On Joint Distributions, Counterfactual Values, and Hidden Variables in Understanding Contextuality

Ehtibar Dzhafarov
Purdue University

Abstract

This paper deals with three traditional ways of defining contextuality: (C1) in terms of (non)existence of certain joint distributions involving measurements made in several mutually exclusive contexts; (C2) in terms of relationship between factual measurements in a given context and counterfactual measurements that could be made if one used other contexts; and (C3) in terms of (non)existence of "hidden variables" that determine the outcomes of all factually performed measurements. It is generally believed that the three meanings are equivalent, but the issues involved are not entirely transparent. Thus, arguments have been offered that C2 may have nothing to do with C1, and the traditional formulation of C1 itself encounters difficulties when measurement outcomes in a contextual system are treated as random variables. I show that if C1 is formulated within the framework of the Contextuality-by-Default (CbD) theory, the notion of a probabilistic coupling, the core mathematical tool of CbD, subsumes both counterfactual values and "hidden variables". In the latter case, a coupling itself can be viewed as a maximally parsimonious choice of a hidden variable.

1 Introduction

The aim of this paper is to consider three historically established ways of understanding (non)contextuality, and relate them to each other from the vantage point of the Contextuality-by-Default (CbD) theory. The reader unfamiliar with CbD can find its latest version in Ref. [18] (and additional details, arguments, and proofs in Refs. [13, 14, 16, 32, 34]). My main point is that the three approaches in question can be viewed as variants or interpretations of the core mathematical tool of CbD — probabilistic couplings.

The first of the three meanings of contextuality considered in this paper can be called *joint-distributional*. It was introduced by Suppes and Zanotti [41] and Fine [22], within the conceptual framework of establishing its equivalence to the more traditional at the time "hidden-variable" meaning of contextuality, as discussed below. Contexts are defined as conditions, or arrangements under which one performs one's measurements (including but not reduced to what other measurements are performed together with a given one). A system of measurements made in varying, mutually exclusive contexts is noncontextual if the random variables representing all these measurements can be considered jointly distributed. This seems to have become a common way of understanding contextuality, as evidenced by numerous contemporary works [1–3, 9, 10, 30, 35, 36, 39]. The CbD theory belongs to the same category, and its specific feature is that all measurement outcomes are consistently treated as contextually labelled random variables. The use of contextual labeling means that measurements made in different contexts are always represented by different random variables, even if they measure the same property. As a result, the possibility of imposing a joint distribution on all random variables in a system is not a restrictive requirement. In CbD, therefore, one is interested in the existence of not just some but only specific joint distributions, those satisfying certain requirements. To match the traditional understanding, derived from Fine's and Suppes and Zanotti's work, the requirement should be that, in the joint distribution imposed on the system, the variables measuring the same property in different contexts should always have equal values (which is achievable only if the random variables in play satisfy the no-disturbance constraint, as discussed below).

The second meaning of contextuality can be called *counterfactual*. It is formulated in terms of whether an outcome of a factual measurement made in some context would have been the same had it been made in another context. It seems that most contemporary researchers take this counterfactual formulation as being equivalent to the joint-distributional meaning of contextuality mentioned above (and to the "hidden-variable" meaning, mentioned below). This is apparent, e.g., in Liang, Speckens, and Wisemen's comprehensive introduction to contextuality [39]. However, this view is not universally accepted. Thus, Griffiths [24] calls the counterfactual meaning of contextuality *Bell-contextuality*, and argues that any system is Bell-noncontextual (see Griffiths's paper in the present issue, [25]). This means that, in Griffiths's opinion, it is always true that had one measured a factually measured property in another context, the result would have been the same; and that this has nothing to do with the existence or nonexistence of joint distributions in the first meaning of contextuality.

The third way of defining (non)contextuality was historically the first. Contextuality (without using this term) was introduced in quantum physics through the notion of hidden variables, primarily by Bell [5, 6] and Kochen and Specker [31]. In particular, Bell demonstrated that one could meaningfully address, using only observable measurements, the question famously discussed in Bohr's [7] critique of Einstein, Podolsky, and Rosen [21]. This question is whether all measurement outcomes in a system of measurements can be presented as being determined by some "hidden" random variable in a context-independent way, i.e., using context-independent mappings of the values of this hidden variable into the values of the observed measurement outcomes. The question has beens historically formulated in terms of "realism", the existence of hidden variables of which all observable outcomes of measurements are functions, and "(non)locality", the (in)dependence of these functions on the contexts. (To include systems in which spatial separation plays no role, e.g., of the Kochen-Specker variety, the term "locality" should be replaced with the broader term "context-independence".) Quantum mechanics is usually said to exclude the conjunction of realism and context-independence, but the culprit in this conjunction is not agreed on by all. Thus, Leggett [37] and Gröblacher, Paterek, Kaltenbaek, Brukner, Żukowski, Aspelmeyer, and Zeilinger [26] show that quantum mechanics rules out realism in conjunction with certain forms of context-dependent mapping.

In this paper, I uphold the prevalent view that all three meanings of contextuality are equivalent. We will see that counterfactual definiteness and the existence of hidden variables can be viewed as philosophically and/or physically laden ways of speaking of probabilistic couplings, the notion that lies at the heart of CbD.

CbD is usually taken to be useful for *inconsistently connected systems* (systems with "disturbance"), where measurements of the same property in different contexts may have differently distributed outcomes [4, 12, 23, 34, 40, 44]. However, if measurements are treated as random variables within the framework of classical probability theory, CbD offers considerable conceptual clarity even for consistently connected systems, those with no "disturbance". To illustrate this, I focus on such systems throughout most of this paper.

The following three aspects of our discussion should be kept in mind. First, the term *measurement* can be replaced with any procedure with generally random outcomes, e.g., responses of a biological organism to stimuli. For this reason I prefer in the following to use the standard CbD term "content" (of a random variable) in place of the "property being measured". Second, the analysis is entirely within the framework of classical probability theory, with classical understanding of random variables. In particular, random variables may but need not be related to observables in Hilbert space. Third, the arguments I present in favor of using CbD should not be misconstrued as criticism of other contemporary approaches to (non)contextuality, such as presented in Refs. [2, 3, 9, 29, 39]. In particular, the terms "traditional" and "historical" used in describing positions contrasted with CbD refer primarily to the literature of the last century.

2 Preliminaries: Terminology and notation

Consider an experiment consisting in measuring several properties, generically called *contents*, under various conditions, called *contexts*. The contents form a set Q, the contexts form a set C, and in each context $c \in C$ one jointly measures some subset Q_c of the properties Q, with $Q_c \cap Q_{c'}$ generally nonempty. If $q \in Q_c$, the result of measuring this content q in a context c is a random variable R_q^c. The set of the random variables double-labeled in this way is a *system* (of random variables).

The random variables belonging to the same context form a set of jointly distributed random variables

$$\left\{R_q^c : q \in Q_c\right\}. \tag{1}$$

Conceptually, the joint distribution of context-sharing variables means that they can be presented as measurable functions on one and the same *domain probability space* (sometimes also referred to as *sample space*). Equivalently, and more conveniently for our purposes, this means that one can choose a random variable H_c for each context c, and functions f_q^c , such that

$$f_q^c\left(H_c\right) = R_q^c,\ q \in Q_c, c \in C. \tag{2}$$

If the sample space for the context-sharing variables in (1) is specified, H_c can be chosen as the identity function on the same space, in which case f_q^c is the measurable function defining R_q^c. However, a practical, and most economic, choice of H_c is (1) itself, which is a random variable in its own right.[1] With this choice, f_q^c is the componentwise projection function.

By contrast with (1), any two random variables picked from different contexts are *stochastically unrelated*, even if sharing a content: H_c and $H_{c'}$ for $c \neq c'$ have no joint distribution (are defined on distinct sample spaces).

As mentioned above, we focus on *consistently connected systems*, defined by the following property: if $q \in Q_c \cap Q_{c'}$ $(c, c' \in C)$, then the distribution of R_q^c is the same as the distribution of $R_q^{c'}$. This is written as

$$R_q^c \sim R_q^{c'}. \tag{3}$$

That is, a random variable's distribution in a consistently connected system is determined by the variable's content only.[2]

[1] As mentioned in Section 5, the distinction between a "single" variable and a set of jointly distributed variables is purely representational and dispensable.

[2] This property is known in physics under a variety of names, such as no-signaling, no-disturbance, parameter

As an example I will use the following system:

$$\begin{array}{|c|c|c||c}
\hline
R_1^1 & R_2^1 & & c=1 \\
\hline
 & R_2^2 & R_3^2 & c=2 \\
\hline
R_1^3 & & R_3^3 & c=3 \\
\hline\hline
q=1 & q=2 & q=3 &
\end{array}. \tag{4}$$

Consistent connectedness means here that $R_1^1 \sim R_1^3$, $R_2^1 \sim R_2^2$, and $R_3^2 \sim R_3^3$. Because the two variables within each context are measured together, they are *operationally* (or *empirically*) jointly distributed.[3] By contrast, random variables picked from different contexts, say R_1^1 and R_3^3 or R_1^1 and R_1^3, are *stochastically unrelated*: there is no empirical procedure for pairing their values; they can be paired "on paper", but not uniquely, with no particular way being privileged.

I will use the example of system (4) in the remainder, sometimes mentioning and sometimes only implying a generalization to any system of random variables. However, the generalization is trivial only if we confine our discussion to systems of random variables with finite numbers of contents and contexts, and to random variables that are *categorical*, i.e., have finite numbers of values. The latter is not a restriction for the CbD approach, in which all random variables should be replaced with sets of jointly distributed *dichotomous* random variables before contextuality analysis can be applied [18] (we need not, however, discuss this construction in this paper).

3 Joint-distributional understanding: Identically connected couplings

The first meaning of contextuality is based on the obvious fact that one can consider a multitude of ways the six random variables in (4) could be jointly distributed "on paper" (knowing that they are not jointly distributed de facto). This formulation (with the words "could be") is hinting at counterfactuality, but we need not go that way: mathematically, we simply consider all sextuples of jointly distributed random variables

$$\begin{array}{|c|c|c|}
\hline
S_1^1 & S_2^1 & \\
\hline
 & S_2^2 & S_3^2 \\
\hline
S_1^3 & & S_3^3 \\
\hline
\end{array}, \tag{5}$$

with the same row-wise distributions as in (4). Any such a sextuple is a *probabilistic coupling* of the system (4), and the set of possible couplings is always nonempty. More generally, given a set $\mathcal{X}$ of random variables, its coupling is defined as a set X of jointly distributed random variables, in a bijective correspondence with $\mathcal{X}$, such that for any $\mathcal{Y} \subseteq \mathcal{X}$, if the elements of $\mathcal{Y}$ are

invariance, etc. [11]. In psychology it is known under the name of simple marginal selectivity. The term "consistent connectedness" is less connotationally loaded than most of these terms. It is a naturally formed term in CbD, because contexts in this theory are represented by isolated "islands" of random variables, and they are only "connected" to each other by the fact that some variables in different contexts have the same content. The set of all variables R_q^c with the same content q is then natural to call a *connection* [15].

[3] The term "measured together" usually means simultaneously, but it could be any empirical scheme by which the values of two variables are paired. They can be measured at different times, as in the Leggett-Garg-type experiments [9,38], and they can be related to the same or different physical objects. This is one reason my analysis is not critically related to such issues as commutativity of the quantum operators generating the random variables.

jointly distributed, then the corresponding subset $Y \subseteq X$ has the same distribution. In particular, corresponding elements of $\mathcal{X}$ and X are identically distributed.[4]

In contextuality analysis of a consistently connected system one is interested in whether among all possible couplings of a system of random variables one can find one with a special property. This property is that in each connection, i.e. each column of (5), the two random variables are equal to each other with probability 1. If such a coupling exists (in CbD it is called *identically connected*), then the system (4) is considered *noncontextual*. Otherwise it is *contextual*. Thus, if the variables in the system (4) are dichotomous, $+1/-1$, then it is known [3, 32] that this system is contextual if and only if

$$\max_{\text{odd } \# \text{ of } -\text{'s}} \left(\pm \left\langle R_1^1 R_2^1 \right\rangle \pm \left\langle R_2^2 R_3^2 \right\rangle \pm \left\langle R_3^3 R_1^3 \right\rangle\right) > 1, \tag{6}$$

where the maximum is taken over all combinations with odd numbers of minus signs (1 or 3). Examples of such systems are readily constructed.

Traditionally, the system (4) would be presented as

R_1	R_2		$c=1$
	R_2	R_3	$c=2$
R_1		R_3	$c=3$
$q=1$	$q=2$	$q=3$	

, (7)

with overlapping sets of random variables (here, each random variable occurs in two different contexts). This is the case, e.g., in what seems to be historically very first joint-distributional analysis of contextuality (without using this term), the 1981 paper by Suppes and Zanotti [41]. The central theorem in that paper says (mutatis mutandis):[5]

> Suppes-Zanotti's Theorem. Let R_1, R_2, R_3 be random variables with possible values 1 and -1. Then a necessary and sufficient condition for the existence of a joint probability distribution of the three random variables is
>
> $$\max_{\text{odd } \# \text{ of } -\text{'s}} \left(\pm \langle R_1 R_2 \rangle \pm \langle R_2 R_3 \rangle \pm \langle R_3 R_1 \rangle\right) \leq 1. \tag{8}$$

While the necessity part of this statement is straightforward, the sufficiency part encounters difficulties. The reason for this is that the relation of being jointly distributed is "agglutinative", in the following sense:

(Agglutinativity) given sets A and B of jointly distributed random variables, if $A \cap B \neq \emptyset$, then $A \cup B$ is a set of jointly distributed random variables.[6]

This property holds essentially by definition of a random variable. It follows that for R_1, R_2, R_3 to be jointly distributed it is sufficient that at least two of the expected values $\langle R_1 R_2 \rangle, \langle R_2 R_3 \rangle, \langle R_3 R_1 \rangle$

[4] This definition is modified with respect to the standard one [42] to better suit contextuality analysis.

[5] Compared to the original formulation, notation is changed, an unnecessary constraint is removed, and the inequality is replaced with an equivalent one to make it comparable to (6).

[6] In a previous publication [16] this property was erroneously called "transitivity". When applied to three random variables, X, Y, Z, transitivity means that joint distributions of (X, Y) and (Y, Z) implies that of (X, Z). Agglutinativity means that joint distribution of (X, Y) and (Y, Z) implies that of (X, Y, Z). This implies transitivity but is not equivalent to it.

be well-defined, i.e. at least two of the pairs $(R_1, R_2), (R_2, R_3), (R_3, R_1)$ be jointly distributed. But the latter would be the case even if

$$\max_{\text{odd } \# \text{ of } -\text{'s}} (\pm \langle R_1 R_2 \rangle \pm \langle R_2 R_3 \rangle \pm \langle R_3 R_1 \rangle) > 1, \tag{9}$$

in which case R_1, R_2, R_3 cannot be jointly distributed. This contradiction shows that a correct formulation of Suppes-Zanotti's theorem should have been as follows:

> Let R_1, R_2, R_3 be random variables with possible values 1 and -1. Then a necessary and sufficient condition for the existence of a joint probability distribution of the three random variables is the existence of a joint distribution of any two of the three pairs $(R_1, R_2), (R_2, R_3), (R_3, R_1)$. If this is the case, (8) is satisfied.

However, then it follows that (9) cannot ever hold. Put differently, if (7) is a *system* (implying, in particular, that the joint distributions within contexts are well-defined), then this system can only be noncontextual. Therefore, if it happens that (9) holds for this system, then one has a true contradiction on one's hands, and this contradiction cannot be resolved within the framework of (7). It can only be resolved by explicating and rejecting some hidden assumptions – and in this case the culprit is the assumption that the random variables measuring the same content in different contexts are the same.

This problem is ubiquitous in the traditional literature succeeding Ref. [41], although its critical analysis is complicated by the fact that many authors would refer to elements of (7) as measurements or observables rather than random variables. I take it as a given, however, that the notion of a distribution of R_i, or the probability of R_i being equal to some value, can only be used if R_i is a random variable. With this in mind, the contradiction just described can only be resolved by using contextual notation, and CbD offers a straightforward way of doing this. However, contextual notation can also be applied to probabilities rather than random variables per se, and this seems to be the way chosen in some of the contemporary literature. Thus, Khrennikov [28, 29] proposes labeling of the form $\Pr[R_i = r \mid c = j]$, calling this "contextual probabilities" and warning against identifying them with conditional probabilities. Liang, Speckens, and Wisemen [39] use essentially the same notation (if one considers quantum preparations part of contexts). Abramsky and colleagues [1, 2] developed a similar system in which e_c denotes the joint probability of all random variables in context c, and distributions of their subsets are treated as specializations: e.g., the distribution of R_1 in $c = 3$ would be denoted $e_{c=3}|_{q=1}$. Contextual notation for probabilities, the same as CbD's contextual notation for random variables, allows one to avoid the difficulties related to the agglutinativity.

4 Counterfactual approach

The second meaning of (non)contextuality is predicated on an affirmative answer to the following question:

(Q1: counterfactual definiteness) when one makes measurements in a given context, can one meaningfully speak of what the outcomes of measurements would have been had one chosen another context?

Using our example system (4), if the chosen context is $c = 2$, one records the values of R_2^2 and R_3^2. Is it meaningful to ask what the recorded values would have been had we chosen $c = 1$ or $c = 3$ instead of $c = 2$? If the answer is negative, there is nothing more to discuss, and the counterfactual meaning of contextuality cannot be formulated. A positive answer means that whenever a measurement is being made, *all random variables in all contexts can be thought of as having definite values.* It makes no difference for any possible consequences whether this assignment of values is understood epistemologically (the experimenter can always assign these counterfactual values, perhaps not uniquely) or ontologically (the random variables in counterfactual contexts have true values, but being unknown in principle, one should consider possibilities, perhaps more than one, of what these "true" values could be).

Assuming the answer to Q1 is positive, one can ask the next, critical question:

(Q2: counterfactual identity) is it possible that in this assignment of values to all random variables in the system any two content-sharing random variables R_q^c and $R_q^{c'}$ be always assigned the same value?

If this question, too, is answered in the affirmative, i.e., if there is an assignment of values that only depends on the variables' contents, rather than also on their contexts, the system is deemed noncontextual in the "counterfactual sense".

Note: in our understanding of Q2, the variables R_q^c and $R_q^{c'}$ measure the same content in two distinct contexts, one of which may be, but need not be factual. One might object to this and argue that one should only be interested in this question if one of the contexts c, c' is the factual context. A comparison of two counterfactual assignments, one might insist, is of no interest. I can see no convincing justification for imposing this restriction. Using our example (4), if one meaningfully contemplates a counterfactual value of R_2^1 when the measurements are factually made in context $c = 2$, one should also be able to consider counterfactual value of R_1^1: after all, R_2^1 and R_1^1 are jointly distributed, it may even be the case that one of them is uniquely determined by the other, e.g., $R_2^1 = -R_1^1$. But then, if one can speak of the value that R_1^1 would have had if one measured $q = 1$ in context $c = 1$ instead of measuring $q = 2$ in context $c = 2$, then one can also speak of what the value of R_1^3 would have been if the same $q = 1$ were measured in context $c = 3$. And the requirement that the assignment of values should only depend on content would then dictate that R_1^3 be assigned the same value as R_1^1.

The picture we arrive at now is: irrespective of what factual measurements are made, all random variables are assigned values, and we ask whether this can be done so that any two R_q^c and $R_q^{c'}$ be assigned the same value. Assume that this is possible, i.e., the system is noncontextual in the counterfactual sense. Then R_q^c and $R_q^{c'}$, being always equal to each other, are jointly distributed. By the agglutinativity of the relation of being jointly distributed, this means that all random variables in the system are jointly distributed. This is obvious in our example, where the identity requirement across contexts together with the joint distributions within contexts yields the following graph of pairwise jointly distributed random variables:

$$
\begin{tikzcd}
R_1^1 \arrow[r, dash] \arrow[dd, dash] & R_2^1 \arrow[d, dash] & \\
& R_2^2 \arrow[r, dash] & R_3^2 \arrow[d, dash] \\
R_1^3 \arrow[rr, dash] & & R_3^3
\end{tikzcd}
\quad . \qquad (10)
$$

Any path in this graph that includes all nodes suffices to establish that the random variables are jointly distributed. Even easier, consider contexts as nodes of a graph and connect two contexts by an edge if they involve at least one common content. In our example it would look like this:

$$\begin{tikzcd} c_1 \arrow[rr, dash] \arrow[dr, dash] & & c_2 \arrow[dl, dash] \\ & c_3 & \end{tikzcd} . \tag{11}$$

If such a graph is constructed for an arbitrary system, and if, as in our example, it is *connected*, then in this system all random variables are jointly distributed. Systems whose context graphs are not connected do not pose difficulties as they can (and should) always be studied as several unrelated to each other systems with connected graphs.

It is now obvious that a system that is noncontextual (or contextual) in the the counterfactual sense is precisely a system for which there exists (respectively, does not exist) an identically connected coupling. The two meanings coincide. One can even say that the notion of a coupling is nothing but a rigorous mathematical meaning of the "counterfactual sense". Instead of making the assumption, some would say distinctly metaphysical in flavor, that a random variable has a definite value even if unmeasured, one can state as a fact, with no assumptions involved, that one can impose a joint distribution on (construct a coupling of) all random variables in the system. And then one can ask whether among all such couplings one can find an identically connected one. Counterfactual statements can be rigorously formalized, but most would agree that their logical status is more involved than that of factual statements. "The system has a coupling with properties X" is more "ordinary" a statement than "Had we measured A in another context its value would have been x." As stated in a Stanford Encyclopedia of Philosophy article [43], "Philosophers, linguists, and psychologists remain fiercely divided on how to best understand counterfactuals." It is hard to be divided over the notion of a coupling.

5 Hidden variables with context-independent mapping

As mentioned in Section 2, several random variables are jointly distributed if and only if they are functions of one and the same random variable. Thus, because in each content c all random variables R_q^c $(q \in Q_c)$ are jointly distributed in the operational sense (measured "together"), one can define a random variable H_c and functions f_q^c $(q \in Q_c)$ such that (2) holds. This is a context-specific hidden-variable construction, and it is obviously nonrestrictive, applicable to any system. The third meaning of (non)contextuality we are discussing now is about the possibility of a single hidden variable for all contexts, and context-independent functions that map it into the random variables comprising the system. In other words, one asks whether there is a random variable H and a set of functions f_q $(q \in Q)$ such that for any $c \in C$ and $q \in Q_c$,

$$R_q^c = f_q(H) \,. \tag{12}$$

If (and only if) such a construction is possible, the system is noncontextual in the "hidden variable" sense.

Applied to our example, the question is about the possibility of replacing (4) with

$$\begin{array}{|c|c|c||c|}
\hline
f_1(H) & f_2(H) & & c=1 \\
\hline
 & f_2(H) & f_3(H) & c=2 \\
\hline
f_1(H) & & f_3(H) & c=3 \\
\hline\hline
q=1 & q=2 & q=3 &
\end{array}. \tag{13}$$

At every measurement, the variable H has some value, and all six random variables, irrespective of what context c is being factually measured, are determined by this value. At that, the values of any two random variables measuring the same content are always equal (because the functions are labeled by their contents only). It is easy to see that this is precisely the same as the existence, for any factual measurement, of the assignment of values to all random variables in the system, such that $R_q^c = R_q^{c'}$ for any $q \in Q_c \cap Q_{c'}$. Of course, in CbD, (12) has to be replaced with

$$R_q^c \sim f_q(H) = S_q, \tag{14}$$

i.e. (13) represents a coupling of the system rather than the system itself.

This completes the demonstration that the three meanings of (non)contextuality considered in this paper are subsumed by the notion of a coupling in the CbD sense.

There is also a "stochastic" version of the hidden-variable hypothesis, in which each variable R_q^c is a function f_q of a common source of randomness H and some specific source of randomness V_q (context-independent),

$$R_q^c \sim f_q(H, V_q). \tag{15}$$

This version, however, is immediately reduced to the previous, "deterministic" version (12), on renaming

$$\left(H, \{V_q\}_{q\in Q}\right) \mapsto H'. \tag{16}$$

The existence of the single underlying random variables H is sometimes referred to as "realism", whereas the context-independence of the mappings f_q is the generalization of what is traditionally referred to as "locality" (when applied to spatially distributed systems of particles). It is worth noting that of these two requirements for a hidden variable theory, it is only context-independence of mappings that has a restrictive effect. Indeed, the variables H_c in (2), of which we know that it is universally applicable, can always be coupled in a variety of ways, e.g. independently. In our example, let us replace $H_{c=1}, H_{c=2}, H_{c=3}$ with a triple of jointly distributed, e.g. independent, random variables $(G_{c=1}, G_{c=2}, G_{c=3})$ such that $G_c \sim H_c$. Being jointly distributed, $G_{c=1}, G_{c=2}, G_{c=3}$ can be presented as functions of some random variables G,

$$G_c = g_c(G), \ c = 1, 2, 3.$$

This variable G can be chosen, e.g., as the vector

$$G = (G_{c=1}, G_{c=2}, G_{c=3}), \tag{17}$$

in which case

$$R_q^c \sim f_q^c(g_c(G)) = h_q^c(G), \ c = 1, 2, 3, q = 1, 2, 3, \tag{18}$$

where g_c is the cth projection function. Denoting

$$h_q^c(G) = S_q^c, \tag{19}$$

we form a coupling of the system. In accordance with [37], this "realist" construction is completely nonrestrictive, applicable to any system. This does not exclude the possibility that some special cases of context-dependent mapping too can be ruled out (e.g., by laws of quantum mechanics), and this was demonstrated in Refs. [37] and [26].

In terms of counterfactual values, the construction (18) does introduce them implicitly, adhering thereby to counterfactual definiteness. However, it does not require counterfactual identity, i.e., it is not necessary that $S_q^c = S_q^{c'}$ whenever $q \in Q_c \cap Q_{c'}$ (e.g., R_q^c and $R_q^{c'}$ are independent if so are $G_{c=1}, G_{c=2}, G_{c=3}$). The assumption of counterfactual identity is equivalent to that of context-independent mappings, this assumption is restrictive and may be empirically violated.

It is perhaps useful here to dispel the naive but not infrequent misconception that the random variable in "realist" representation must be "single", so that, e.g., G in (17) is not a legitimate choice. Any set of jointly distributed random variables can always be replaced with a "single" one. In the special case of a countable set of random variables defined on reals endowed with Borel sigma-algebra this random variable can always be chosen, if one so wishes, as a single variable uniformly distributed between 0 and 1 [8, 19, 27]. "Probabilistic dimensionality" (the number of components of a random entity) is entirely a matter of one's choice.

The obvious statement that jointly distributed random variables X and Y are functions of one and the same random variable Z becomes even more obvious if one realizes that Z can always be chosen as (X, Y). In particular, a coupling S of a system can itself be viewed as a hidden variable, and the most "economically chosen" one at that. Thus, in the classical proof of the Bell theorem [5], for three dichotomous random variables A, B, C, the hidden variable λ could be chosen, with no loss of generality, as the eight-valued $\lambda = (A, B, C)$. To see that this is the most economic choice, note that any λ such that $A = f_A(\lambda), B = f_B(\lambda), C = f_C(\lambda)$ can be presented as the jointly distributed quadruple (A, B, C, λ) of which A, B, C are the first three projections. Obviously, no choice of λ can make this quadruple simpler than eliminating λ altogether.

6 Conclusion

The language of probabilistic couplings used in CbD is a rigorous and parsimonious way of talking about counterfactuals and hidden variables with context-(in)dependent mapping. It is also conducive to expanding the sphere of applicability and depth of contextuality analysis [13–16, 18]. If an identically connected coupling of (4) does not exist, other couplings do, and one can profitably study this set of possible couplings, e.g. to compute the degree of contextuality. Using our example (4), if the probabilities with which $S_1^1 = S_1^3$, $S_2^1 = S_2^2$, and $S_3^2 = S_3^3$ in a coupling (5) cannot all be 1, one can be naturally interested in the maximal values of the sum

$$\Delta = \Pr\left[S_1^1 = S_1^3\right] + \Pr\left[S_2^1 = S_2^2\right] + \Pr\left[S_3^2 = S_3^3\right] \tag{20}$$

that can be achieved among all possible couplings (5). Then the difference $3 - \Delta$ will serve as a possible measure of contextuality of system (4). For arbitrary consistently connected systems, Δ is the sum of $\Pr\left[S_q^c = S_q^{c'}\right]$ for all (c, c', q) such that $q \in Q_c \cap Q_{c'}$, and the measure of contextuality is $N - \Delta$, where N is the number of all such (c, c', q).

As we know, the main motivation for developing CbD, was that the language of probabilistic couplings allows one to "smoothly" go beyond the class of consistently connected systems [13, 16–18, 20, 32, 34]. An arbitrary system $\left\{R_q^c : q \in Q_c, c \in C\right\}$ is considered noncontextual if and only if it has a coupling $\left\{S_q^c : q \in Q_c, c \in C\right\}$ in which, for any (c, c', q) with $q \in Q_c \cap Q_{c'}$, the probability of $S_q^c = S_q^{c'}$ is maximal possible. Denoting this sum of all these maximal values by Δ_0, the measure of contextuality mentioned above is generalized as $\Delta_0 - \Delta$. It is, of course, only one of many possible measures of contextuality, other measures being described, e.g., in Refs. [15,18]. An important extension of measures of contextuality into measures of noncontextuality, if a system is noncontextual, is addressed in another paper in the present issue [33].

Acknowledgements.

The author is grateful to the participants of the Purdue Winer Memorial Lectures 2018 for many fruitful discussions, and to one of the anonymous reviewers for constructive criticism.

References

[1] Abramsky S, Brandenburger A 2011 The sheaf-theoretic structure of non-locality and contextuality. New J. Phys. 13: 113036-113075.

[2] Abramsky S., Barbosa R. S., Mansfield S 2017 The contextual fraction as a measure of contextuality. Phys. Rev. Lett. 119: 050504.

[3] Araújo M, Quintino MT, Budroni C, Cunha MT, Cabello A 2013. All noncontextuality inequalities for the n-cycle scenario, Phys. Rev. A 88: 022118.

[4] Arias M, Canas G, Gomez ES, Barra JB, Xavier GB, Lima G, D'Ambrosio V, Baccari F, Sciarrino F, Cabello A 2015 Testing noncontextuality inequalities that are building blocks of quantum correlations. Phys. Rev. A 92: 032126.

[5] Bell J. 1964 On the Einstein-Podolsky-Rosen paradox. Physics, 1: 195-200.

[6] Bell J. 1966 On the problem of hidden variables in quantum mechanics. Rev. Mod. Phys. 38: 447-453.

[7] Bohr N. 1935 Can quantum-mechanical description of physical reality be considered complete? Phys. Rev., 48: 696-702.

[8] Brandenburger A, Keisler HJ 2018 A canonical hidden-variable space. Annals of Pure and Applied Logic 169: 1295-1302.

[9] Budroni C. 2016 Temporal Quantum Correlations and Hidden Variable Models. Springer: Heidelberg.

[10] Cabello A 2013 Simple explanation of the quantum violation of a fundamental inequality. Phys. Rev. Lett., 110: 060402.

[11] Cereceda J 2000 Quantum mechanical probabilities and general probabilistic constraints for Einstein–Podolsky–Rosen–Bohm experiments. Found. Phys. Lett. 13: 427–442.

[12] Crespi A, Bentivegna M, Pitsios I, Rusc D, Poderini D, Carvacho G, D'Ambrosio V, Cabello A, Sciarrino F, Osellame R 2017 Single-photon quantum contextuality on a chip. ACS Phot. 4: 2807-2812.

[13] Dzhafarov EN 2017 Replacing nothing with something special: Contextuality-by-Default and dummy measurements. In A. Khrennikov & T. Bourama (Eds.), Quantum Foundations, Probability and Information. Springer.

[14] Dzhafarov EN, Kujala JV 2014 Contextuality is about identity of random variables. Phys. Scr. T163: 014009.

[15] Dzhafarov EN, Kujala JV 2016 Context-content systems of random variables: The Contextuality-by- Default theory. J. Math. Psych. 74: 11-33.

[16] Dzhafarov EN, Kujala JV 2017 Probabilistic foundations of contextuality. Fort. Phys. 65: 1-11.

[17] Dzhafarov EN, Kujala JV 2018 Contextuality analysis of the double slit experiment (with a glimpse into three slits). Entropy 20: 278; doi:10.3390/e20040278.

[18] Dzhafarov EN, Cervantes VH, Kujala JV 2017 Contextuality in canonical systems of random variables. Phil. Trans. Roy.Soc. A 375: 20160389.

[19] Dzhafarov EN, Gluhovsky I 2006 Notes on selective influence, probabilistic causality, and probabilistic dimensionality. J. Math. Psych. 50: 390–401.

[20] Dzhafarov EN, Kujala JV, Larsson, J-Å 2015 Contextuality in three types of quantum-mechanical systems. Found. Phys. 7: 762-782.

[21] Einstein A, Podolsky B, Rosen N 1935 Can quantum-mechanical description of physical reality be considered complete? Phys. Rev., 47: 777-780.

[22] Fine A 1982 Hidden variables, joint probability, and the Bell inequalities. J. Math. Phys., Vol. 23: 1306–1310.

[23] Flühmann C, Negnevitsky V, Marinelli M, Home JP 2018 Sequential modular position and momentum measurements of a trapped ion mechanical oscillator. Phys. Rev. X 8: 021001.

[24] Griffiths RB 2017 What quantum measurements measure. Phys. Rev. A 96: 032110.

[25] Griffiths RB 2019 Quantum measurements and contextuality. Phil. Trans. Roy.Soc. A, present issue.

[26] Gröblacher S, Paterek T, Kaltenbaek R, Brukner Č, Żukowski M, Aspelmeyer M, Zeilinger A 2007 An experimental test of non-local realism. Nature 446: 871–875.

[27] Kechris AS 1995 Classical Descriptive Set Theory. New York: Springer.

[28] Khrennikov A 2005 The principle of supplementarity: A contextual probabilistic viewpoint to complementarity, the interference of probabilities, and the incompatibility of variables in quantum mechanics. Found. Phys., 35: 1655 - 1693.

[29] Khrennikov A 2009 Contextual approach to quantum formalism. Springer: Berlin.

[30] Klyachko AA, Can MA, Binicioğlu S, Shumovsky AS 208 Simple test for hidden variables in spin-1 systems. Phys. Rev. Lett. 101: 020403.

[31] Kochen S, Specker EP 1967 The problem of hidden variables in quantum mechanics. J. Math. Mech. 17: 59–87.

[32] Kujala JV, Dzhafarov EN 2016 Proof of a conjecture on contextuality in cyclic systems with binary variables. Found. Phys. 46: 282-299.

[33] Kujala JV, Dzhafarov EN 2019 Measures of contextuality and noncontextuality. Phil. Trans. Roy.Soc. A, present issue.

[34] Kujala JV, Dzhafarov EN, Larsson J-Å 2015 Necessary and sufficient conditions for extended noncontextuality in a broad class of quantum mechanical systems. Phys. Rev. Lett. 115: 150401.

[35] Kurzynski P, Cabello A, Kaszlikowski D. 2014 Fundamental monogamy relation between contextuality and nonlocality. Phys. Rev. Lett. 112: 100401.

[36] Kurzynski P, Ramanathan R, Kaszlikowski D 2012 Entropic test of quantum contextuality. Phys. Rev. Lett. 109: 020404.

[37] Leggett AJ 2003 Nonlocal hidden-variable theories and quantum mechanics: An incompatibility theorem. Found. Phys. 33: 1469

[38] Leggett AJ, Garg A 1985 Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? Phys. Rev. Lett. 54: 857–860.

[39] Liang Y-C. Spekkens RW, Wiseman HM 2011 Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity. Phys. Rep. 506: 1-39.

[40] Malinowski M, Zhang C, Leupold FM, Alonso J, Home JP, Cabello A 2018 Probing the limits of correlations in an indivisible quantum system. Phys. Rev. A 98: 050102.

[41] Suppes P., Zanotti M 1981 When are probabilistic explanations possible? Synthese 48: 191–199.

[42] Thorisson H 2000 Coupling, Stationarity, and Regeneration. Springer, New York.

[43] William S 2019 Counterfactuals. The Stanford Encyclopedia of Philosophy (Spring 2019 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/spr2019/entries/counterfactuals/.

[44] Zhan X, Kurzyński P, Kaszlikowski D, Wang K, Bian Zh, Zhang Y, Xue P 2017 Experimental detection of information deficit in a photonic contextuality scenario. Phys. Rev. Lett. 119: 220403.