcome of B_1 is indicated by the value of the appropriate segment of the green or red lines connecting the fibre of B_1 with the fibre of A_0 or A_1). Thus, the model is outcome deterministic. However, it is simultaneous contextual:

$$\delta_{Y,\Lambda} = p(Y|B_1, \Lambda) \neq p(Y|A_1 \wedge B_1, \Lambda) = \delta_{Y\oplus 1,\Lambda} \tag{21}$$

That is performing the measurement B_1 in the "green" ontic state, Λ_1, together with A_1, the outcome of B_1 will be Y_0, while performing B_1 together with A_0, the outcome will be Y_1; and *vice versa* for the "red" ontic state, Λ_0. Note that the model cannot be made simultaneously noncontextual even at the price of giving up outcome determinism: set $p(Y|B_1, \Lambda)$ as you wish, it cannot agree with both $p(Y|A_0 \wedge B_1, \Lambda)$ and $p(Y|A_1 \wedge B_1, \Lambda)$ since they are different. Since simultaneous contextuality implies measurement contextuality for non-disturbing theories, the model for the PR-box will also be measurement contextual. Indeed,

$$p(Y|B_1, P_{\mathrm{PR}}) = p(X|A_1 \wedge B_1, P_{\mathrm{PR}}) \tag{22}$$

despite the fact that inequality (21) holds. We can recover the PR-box theory again by setting the probability of both ontic states to $\frac{1}{2}$.

Finally, the bundle diagram in the middle of Figure 5 represents an ontological model for the EPR-Bell scenario. Here we have four ontic states portrayed by lines of different color and style. The "green" and "red" ontic states are outcome deterministic and noncontextual in both senses. The "blue" and "brown" ontic states, however, are outcome deterministic but simultaneous and hence measurement contextual: their lines do not close on the fibre of B_1. This means that in these ontic states the outcome of B_1 will be different when measured alone and when co-measured with A_0 or A_1. The dots at one end of the lines indicate the outcomes the outcome of B_1 when measured alone. By setting the probability of the two noncontextual ontic states to $\frac{3}{8}$ and the probability of the two contextual ontic states to $\frac{1}{8}$, the probabilities of the EPR-Bell scenario can be recovered (see Appendix).

To sum up, we constructed three outcome-deterministic ontological models for the three operational theories such that the model for the classical theory is noncontextual (in both senses) and the models for other two theories are contextual (again, in both senses). This is in tune with the satisfaction and violation of the CHSH inequality for the different theories.

6 Three operational theories with trivial compatibility structure

The three operational theories in the previous Section were non-trivial, they had a non-trivial compatibility structure. Let us now "trivialize" them in the way outlined in Section 2. This means that we introduce four new measurements associated with the four old contexts ($i, j = 0, 1$):

$$\text{Context:}\,\{A_i, B_j\} \qquad \longrightarrow \qquad \text{Basic measurement:}\,C_{ij} \tag{23}$$

Thus, the new set of measurements is $\mathcal{C} = \big\{C_{00}, C_{01}, C_{10}, C_{11}\big\}$ with trivial compatibility structure, each measurement having the same outcome space $\mathcal{Z} = \big\{Z_{00}, Z_{01}, Z_{10}, Z_{11}\big\}$. Let C be a random variable over $\mathcal{C}$ assigning to every measurement its index pair. Similarly, let Z be a random variable over $\mathcal{Z}$ assigning to every outcome its index pair. Both C and Z can be expressed as Cartesian products: $C = C_1 \times C_2$ and $Z = Z_1 \times Z_2$ where C_1 and Z_1 assign to every measurement or outcome its first index and C_2 and Z_2 assign the second index.

The following marginalizations of the new basic measurements are operationally equivalent and also equivalent to the old basic measurements:

$$\begin{aligned}
\sum_2 C_{00} \sim \sum_2 C_{01} \sim A_0, &\qquad \sum_2 C_{10} \sim \sum_2 C_{11} \sim A_1 \\
\sum_1 C_{00} \sim \sum_1 C_{10} \sim B_0, &\qquad \sum_1 C_{01} \sim \sum_1 C_{11} \sim B_1
\end{aligned}$$

The line graph of the trivial theory is depicted on the right side of Figure 4.

The three trivial operational theory can be characterized by the following conditional probabilities:

$$p(Z|C, P_{\text{CL}}) \;=\; \begin{cases} \frac{1}{2} & \text{if } Z \oplus Z_2 = 0 \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

$$p(Z|C, P_{\text{EPR}}) \;=\; \begin{cases} \frac{3}{8} & \text{if } Z \oplus Z_2 = 0 \text{ and } C \cdot C_2 = 0 \\ \frac{1}{8} & \text{if } Z \oplus Z_2 = 1 \text{ and } C \cdot C_2 = 0 \\ \frac{1}{2} & \text{if } Z \oplus Z_2 = 0 \text{ and } C \cdot C_2 = 1 \\ 0 & \text{if } Z \oplus Z_2 = 1 \text{ and } C \cdot C_2 = 1 \end{cases} \tag{25}$$

$$p(Z|C, P_{\text{PR}}) \;=\; \begin{cases} \frac{1}{2} & \text{if } Z \oplus Z_2 = C_1 \cdot C_2 \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

Observe that the probabilistic description of trivial operational theory is formally analogous with the non-theory of the previous section: we obtain equations (15)-(17) from (24)-(26) by simply replacing C_1, C_2, Z_1, Z_2 with A, B, X, Y, respectively. The measurements and outcomes, however, are different in the two theories.

All three operational theories are non-disturbing in a trivial sense: there are no simultaneous measurements. Therefore, the CHSH inequalities cannot be defined. Again, one can construct an ontological model for each operational theory. The distribution of ontic states is the same as in the models for the non-trivial theories. The response functions are obtained from those of the non-trivial theory by simply replacing A, B, X, Y with C_1, C_2, Z_1, Z_2. All this is specified in the Appendix and visualized in Figure 6. As can be seen, the lines representing the outcomes of

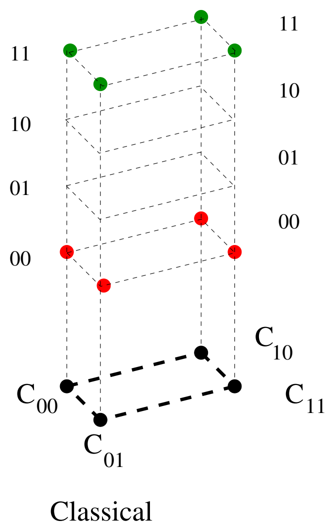

Figure 6: Bundle diagrams of the ontological models for the operational theories with non-trivial compatibility structure

simultaneous measurements have disappeared. Each basic measurement has a definite outcome in every ontic state denoted by a dot at the appropriate height on the fibre corresponding to the measurement. In the classical theory and in the PR box there are two ontic states ("green" and "red"), in the EPR-Bell scenario there are four ontic states ("green", "red", "blue" and "brown"). All three models are outcome-deterministic and simultaneously non-contextual since there are no simultaneous measurement. But the non-classical (EPR and PR) models are measurement contextual. Certain marginalizations of the measurements, for example, $\sum_1 C_{01}$ and $\sum_1 C_{11}$ are operationally equivalent. Still, both the "blue" and "brown" ontic states in the EPR model and "green" and "red" ontic states in the PR model assign different outcomes for them. This shows, that measurement noncontextuality is a stronger concept than simultaneous noncontextuality.

7 The causal structure of the ontological models

Let us turn now to the causal structure of the ontological models. Since these models provide information only about the probabilistic relations of the events and not about their spatiotemporal or other relations, the reconstruction of the causal structure will rely solely on these probabilistic information. The machinery to deduce causal relations from probabilistic relations is known as *causal discovery algorithms* and was introduced in (Pearl, 2009; Spirtes, Glymour, Scheines, 2001). These algorithms do not make use of the full probabilistic setting, they use only the conditional and unconditional independence relations to construct a causal graph. A causal graph is a directed acyclic graph (DAG),[5] where the vertices represent random variables and the directed edges represent causal relevance between these variables. For a variable X, the set of vertices that have directed edges in X is called the parents of X, denoted by $Par(X)$, and the set of vertices that are endpoints of a directed paths from X is called the descendants of X, denoted by $Des(X)$. A set V of random variables (on a classical probability space) is said to satisfy the *Causal Markov Condition* relative to a causal graph G if for any $X \in V$ and $Y \notin Des(X)$:

$$p(X|Par(X), Y) = p(X|Par(X))$$

That is, conditioning on its parents any random variable will be probabilistically independent from any of its non-descendants.

Now, causal discovery algorithms take as input a set of conditional and unconditional independence relations among random variables and provide a causal graph G as output which returns these independence relations if the Causal Markov Condition is applied to the graph.[6] Here we do not enter into the details of these algorithms; rather we simply list the independence relations of the ontological models of the non-trivial and trivial operational theory and the corresponding causal graphs.[7]

The conditional independence relations in the ontological models of our three *non-trivial* theories are the following:

$$\begin{aligned}
p(X|A,B) &= p(X|A) &(27)\\
p(Y|A,B) &= p(Y|B) &(28)\\
p(X|A,Y,\Lambda) &= p(X|A,\Lambda) &(29)\\
p(Y|X,B,\Lambda) &= p(Y|B,\Lambda) &(30)\\
p(X|A,B,\Lambda) &= p(X|A,\Lambda) &(31)\\
p(Y|A,B,\Lambda) &\overset{\text{(CL)}}{=} p(Y|B,\Lambda) &(32)
\end{aligned}$$

The first two relations are just the non-disturbance equations (18)-(19), the subsequent relations follow from the appropriate response functions (42)-(44), (47)-(49), and (52)-(54) of the models specified in the Appendix. The first five conditional independence relations (27)-(31) hold for all the three models but the last relation (32) holds only for the classical model.

The causal graphs which return the independences for the three models are depicted in Figure 7. These graphs are *minimal* in the sense that no subgraph can return all the independence relations. Applying the Causal Markov Condition to the graphs, one obtains also an extra

[5]Note that these causal graphs are different from the graphs and line graphs used in the previous Sections to represent compatibility structure and common marginalization.

[6]More precisely, the independence relations are returned if all those graphical criteria are applied to the graph which can be derived from the Causal Markov Condition plus the semi-graphoid axioms. These criteria are captured by the so-called *d*-separation criterion (see Pearl, 2009, Ch. 1).

[7]For the application of the causal discovery algorithm for the EPR-Bell scenario, see (Suarez and SanPedro, 2009; Wood and Spekkens, 2015).

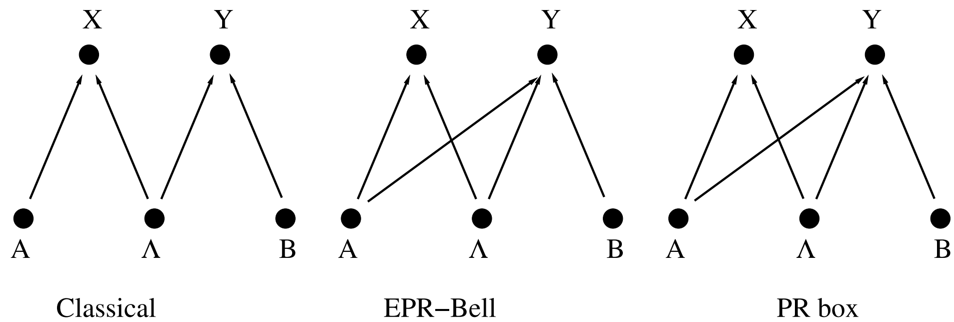

Figure 7: Causal structure of the ontological models with non-trivial compatibility structure

unconditional independence relation among the exogenous variables (that is variables which have no parents):

$$p(A, B, \Lambda) \quad = \quad p(A)p(B)p(\Lambda) \tag{33}$$

These relations are not specified in the model but are consistent with it. They are a special case of the *no-conspiracy* condition (6).

Observe that there is an edge in the graph of the non-classical models connecting A and Y. This edge represents the causal influence responsible for simultaneous contextuality: the value of Y causally depends not only on the value of X and Λ but also on the value of Y. If A and Y are spacelike separated, this edge represents a non-local causal influence. Note again, however, that in constructing the graphs, we relied only on the probabilistic features of the models and not on the spatiotemporal localizations of the events—in strong contrast to the usual EPR-Bell analysis.

A further difference between the classical and non-classical models concerns fine-tuning. To see this, first recall that any joint probability distribution of the random variables which is compatible with the corresponding causal graph in Figure 7 is of the form

$$p(X, Y, A, B, \Lambda) \stackrel{\text{(CL)}}{=} p(X|A, \Lambda)p(Y|B, \Lambda)p(A)p(B)p(\Lambda) \tag{34}$$

for the classical model and of the form

$$p(X, Y, A, B, \Lambda) \stackrel{\text{(EPR, PR)}}{=} p(X|A, \Lambda)p(Y|A, B, \Lambda)p(A)p(B)p(\Lambda) \tag{35}$$

for the non-classical models. In both equations, the conditional probabilities (the response functions) are called *causal parameters* and the unconditional probabilities are called *statistical parameter* (where $p(\Lambda)$ is just a short hand for $p(\Lambda|P)$). By manipulating these parameters, one obtains all the joint distributions compatible with the causal graphs. Since causal discovery algorithms are sensitive only to the independence relations and not to the full joint probability distribution, the question arises, whether these independence relations are robust enough against the perturbation of the causal-statistical parameters, that is whether they continue to hold when these parameters are not those specified in the Appendix but take on other values. If so, the graph is said to be *faithful*, if not, it is said to be *fine-tuned*.

Now, for the classical model all the conditional independences (27)-(32) can be derived from the joint probability distribution equation (34) plus the theorem of total probability. This means that the conditional independences hold for any choice of the parameters. Thus, the classical model is faithful. The crucial step in the derivation of the conditional independences is the factorization

$$p(X, Y|A, B, \Lambda) \overset{\text{(CL)}}{=} p(X|A, \Lambda)p(Y|B, \Lambda)$$

By summing up for the different variables, one recovers the different conditional independences (27)-(32). In the non-classical models, however, instead of the factorization one has

$$p(X, Y|A, B, \Lambda) \overset{\text{(EPR, PR)}}{=} p(X|A, \Lambda)p(Y|A, B, \Lambda)$$

and hence summing up does not recover (32) and the non-disturbance (28). And indeed, for a non-zero measure of the choice of the parameters, these conditional independences will fail to hold. Therefore, the non-classical models are fine-tuned.

These facts are in tune with Cavalcanti's (2018) theorem on bipartite Bell scenarios stating that every causal model for a non-disturbing operational theory violating the CHSH inequality requires fine-tuning. Cavalcanti's result points out a deep connection between simultaneous contextuality of the model and fine tuning of the corresponding graph. At the end of his paper, he asks whether also *measurement* noncontextuality can be understood as arising from the no-fine-tuning condition. To answer Cavalcanti's question, let us now turn to the causal structure of the ontological models of the *trivial theory*. In these models there are no conditional independence relations, except that among the exogenous variables:

$$p(C, \Lambda) = p(C)p(\Lambda) \tag{36}$$

which is again consistent with the models. The causal graph which is compatible with (36) is depicted in Figure 8. Note, that the graph is the same for all three models. The four measurements

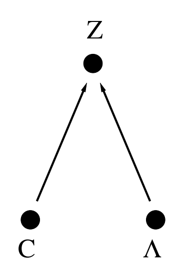

Figure 8: Causal structure of the ontological models with trivial compatibility structure

cannot be simultaneously performed, therefore the models are (trivially) simultaneously noncontextual. The models are also faithful since any choice of the parameters in the joint probability distribution equation

$$p(Z, C, \Lambda) = p(Z|C, \Lambda)p(C)p(\Lambda) \tag{37}$$

compatible with the graph in Figure 8 will return the same independence relations, that is (36). Thus, the answer to Cavalcanti's question is no. Both the models of the non-trivial and trivial non-classical operational theories are measurement contextual, still the causal graph is fine-tuned for the former and faithful for the latter. Fine-tuning relates to simultaneous contextuality but not to measurement contextuality.

To sum up, the causal graph of the classical and non-classical models of the non-trivial operational theory are different; the graph of the non-classical models is fine-tuned and contains an edge representing simultaneous contextuality. This difference between the graphs collapses upon trivializing the theories; the graph of all three models will be the same and will be trivial and faithful.

Quantum mechanics

Quantum mechanics is an operational theory in a special linear algebraic representation. Therefore, it is instructive to see the quantum mechanics represents the EPR-Bell scenario and see how this relates to the Bridgmannian and the standard identification of observables. The probabilities of the both the non-trivial operational theory (14) and (16) and the trivial operational theory (25) are generated quantum mechanically as follows:

$$\langle \Psi_s | (\mathbf{X^A} \otimes \mathbf{I}) \Psi_s \rangle = p(X|A, P_{\text{EPR}}) \tag{38}$$

$$\langle \Psi_s | (\mathbf{I} \otimes \mathbf{Y^B}) \Psi_s \rangle = p(Y|B, P_{\text{EPR}}) \tag{39}$$

$$\langle \Psi_s | (\mathbf{X^A} \otimes \mathbf{Y^B}) \Psi_s \rangle = p(X, Y|A, B, P_{\text{EPR}}) \tag{40}$$

where $|\Psi_s\rangle$ is the singlet state representing the preparation P_{EPR} in the Hilbert space $H_2 \otimes H_2$; $\mathbf{I}$ is the unit operator in H_2; and $\mathbf{X^A}$ and $\mathbf{Y^B}$ scroll over eight projections

$$\mathbf{X_0}^{\mathbf{A_0}}, \mathbf{X_1}^{\mathbf{A_0}}, \mathbf{X_0}^{\mathbf{A_1}}, \mathbf{X_1}^{\mathbf{A_1}}$$
$$\mathbf{Y_0}^{\mathbf{B_0}}, \mathbf{Y_1}^{\mathbf{B_0}}, \mathbf{Y_0}^{\mathbf{B_1}}, \mathbf{Y_1}^{\mathbf{B_1}}$$

corresponding to eight unit vectors $|X^A\rangle$ and $|Y^B\rangle$ in H_2 such that

$$|\langle X^A | Y^B \rangle|^2 = \begin{cases} \frac{3}{4} & \text{if } X \oplus Y = 0 \text{ and } A \cdot B = 0 \\ \frac{1}{4} & \text{if } X \oplus Y = 1 \text{ and } A \cdot B = 0 \\ 1 & \text{if } X \oplus Y = 0 \text{ and } A \cdot B = 1 \\ 0 & \text{if } X \oplus Y = 1 \text{ and } A \cdot B = 1 \end{cases}$$

The operators representing the four measurements are:

$$\mathbf{A_0} = \mathbf{X_0}^{\mathbf{A_0}} - \mathbf{X_1}^{\mathbf{A_0}}, \qquad \mathbf{A_1} = \mathbf{X_0}^{\mathbf{A_1}} - \mathbf{X_1}^{\mathbf{A_1}}$$
$$\mathbf{B_0} = \mathbf{Y_0}^{\mathbf{B_0}} - \mathbf{Y_1}^{\mathbf{B_0}}, \qquad \mathbf{B_1} = \mathbf{Y_0}^{\mathbf{B_1}} - \mathbf{Y_1}^{\mathbf{B_1}}$$

with eigenvalues ± 1.

The operators, however, represent different measurements in the non-trivial and trivial operational theory. Consider, for example, the quantum optical realization of the EPR-Bell scenario. In both operational theories, one prepares an ensemble of photon pairs in singlet state and performs certain polarization measurements on the pairs. In the *non-trivial* theory, however, one has four *local* measurements: two linear polarization measurements on the left photon, A_0 and A_1, and two linear polarization measurements on the right photon, B_0 and B_1. These measurements are the following:

A_0 : Measure the linear polarization of the left photon along a given transverse axis a_0 (with outcome $+1$ if the photon passes the polarizer and -1 if not)

A_1 : Measure the linear polarization of the left photon along a transverse axis a_1 at $60°$ from the axis a_0

B_0 : Measure the linear polarization of the right photon along a transverse axis b_0 at $60°$ from the axis both a_0 and a_1

B_1 : Measure the linear polarization of the right photon along the transverse axis $b_1 = a_1$

The polarization measurements on the left subsystem can be simultaneously performed with the polarization measurements on the right subsystem realizing the simultaneous measurements $A_0 \wedge B_0$, $A_0 \wedge B_1$, $A_1 \wedge B_0$, and $A_1 \wedge B_1$. The local measurements do not disturb one another, still the ontological model constructed above is simultaneous contextual: performing measurement A_0 or A_1 causally influences the outcomes of B_0 and B_1. Since the the events A and Y are spacelike separated, this is a clear violation of local causality.

In the *trivial* operational theory, however, one has different measurements. Here one has four *global* measurements, each with four outcomes represented by four orthogonal unit vectors in $H_2 \otimes H_2$:

C_{00} : Perform a global polarization measurement on the photon pair with four outcomes corresponding to the basis $\left\{|X_0^{A_0}\rangle \otimes |Y_0^{B_0}\rangle, |X_0^{A_0}\rangle \otimes |Y_1^{B_0}\rangle, |X_1^{A_0}\rangle \otimes |Y_0^{B_0}\rangle, |X_1^{A_0}\rangle \otimes |Y_1^{B_0}\rangle\right\}$

C_{01} : Perform a global polarization measurement on the photon pair with four outcomes corresponding to the basis $\left\{|X_0^{A_0}\rangle \otimes |Y_0^{B_1}\rangle, |X_0^{A_0}\rangle \otimes |Y_1^{B_1}\rangle, |X_1^{A_0}\rangle \otimes |Y_0^{B_1}\rangle, |X_1^{A_0}\rangle \otimes |Y_1^{B_1}\rangle\right\}$

C_{10} : Perform a global polarization measurement on the photon pair with four outcomes corresponding to the basis $\left\{|X_0^{A_1}\rangle \otimes |Y_0^{B_0}\rangle, |X_0^{A_1}\rangle \otimes |Y_1^{B_0}\rangle, |X_1^{A_1}\rangle \otimes |Y_0^{B_0}\rangle, |X_1^{A_1}\rangle \otimes |Y_1^{B_0}\rangle\right\}$

C_{11} : Perform a global polarization measurement on the photon pair with four outcomes corresponding to the basis $\left\{|X_0^{A_1}\rangle \otimes |Y_0^{B_1}\rangle, |X_0^{A_1}\rangle \otimes |Y_1^{B_1}\rangle, |X_1^{A_1}\rangle \otimes |Y_0^{B_1}\rangle, |X_1^{A_1}\rangle \otimes |Y_1^{B_1}\rangle\right\}$

Note that these global polarization measurement require a complicated arrangement of beam splitters, polarizing beam splitters, wave plates, photo detectors and other non-linear optical devices (Mattle et al., 1996; Lütkenhaus et al., 1999; Weihs and Zeilinger 2001). What is important, is that C_{00} is *not* simply performing a linear polarization measurement on the left photon along axis a_0 and performing a linear polarization measurement on the right photon along a given transverse axis b_0. In other words, C_{00} is *not* the same measurement as $A_0 \wedge B_0$. They are operationally equivalent but not the same. Consequently, $\sum_2 C_{00}$ will not be the same as A_0; they will be only operationally equivalent.

This new operational theory has a trivial compatibility structure: C_{01} and C_{11} cannot be performed simultaneously, that is, they cannot be performed on the same pair of photons. Consequently, any ontological model for the theory is (trivially) simultaneously noncontextual. But the model we provided will be measurement contextual: some ontic states will provide different outcomes for the C_{01} and C_{11} contrary to the fact that $\sum_1 C_{01}$ and $\sum_1 C_{11}$ are operationally equivalent. Note, however, that measurement contextuality does not lead to the violation of local causality.

In the Introduction, we discerned the Bridgmannian and the standard identification of observables. In the first case, we identified observables with operators, in the second, with measurements. Applying this distinction to the EPR-Bell scenario, one gets the following schema:

$$
\begin{tikzcd}
 & & \textit{Bridgmannian} & & & \textit{Standard} & \\
\textit{Operator:} & & \mathbf{A_0} \arrow[dl] \arrow[dr] & & & \mathbf{A_0} \arrow[d] & \\
\textit{Observable:} & \mathcal{O}_1 \arrow[d] & & \mathcal{O}_2 \arrow[d] & & \mathcal{O} \arrow[dl] \arrow[dr] & \\
\textit{Measurement:} & A_0 & & \sum_2 C_{00} & A_0 & & \sum_2 C_{00}
\end{tikzcd}
$$

The local and global measurements are represented by the same operator in quantum mechanics. But do they measure the same observable? According to the standard interpretation yes; according to the Brigdmannian interpretation: no.

Conclusions

Quantum mechanics, at least in the minimalist interpretation, is an operational theory in a special linear algebraic representation. A distinctive feature of this theory is operational equivalence, the representation of different (and not necessarily simultaneous) measurements providing the same outcome statistics in every preparation by the same self-adjoint operator (or POVM). From the perspective of strict operationalism, the identity of the representation of such measurements does not mean the identity of the measured observables. In this paper, I intended to explore some of the consequences of abandoning operational equivalence in quantum theory and in general operational theories. We saw, how some of the main properties of the underlying ontological models will change if some measurements are replaced with other operationally equivalent measurements. To illustrate this change, we took the example of the EPR-Bell scenario and compared the ontological models of the non-trivial and the trivial operational theories realizing the EPR-Bell scenario by local and global measurements, respectively. The EPR-Bell situation, however, was not peculiar whatsoever, we could have equally well used the GHZ or the Peres-Mermin case to this goal. The four commuting operators in the horizontal line of the GHZ pentagram

$$\boldsymbol{\sigma}_z \otimes \boldsymbol{\sigma}_z \otimes \boldsymbol{\sigma}_z \qquad \boldsymbol{\sigma}_z \otimes \boldsymbol{\sigma}_x \otimes \boldsymbol{\sigma}_x \qquad \boldsymbol{\sigma}_x \otimes \boldsymbol{\sigma}_z \otimes \boldsymbol{\sigma}_x \qquad \boldsymbol{\sigma}_x \otimes \boldsymbol{\sigma}_x \otimes \boldsymbol{\sigma}_z$$

or the three commuting operators in the third column of the Peres-Mermin square

$$\boldsymbol{\sigma}_z \otimes \boldsymbol{\sigma}_z \qquad \boldsymbol{\sigma}_y \otimes \boldsymbol{\sigma}_y \qquad \boldsymbol{\sigma}_x \otimes \boldsymbol{\sigma}_x$$

can also be represented both by local measurements on individual photons (represented by the graphs in the Figure 1) and also by complicated global GHZ or Bell state measurements on photon pairs or triples (represented by the linegraphs in the Figure 3). These local and global measurements are different and so are the ontological models. All ontological models will be measurement contextual, but those for global measurements will be simultaneously noncontextual and will have a trivial causal structure. All these results point in the same direction which is also the main message of this paper: Operationally equivalent families of measurements represented by the same operators in quantum mechanics can give rise to ontological models with highly different features. Thus, to study these models, it is not enough to simply investigate quantum mechanics at an abstract mathematical level; we also need to take into consideration the measurements represented by the operators. This is the lesson that we can learn from Bridgman.

Appendix

Three outcome-deterministic ontological model for the three operational theories with non-trivial compatibility structure:

Classical theory.

- Set of ontic states: $\mathcal{L} = \{\Lambda_0, \Lambda_1\}$
- Random variable on $\mathcal{L}$: $\Lambda = 0, 1$
- Probability distribution:

$$p(\Lambda|P_{\text{CL}}) = \frac{1}{2} \tag{41}$$

- Response functions of the non-trivial theory:

$$p(X|A,\Lambda) = \delta_{X,\Lambda} \tag{42}$$
$$p(Y|B,\Lambda) = \delta_{Y,\Lambda} \tag{43}$$
$$p(X,Y|A,B,\Lambda) = \delta_{X,\Lambda} \cdot \delta_{Y,\Lambda} \tag{44}$$

  where δ is the Kronecker delta function.

- Response functions of the trivial theory:

$$p(Z|C,\Lambda) = \delta_{Z_1,\Lambda} \cdot \delta_{Z_2,\Lambda} \tag{45}$$

The EPR-Bell scenario.

- Set of ontic states: $\mathcal{L} \times \mathcal{L}$ where $\mathcal{L} = \{\Lambda_0, \Lambda_1\}$
- Random variable on $\mathcal{L} \times \mathcal{L}$: $\Lambda_1 \times \Lambda_2$ with $\Lambda_1, \Lambda_2 = 0, 1$
- Probability distribution:

$$p(\Lambda_1, \Lambda_2 | P_{\text{EPR}}) = \begin{cases} \frac{1}{8} & \text{if } \Lambda_1 \oplus \Lambda_2 = 1 \\ \frac{3}{8} & \text{otherwise} \end{cases} \tag{46}$$

- Response functions of the non-trivial theory:

$$p(X|A, \Lambda_1, \Lambda_2) = \delta_{X,\Lambda_1} \tag{47}$$
$$p(Y|B, \Lambda_1, \Lambda_2) = \delta_{Y,\Lambda_2} \tag{48}$$
$$p(X, Y|A, B, \Lambda_1, \Lambda_2) = \delta_{X,\Lambda_1} \cdot (\delta_{Y \oplus (A \cdot B),\Lambda_2} \cdot \delta_{\Lambda_1 \oplus \Lambda_2,1} + \delta_{Y,\Lambda_2} \cdot \delta_{\Lambda_1 \oplus \Lambda_2,0}) \tag{49}$$

- Response functions of the trivial theory:

$$p(Z|C, \Lambda) = \delta_{Z_1,\Lambda_1} \cdot (\delta_{Z_2 \oplus (C_1 \cdot C_2),\Lambda_2} \cdot \delta_{\Lambda_1 \oplus \Lambda_2,1} + \delta_{Z_2,\Lambda_2} \cdot \delta_{\Lambda_1 \oplus \Lambda_2,0}) \tag{50}$$

PR box.

- Set of ontic states: $\mathcal{L} = \{\Lambda_0, \Lambda_1\}$
- Random variable on $\mathcal{L}$: $\Lambda = 0, 1$
- Probability distribution:

$$p(\Lambda | P_{\text{PR}}) = \frac{1}{2} \tag{51}$$

- Response functions of the non-trivial theory:

$$p(X|A, \Lambda) = \delta_{X,\Lambda} \tag{52}$$
$$p(Y|B, \Lambda) = \delta_{Y,\Lambda} \tag{53}$$
$$p(X, Y|A, B, \Lambda) = \delta_{X,\Lambda} \cdot \delta_{Y \oplus (A \cdot B),\Lambda} \tag{54}$$

- Response functions of the trivial theory:

$$p(Z|C, \Lambda) = \delta_{Z_1,\Lambda} \cdot \delta_{Z_2 \oplus (C_1 \cdot C_2),\Lambda} \tag{55}$$

References

Abramsky S., and A. Brandenburger, (2011). The sheaf-theoretic structure of non-locality and contextuality, New J. Phys, 13, 113036.

Abramsky, S., R. S. Barbosa, K. Kishida, R. Lal, S. Mansfield, (2017). Contextuality, cohomology and paradox, URL = https://arxiv.org/abs/1502.03097.

Bridgman, P. W. (1958). The Logic of Modern Physics (New York: The Macmillan Company).

Cavalcanti, E. (2018). Classical Causal Models for Bell and Kochen-Specker Inequality Violations Require Fine-Tuning, Phys. Rev. X, 8, 021018.

Chang, H. (2019). Operationalism, Stanford Encyclopedia of Philosophy. URL = https://plato.stanford.edu/entries/operationalism.

Clauser, J. F., M.A. Horne, A. Shimony and R. A. Holt, (1969). Proposed experiment to test local hidden-variable theories, Phys. Rev. Lett., 23, 880-884.

Glymour, C., Scheines, R., and Spirtes, P. (2001). Causation, Prediction, and Search, (Cambridge: The MIT Press).

Greenberger, D. M., Horne, M. A., Shimony, A. és Zeilinger, A. (1990). Bell's theorem without inequalities, Am. J. Phys. 58, 1131

Hofer-Szabó, G. (2021a). Commutativity, comeasurability, and contextuality in the Kochen-Specker arguments, Phil. Sci., 88, 483-510.

Hofer-Szabó, G. (2021b). Two concepts of noncontextuality in quantum mechanics, (submitted).

Hofer-Szabó, G. (2021c). Three noncontextual hidden variable models for the Peres-Mermin square, Eur. J. Phil. Sci., 11, 30.

Kochen, S., and E. P. Specker (1967). The problem of hidden variables in quantum mechanics, J. Math. Mech., 17, 59–87.

Lütkenhaus, N., Calsamiglia, J., and Suominen, K-A. (1999). On Bell measurements for teleportation, Phys. Rev. A, 59, 3295.

Mattle, K., Weinfurter, H., Kwiat, P. G., and Zeilinger A. (1996). Dense Coding in Experimental Quantum Communication, Phys. Rev. Lett, 76 (25), 4656-4659.

Mermin, D. (1993). Ontological states and the two theorems of John Bell, Rev. Mod. Phys., 65 (3), 803-815.

Pearl, J. (2009). Causality: Models, Reasoning, and Inference, (Cambridge: Cambridge University Press)

Peres, A. (1990). Incompatible Results of Quantum Measurements," Phys. Lett. A, 151, 107-108.

Popescu, S., and D. Rohrlich. (1994). Nonlocality as an axiom, Found. Phys., 24, 379-385.

Schlick, M. (1930 [1979]). On the Foundations of Knowledge, in Philosophical Papers, vol. 2 (1925–1936), H. L. Mulder and B. F. B. van de Velde-Schlick (eds.), Dordrecht: Reidel, pp. 370–387.

Spekkens, R. W. (2005). Contextuality for preparations transformations and unsharp measurements, Phys. Rev. A 71:052108.

Suarez, M. and SanPedro, I. (2009). Causal Markov, Robustness and the Quantum Correlations, in: M. Suarez (ed.), Probabilities, Causes and Propensities in Physics, Synthese Library, Springer, Ch. 8., 173-193.

Weihs, G., and Zeilinger, A. (2001). Photon statistics at beam splitters: an essential tool in quantum information and teleportation, in: Jan Perina (ed), Coherence and Statistics of Photons and Atoms, (John Wiley and Son, Inc, New York) 262-288.

Wood, C. J., and Spekkens, R. W. (2015). The lesson of causal discovery algorithms for quantum correlations: causal explanations of Bell-inequality violations require fine-tuning, New J. Phys. 17, 033002.