Contextuality in Three Types of Quantum-Mechanical Systems

Ehtibar N. Dzhafarov[1], Janne V. Kujala[2], and Jan-Åke Larsson[3]

[1]Purdue University, ehtibar@purdue.edu
[2]University of Jyväskylä, jvk@iki.fi
[3]Linköping University, jan-ake.larsson@liu.se

Abstract

We present a formal theory of contextuality for a set of random variables grouped into different subsets (contexts) corresponding to different, mutually incompatible conditions. Within each context the random variables are jointly distributed, but across different contexts they are stochastically unrelated. The theory of contextuality is based on the analysis of the extent to which some of these random variables can be viewed as preserving their identity across different contexts when one considers all possible joint distributions imposed on the entire set of the random variables. We illustrate the theory on three systems of traditional interest in quantum physics (and also in non-physical, e.g., behavioral studies). These are systems of the Klyachko-Can-Binicioglu-Shumovsky-type, Einstein-Podolsky-Rosen-Bell-type, and Suppes-Zanotti-Leggett-Garg-type. Listed in this order, each of them is formally a special case of the previous one. For each of them we derive necessary and sufficient conditions for contextuality while allowing for experimental errors and contextual biases or signaling. Based on the same principles that underly these derivations we also propose a measure for the degree of contextuality and compute it for the three systems in question.

KEYWORDS: CHSH inequalities; contextuality; Klyachko inequalities; Leggett-Garg inequalities; probabilistic couplings; signaling.

1 Introduction

A deductive mathematical theory is bound to begin with definitions and/or axioms, and one is free not to accept them. We propose a certain definition of contextuality which may or may not be judged "good." Ultimately, its utility will be determined by whether it leads to fruitful mathematical developments and interesting applications. Our definition applies to situations where contextuality is traditionally investigated in quantum physics: Klyachko-Can-Binicioglu-Shumovsky-type systems of measurements [1], Einstein-Podolsky-Rosen-Bell-type systems [2–6], and Suppes-Zanotti-Leggett-Garg-type systems [7,8]. We will refer to these systems by abbreviations KCBS, EPRB, and SZLG, respectively. In the absence of what we call "inconsistency," our contextuality criteria (necessary and sufficient conditions) coincide with the traditional inequalities. But our criteria also apply to situations with measurement errors, contextual biases, and interaction among jointly measured physical properties ("signaling"). Moreover, the logic of constructing our criteria of contextuality leads to a natural quantification of the degree of contextuality in the three types of systems considered.

This paper can be viewed as a companion one for Ref. [9], in which we prove a general criterion for contextuality in "cyclic" systems of which the systems just mentioned (KCBS, EPRB, and SZLG) are special cases. However, we use here a different criterion for contextuality in these three types of systems, whose advantage is in that it is directly related to the notion of the degree of contextuality. At the end of this paper we conjecture (see Remark 40) a generalization of the criterion and the measure of the degree of contextuality to all "cyclic" systems.

The notion of probabilistic contextuality is usually understood to be about "sewing together" random variables recorded under different conditions. That is, it is viewed as answering the question: given certain sets of jointly distributed random variables, can a joint distribution be found for their union? The key aspect and difficulty in answering this questions is that different sets of random variables generally pairwise overlap, share some of their elements. In Ernst Specker's [10] well-known example with three magic boxes containing (or not containing) gems, which we present here in probabilistic terms, we have three binary random variables, A, B, C, that can only be recorded in pairs,

$$X = (A, B)\,,\; Y = (B, C)\,,\; Z = (A, C)\,. \tag{1}$$

That is, the joint distribution of A and B in X is known, and the same is true for the components of Y and Z. We ask whether there is a joint distribution of all three of them, (A, B, C), that agrees with the distributions of X, Y, and Z as its 2-component marginals. In Specker's example the boxes are magically rigged so that (assuming A, B, C attain values $+1$/-1, denoting the presence/absence of a gem in the respective box)

$$\Pr[A = -B] = 1,\; \Pr[-B = C] = 1,\; \Pr[C = -A] = 1, \tag{2}$$

which, obviously, precludes the existence of a jointly distributed (A, B, C). We may say then that the system of random variables (1) exhibits contextuality.

On a deeper level of analysis, however, contextuality is better to be presented as a problem of determining *identities of the random variables recorded under different conditions*. That is, it answers the question: is this random variable (under this condition), say, A in X, "the same as" that one (under another condition), say, A in Y, or is the former at least "as close" to the latter as their distributions in the two pairs allow?

This deeper view is based on the principle we dubbed *Contextuality-by-Default*, developed through a series of recent publications [11–18]. According to this principle, any two random variables recorded under different (i.e., mutually exclusive) conditions (treatments) are labeled by these conditions and considered *stochastically unrelated* (defined on different sample spaces, possessing no joint distribution). Thus, in Specker's example with the magic boxes, we need to denote the observed three pairs of random variables not as in (1), but as

$$X = (A_X, B_X)\,,\; Y = (B_Y, C_Y)\,, Z = (A_Z, C_Z)\,. \tag{3}$$

Of course, any other unique labeling making random variables in one context distinct from random variables in another context would do as well. The notion of stochastic unrelatedness within the framework of the Kolmogorovian probability theory has been explored in the quantum-theoretic literature, notably by A. Khrennikov [22–24].

The use of this notion within the present conceptual framework is based on the fact that stochastically unrelated random variables can always be *coupled* (imposed a joint distribution upon) [13, 14, 19–21]. This can generally be done in multiple ways, and no couplings are privileged a priori. For Specker's example, one constructs a random 6-tuple

$$S = (A_X, B_X, B_Y, C_Y, A_Z, C_Z) \tag{4}$$

such that its 2-marginals X, Y, Z in (3) are consistent with the observed probabilities. In particular, they should satisfy

$$\Pr[A_X = -B_X] = 1,\ \Pr[-B_Y = C_Y] = 1,\ \Pr[C_Z = -A_Z] = 1. \tag{5}$$

Such a coupling S can be constructed in an infinity of ways. To match this representation with Specker's original meaning, we have to impose additional constraints on the possible couplings. Namely, we have to require that S in (4) be constructed subject to the following "identity hypothesis":

$$\Pr[A_X = A_Z] = \Pr[B_X = B_Y] = \Pr[C_Y = C_Z] = 1. \tag{6}$$

Such a coupling S, as we have already determined, does not exist, and we can say that the system of the random variables (3) exhibits contextuality with respect to the identity hypothesis (6).

One might wonder whether this re-representation of the problem is useful. Aren't the questions

> "Let me see if I can 'sew together' (A, B), (B, C), and (A, C) into a single (A, B, C),"

and

> "Let me see if I can put together (A_X, B_X), (B_Y, C_Y), and (A_Z, C_Z) into a single S in (4) under the identity hypothesis (6),"

aren't they one and the same question in two equivalent forms? Clearly, they are. But there are two (closely related) advantages of the second formulation:

1. It can be readily generalized by replacing the perfect identities in (6) with less stringent or altogether different constraints; and

2. for any given constraint, if a coupling satisfying it does not exist, this approach allows one to gauge how close one can get to satisfying it, i.e., one has a principled way for constructing a measure for the degree of contextuality the system exhibits.

To illustrate these interrelated points on Specker's example, observe that the identity hypothesis (6) cannot be satisfied if the system is "inconsistently connected," i.e., if the marginal distribution of, say, A_X is not the same as that of A_Z. This may happen if the magic boxes somehow physically communicate (e.g., the gem can be transposed from one of the boxes being opened to another), and the probability of finding a gem in the first box ($A = 1$) is affected differently by the opening of the second box (i.e., in context X) and of the third box (in context Z). A_X and A_Z may have different distributions also as a result of (perhaps magically induced) errors in correctly identifying which of the two open boxes contains a gem: e.g., when boxes i and j are open ($i < j$), one may with some probability erroneously see/record the gem contained in the ith box as being in the jth box. We may speak of "signaling" between the boxes in the former case, and of "contextual measurement biases" in the latter. In either case, the requirement (6) cannot be satisfied for A_X

and A_Z, and to determine this one does not even have to look at the observed distributions of (A_X, B_X), (B_Y, C_Y), and (A_Z, C_Z). However, in either of the two cases one can meaningfully ask: what is the maximum possible value of $\Pr[A_X = A_Z]$ that is consistent with the distributions of A_X and A_Z (and analogously for B_X, B_Y and C_Y, C_Z), and are these maximum possible values consistent with the observed distributions of (A_X, B_X), (B_Y, C_Y), and (A_Z, C_Z)?

We will proceed now to formulate these ideas in a more rigorous way.

2 Systems, Random Bunches, and Connections

Let $X = (A, B, C, \ldots)$ be a (generalized) sequence[1] of jointly distributed random variables, called *components* of X. We will refer to X as a *(random) bunch*. Let $\mathfrak{S}$ be a set of random bunches

$$X = (A_X, A'_X, A''_X, \ldots), Y = (B_Y, B'_Y, B''_Y, \ldots), Z = (C_Z, C'_Z, C''_Z, \ldots), \ldots \tag{7}$$

(of arbitrary cardinalities), with the property that they are *pairwise componentwise stochastically unrelated*. The term means that no component of one random bunch is jointly distributed with any component of another.

Remark 1. Intuitively, each random bunch corresponds to certain *conditions* under which (or *contexts* in which) the components of the bunch are jointly recorded; and the conditions corresponding to different random bunches are *mutually exclusive*.

Any pair $\{A, B\}$ such that A and B are components of two *distinct* random bunches in $\mathfrak{S}$ is called a (simple) *connection*. A set $\mathfrak{C}$ of pairwise disjoint connections is called a *simple set of (simple) connections*.

Remark 2. Intuitively, a connection indicates a pair of random variables A and B that represent "the same" physical property, because of which, ideally, they should be "one and the same" random variable in different contexts. However, the distributions of A and B may be different due to signaling (from other random variables in their contexts) or due to contextual measurement biases. Note that the elements A and B of a connection never co-occur, i.e., they possess no joint distribution, and their "identity" therefore can never be verified by observation.

Together, $(\mathfrak{S}, \mathfrak{C})$ form a *system* (of measurements, or of random bunches). Without loss of generality, we can assume that $\mathfrak{S}$ contains no "non-participating" bunches, i.e., each random bunch X has at least one component A that belongs to some connection $\{A, B\}$ in $\mathfrak{C}$. In particular, if set $\mathfrak{C}$ is finite, then so is set $\mathfrak{S}$ (even if the number of components in some of the bunches in $\mathfrak{S}$ is not finite).

Example 3 (KCBS-system)**.** A KCBS-system [1] consists of five pairs of binary (± 1) random variables,

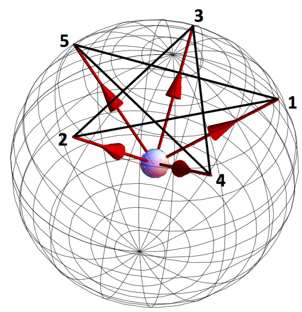
$$\mathfrak{S} = \{(V_1, W_2), (V_2, W_3), (V_3, W_4), (V_4, W_5), (V_5, W_1)\}. \tag{8}$$

Abstracting away from the physical meaning, the schematic picture below shows five radius-vectors, each corresponding to a distinct physical property represented by a binary random variable. They can only be recorded in pairs (8), and each of these pairs corresponds to vertices connected by an

[1] A sequence is an indexed set, and "generalized" means that the indexing is not necessarily finite or countable. We try to keep the notation simple, omitting technicalities. A sequence of random variables that are jointly distributed is a random variable, if the latter term is understood broadly, as anything with a well-defined probability distribution, to include random vectors, random sets, random processes, etc.

edge of the pentagram. In accordance with the Contextuality-by-Default principle, we label each variable both by index $i \in \{1, \ldots, 5\}$ indicating the radius-vector (physical property) it corresponds to, and by the context, defined by which of the two pairs it enters. We use notation V_i in one of these pairs and W_i in another. For instance, $i = 2$ is used to label V_2 in the pair (V_2, W_3) and W_2 in the pair (V_1, W_2). With this notation, the simple set of the connections of interest in this system is

$$\mathfrak{C} = \{(V_1, W_1), (V_2, W_2), (V_3, W_3), (V_4, W_4), (V_5, W_5)\}. \tag{9}$$



In the ideal KCBS-system, each recorded pair, say, (V_1, W_2), can attain values $(+1, -1)$, $(-1, +1)$, $(-1, -1)$, but not $(+1, +1)$. In our analysis, however, we allow for experimental errors, so that the "pure" KCBS-system is a special case of a more general system in which $\Pr[V_1 = +1, W_2 = +1]$ may be non-zero. In the ideal KCBS-system the probabilities are computed in accordance with the principles of quantum mechanics, so that the distribution of (V_i, W_j) in (8) is determined by the angle between the radius vectors i and j, and the distributions of V_i is always the same as the distribution of W_i ($i = 1, \ldots, 5$). In our analysis, however, we allow for "signaling" between the detectors and/or for "contextual measurement biases," so that, e.g., V_1 in (V_1, W_2) and W_1 in (V_5, W_1) may have different distributions.

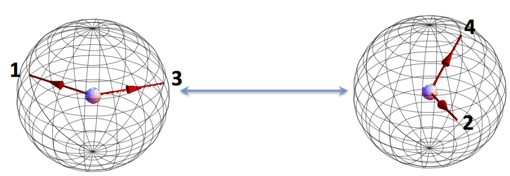
Remark 4. There is no "traditional" contextual notation for the KCBS system, but one can think of a variety of alternatives to our $V - W$ scheme, e.g., denoting the ith measurement in the context of being conjoint with the jth measurement by R_i^j, as we do in Ref. [9].

Example 5 (EPRB-systems)**.** An EPRB-system [2–6] consists of four pairs of binary (± 1) random variables,

$$\mathfrak{S} = \{(V_1, W_2), (V_2, W_3), (V_3, W_4), (V_4, W_1)\}. \tag{10}$$

Again we abstract away from the physical meaning, involving spins of entangled particles. In the schematic picture below each direction (1 or 3 in one particle and 2 or 4 in another) corresponds to a binary random variable. They are recorded in pairs $\{1, 3\} \times \{2, 4\}$, so each random variable participates in two contexts, and is denoted either V_i or W_i ($i \in \{1, 2, 3, 4\}$) accordingly. The simple set of connections of interest is

$$\mathfrak{C} = \{(V_1, W_1), (V_2, W_2), (V_3, W_3), (V_4, W_4)\}. \tag{11}$$



Here, (V_i, W_j) may attain all four possible values $(\pm 1, \pm 1)$. In the ideal system with space-like separation between the recordings of V_i and W_j, the distribution of V_i is always the same as that of W_i $(i = 1, \ldots, 4)$. However, we allow for the possibility that the measurements are time-like separated (so that direct signaling is possible), as well as for the possibility that the results of the two measurements are recorded by someone who may occasionally make errors and be contextually biased. Thus, one may erroneously assign $+1$ to, say, $V_1 = -1$ more often than to $W_1 = -1$.

Remark 6. The contextual notation for the EPRB-systems adopted in our previous papers [11–18] is (A_{ij}, B_{ij}), $i, j \in \{1, 2\}$, where A and B refer to measurements on the first and second particles, respectively. The first index refers to one of the two A-measurements (1 or 2), the second index refers to one of the two B-measurements (1 or 2). So the non-contextual (misleading) notation for (A_{ij}, B_{ij}) would be (A_i, B_j). In relation to our present notation, A_{11} corresponds to V_1, and A_{12} (the same property in another context) to W_1; A_{21} corresponds to W_3, and A_{22} (the same property in another context) to V_3; and analogously for B_{ij}.

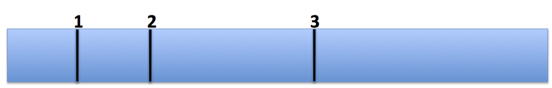
Example 7 (SZLG-system)**.** An SZLG-system [7, 8] consists of three pairs of binary (± 1) random variables,

$$\mathfrak{S} = \{(V_1, W_2), (V_2, W_3), (V_3, W_1)\}. \tag{12}$$

The three random variables are recorded in pairs, the logic of the notation being otherwise the same as above. The simple set of connections of interest in this system is

$$\mathfrak{C} = \{(V_1, W_1), (V_2, W_2), (V_3, W_3)\}. \tag{13}$$

In the Leggett-Garg paradigm proper [7], the three measurements are made at three moments of time, fixed with respect to some zero point, as shown in the schematic picture below. This is, however, only one possible physical meaning, and we can think of any three identifiable measurements performed two at a time.



Again, we allow for the possibility of signaling (which is predicted by the laws of quantum mechanics in some cases, e.g., for pure initial states, as shown in Ref. [26]), i.e., earlier measurements may influence later ones. And, again, we allow for measurement errors and contextual biases: knowing, e.g., that W_2 is preceded by V_1 and is not followed by another measurement, and that V_2 is followed by W_2 and is not preceded by another measurement, may lead one to record V_2 and W_2 differently even if they are identically distributed "in reality."

Remark 8. The contextual notation for the LG-systems adopted in Ref. [16–18] is (Q_{ij}, Q_{ji}), $i, j \in \{1, 2, 3\}$ $(i < j)$, where the first index refers to the earlier of the two measurements. Thus, Q_{12} corresponds to V_1 in our present notation, and Q_{13} (the same property in another context) to W_1; Q_{23} corresponds to V_2, and A_{21} (the same property in another context) to W_2; and analogously for Q_{31} and Q_{32} (resp., V_3 and W_3). In Ref. [26] the notation used is $Q_i^{\{i,j\}}$, where the superscript indicates the context and the subscript the physical property.

3 Contextuality

This section contains our main definitions: of a maximal connection, of (in)consistent connectedness, and of contextuality.

3.1 Couplings

Definition 9. A *coupling* of a set of random variables $X, Y, Z, \ldots$ is a random bunch $(X^*, Y^*, Z^*, \ldots)$ (with jointly distributed components), such that

$$X^* \sim X, Y^* \sim Y, Z^* \sim Z, \ldots, \tag{14}$$

where $\sim$ stands for "has the same distribution as." In particular, a coupling S for $\mathfrak{S}$ is a random bunch coupling all elements (random bunches) of $\mathfrak{S}$.

Example 10. For two binary (± 1) random variables A, B, any random bunch (A^*, B^*) with the distribution

$$\begin{array}{l} r_{11} = \Pr\left[A^* = +1, B^* = +1\right] \\ r_{10} = \Pr\left[A^* = +1, B^* = -1\right] \\ r_{01} = \Pr\left[A^* = -1, B^* = +1\right] \\ r_{00} = \Pr\left[A^* = -1, B^* = -1\right], \end{array} \tag{15}$$

such that

$$\begin{array}{l} r_{11} + r_{10} = \Pr\left[A = +1\right], \\ r_{11} + r_{01} = \Pr\left[B = +1\right], \end{array} \tag{16}$$

is a coupling.

Remark 11. It is a simple but fundamental theorem of Kolmogorov's probability theory [6,8,21,25] that a coupling $(X^*, Y^*, Z^*, \ldots)$ of $X, Y, Z, \ldots$ exists if and only if there is a random variable R and a sequence of measurable functions $(f_X, f_Y, f_Z, \ldots)$, such that

$$X \sim f_X(R), Y \sim f_Y(R), Z \sim f_Z(R), \ldots. \tag{17}$$

In quantum mechanics, R is referred to as a *hidden variable*. (In John Bell's pioneering work [2], he considers the question of whether such a representation exists for four binary random variables A_1, A_2, B_1, B_2 with known distributions of (A_i, B_j), $i, j \in \{1, 2\}$. He imposes no constraints on R, but it is easy to see that the existence of *some* R in his problem is equivalent to the existence of an R with just 16-values.)

3.2 Maximally Coupled Connections and Consistent Connectedness

Definition 12. A coupling (A^*, B^*) of a connection $\{A, B\} \in \mathfrak{C}$ is called *maximal* if

$$\Pr\left[A^* = B^*\right] \geq \Pr\left[A^{**} = B^{**}\right] \tag{18}$$

for any coupling (A^{**}, B^{**}) of $\{A, B\}$.

Definition 13. A system $(\mathfrak{S}, \mathfrak{C})$ is *consistently connected* (CC) if $A \sim B$ in any connection $\{A, B\} \in \mathfrak{C}$. Otherwise the system $(\mathfrak{S}, \mathfrak{C})$ is *inconsistently connected* (not CC).

Remark 14. In physics, the CC condition is sometimes referred to as "no-signaling" [27–29], the term we are going to avoid because then non-CC systems should be referred to as "signaling." The latter term has strong connotations making its use in our technical meaning objectionable to physicists. Inconsistent connectedness may be due to signaling in the narrow physical meaning, but it may also indicate measurement biases due to context (so that one measures A differently when one also measures B than when one also measures C). We make no distinction between "ideal" random variables and those measured "incorrectly."

Lemma 15. *In a CC system, a maximal coupling* (A^*, B^*) *for any connection* $\{A, B\} \in \mathfrak{C}$ *exists, and in this coupling* $\Pr[A^* = B^*] = 1$. *If the system is not CC, and* $A \not\sim B$ *in a connection* $\{A, B\}$, *then in a maximal coupling* (A^*, B^*), *if it exists,* $\Pr[A^* = B^*] < 1$.

A proof is obvious. We focus now on *binary systems*, in which all components of the random bunches are binary (± 1) variables. The three systems mentioned in the opening section, KCBS, LG, and EPRB-type ones, are binary. (A generalization to components with finite but arbitrary numbers of values is straightforward.)

Lemma 16. *For a connection* $\{A, B\}$ *with binary* (± 1) A, B *and*

$$p = \Pr[A = 1] \geq \Pr[B = 1] = q, \tag{19}$$

a maximal coupling (A^*, B^*) *exists, and its distribution is*

$$\begin{array}{l} r_{11} = q \\ r_{10} = p - q \\ r_{01} = 0 \\ r_{00} = 1 - p, \end{array} \tag{20}$$

where

$$r_{ab} = \Pr[A^* = 2a - 1, B^* = 2b - 1]. \tag{21}$$

Proof. For given values $p \geq q$, the maximum possible value of r_{11} is $\min(p, q) = q$, and the maximum possible value of r_{00} is $\min(1 - p, 1 - q) = 1 - p$; these values are attained in distribution (20), with r_{01}, r_{10} determined uniquely. $\Pr[A^* = B^*] = r_{11} + r_{00}$ in this distribution has the maximum possible value, $1 - (p - q)$. □

Remark 17. In a maximal coupling (A^*, B^*) of two binary random variables A, B the expectation

$$\langle A^* B^* \rangle = 2(r_{11} + r_{00}) - 1 \tag{22}$$

attains its maximum possible value (assuming $p \geq q$)

$$\langle A^* B^* \rangle = 1 - 2(p - q) = 1 - (\langle A \rangle - \langle B \rangle). \tag{23}$$

.

Remark 18. In some cases it is more convenient to speak of the minimum value of $\Pr[A^* \neq B^*] = r_{10} + r_{01}$ rather than the maximum value of $\Pr[A^* = B^*] = r_{11} + r_{00}$. This minimum can be presented as

$$\Pr[A^* \neq B^*] = p - q = \frac{1}{2}(\langle A \rangle - \langle B \rangle). \tag{24}$$

Remark 19. In a maximal coupling (A^*, B^*) of two binary random variables A, B,

$$\Pr[A^* = B^*] = 1 \tag{25}$$

if and only if $A \sim B$, i.e.,

$$p = \Pr[A = 1] = \Pr[B = 1] = q. \tag{26}$$

3.3 Contextuality

Definition 20. Let maximal couplings exist for all connections in $\mathfrak{C}$. A system $(\mathfrak{S}, \mathfrak{C})$ has a *(maximally) noncontextual description* if there exists a coupling S for $\mathfrak{S}$ in which all 2-marginals (A^*, B^*) that couple the connections in $\mathfrak{C}$ are maximal couplings. If such a coupling S does not exist, the system is *contextual*.

Remark 21. We will omit the qualifier "maximally" when speaking of the existence of a maximally noncontextual description.

Remark 22. In particular, if the system is CC, it has a noncontextual description if and only if there is a coupling S for $\mathfrak{S}$ in which $\Pr[A^* = B^*] = 1$ for all connections $\{A, B\} \in \mathfrak{C}$. This is essentially the traditional use of the term "(non)contextuality."

Example 23. Let system $(\mathfrak{S}, \mathfrak{C})$ consist of random bunches

$$(A, C), (B, D),$$

with all components binary (± 1), and a single connection $\{A, B\}$. To determine if the system is contextual, we consider all possible couplings for A, B, C, D, i.e., all possible random bunches

$$(A^*, B^*, C^*, D^*)$$

such that $(A^*, C^*) \sim (A, C)$ and $(B^*, D^*) \sim (B, D)$. Denoting, for $a, b, c, d \in \{0, 1\}$,

$$\begin{array}{r} s_{ac} = \Pr[A = 2a - 1, C = 2c - 1], \\ t_{bd} = \Pr[B = 2b - 1, D = 2d - 1], \\ u_{abcd} = \Pr[A^* = 2a - 1, B^* = 2b - 1, C^* = 2c - 1, D^* = 2d - 1], \end{array} \tag{27}$$

the distributional equations $(A^*, C^*) \sim (A, C)$ and $(B^*, D^*) \sim (B, D)$ translate into the following 8 equations for 16 probabilities u_{abcd}:

$$\sum_{b=0}^{1} \sum_{d=0}^{1} u_{abcd} = s_{ac}, \quad a, c \in \{0, 1\}, \tag{28}$$

$$\sum_{a=0}^{1} \sum_{c=0}^{1} u_{abcd} = t_{bd}, \quad b, d \in \{0, 1\}. \tag{29}$$

The requirement that the coupling for $\{A, B\}$ be maximal translates into the additional four equations

$$\sum_{c=0}^{1} \sum_{d=0}^{1} u_{abcd} = r_{ab} = \Pr[A^* = 2a - 1, B^* = 2b - 1], \tag{30}$$

where, in view of Lemma (16), r_{ab} is given by (20), with the same meaning of p, q and the same convention $p = \Pr[A = 1] \geq \Pr[B = 1] = q$. The problem of contextuality therefore reduces to one of determining whether the system of 12 equations (28)-(29)-(30) for the 16 unknown $u_{abcd} \geq 0$ has a solution. The answer in this case can be shown to be affirmative, so the system considered has a noncontextual description.

Definition 20 is sufficient for all subsequent considerations in this paper, but we note that it can be extended to situations when maximal couplings for connections do not necessarily exist. Let us associate with each connection for A, B a supremal number

$$p_{AB} = \sup_{\text{all couplings } (A^*,B^*)} \Pr[A^* = B^*].$$

Definition 24 (extended)**.** A system $(\mathfrak{S}, \mathfrak{C})$ has a *(maximally) noncontextual* description (*is contextual*) if there exists (resp., does not exist) a sequence of couplings $S_1, S_2, \ldots$ for $\mathfrak{S}$ in which $\Pr[A^* = B^*]$ for all connections $\{A, B\}$ in $\mathfrak{C}$ uniformly converge to the corresponding supremal numbers p_{AB}.

3.4 Measure of Contextuality for Binary Systems with Finite Simple Sets of Connections

We will assume that in the binary systems we are dealing with the simple set of connections $\mathfrak{C}$ is finite:

$$\mathfrak{C} = \{\{A_i, B_i\} : i \in \{1, \ldots, n\}\}. \tag{31}$$

Lemma 25. *Given a finite simple set of connections $\{\{A_i, B_i\} : i \in \{1, \ldots, n\}\}$ in a binary system, the respective couplings in the set $\{(A_i^*, B_i^*) : i \in \{1, \ldots, n\}\}$ are all maximal if and only if*

$$\sum_{i=1}^{n} \Pr[A_i^* \neq B_i{}^*] = \frac{1}{2} \sum_{i=1}^{n} |\langle A_i^* \rangle - \langle B_i^* \rangle|. \tag{32}$$

Proof. Immediately follows from Lemma (16) and Remark (18). □

Notation. We denote

$$\Delta_0(\mathfrak{C}) = \frac{1}{2} \sum_{i=1}^{n} |\langle A_i^* \rangle - \langle B_i^* \rangle|, \tag{33}$$

and this quantity is to play a central role in the subsequent computations.

Definition 26. Let $\Delta_{\min}(\mathfrak{S}, \mathfrak{C})$ for a system with $\mathfrak{C} = \{\{A_i, B_i\} : i \in \{1, \ldots, n\}\}$ be the infimum for

$$\sum_{i=1}^{n} \Pr[A_i^* \neq B_i{}^*]$$

across all possible couplings S for $\mathfrak{S}$.

This is another quantity to play a central role in subsequent computations.

Theorem 27. *For a binary system with a finite simple set of connections, the value* $\Delta_{\min}(\mathfrak{S},\mathfrak{C})$ *is achieved in some coupling* S*, and*

$$\Delta_{\min}(\mathfrak{S},\mathfrak{C}) \geq \Delta_0(\mathfrak{C}). \tag{34}$$

The system has a noncontextual description if and only if

$$\Delta_{\min}(\mathfrak{S},\mathfrak{C}) = \Delta_0(\mathfrak{C}). \tag{35}$$

Proof. That $\Delta_{\min}(\mathfrak{S},\mathfrak{C})$ is an achievable minimum follows from the fact that any coupling S is described by a system of linear inequalities relating to each other

$$\Pr\left[A^* = a, B^* = b, C^* = c, \ldots\right]$$

for all possible values $(a, b, c, \ldots)$ of all the random variables involved (the union of all random bunches in $\mathfrak{S}$). $\Delta_{\min}(\mathfrak{S},\mathfrak{C})$ being a linear combinations of these probabilities, its infimum has to be a minimum. (34) and (35) are obvious. □

This theorem allows one to construct a convenient definition for the degree of contextuality in binary systems with finite number of connections.

Definition 28. In a binary system $(\mathfrak{S},\mathfrak{C})$ with a finite simple set of connections the *degree of contextuality* is

$$\mathsf{CNTX}(\mathfrak{S},\mathfrak{C}) = \Delta_{\min}(\mathfrak{S},\mathfrak{C}) - \Delta_0(\mathfrak{C}) \geq 0. \tag{36}$$

In the subsequent sections of this paper we show how this definition of contextuality applies to KCBS, LG, and EPRB-systems.

Remark 29. Clearly, a measure of contextuality could also be constructed as $\left(1+\Delta_{\min}(\mathfrak{S},\mathfrak{C})\right)/\left(1+\Delta_0(\mathfrak{C})\right)-1$, $\left(\Delta_{\min}(\mathfrak{S},\mathfrak{C})-\Delta_0(\mathfrak{C})\right)/\left(\Delta_{\min}(\mathfrak{S},\mathfrak{C})+\Delta_0(\mathfrak{C})\right)$, and in a variety of other ways. The only logically necessary aspect of the definition is that $\mathsf{CNTX}(\mathfrak{S},\mathfrak{C})$ is zero when $\Delta_{\min}(\mathfrak{S},\mathfrak{C}) = \Delta_0(\mathfrak{C})$ and positive otherwise. The simple difference is chosen because it has been shown to have certain desirable properties [16], but this choice is not critical for the present paper.

3.5 Conventions

3.5.1 Abuse of language

To simplify notation we adopt the following convention: in a coupling (X^*, Y^*) of two random variables X, Y we drop the asterisks and write simply (X, Y). The abuse of language thus introduced is common, if not universally accepted in quantum physics, and we too conveniently resorted to it when discussing Specker's magic boxes in our introductory section.

3.5.2 Functions s_{even} and s_{odd}

We will make use of the following notation. For any finite sequence of real numbers $(a_i : i \in \{1, \ldots, n\})$ we denote by

$$\mathsf{s}_{even}(a_i : i \in \{1,\ldots,n\}) = \mathsf{s}_{even}(a_1,\ldots,a_n) = \max_{\text{even number of } -\text{'s}} \sum_{i=1}^{n} (\pm a_i), \tag{37}$$

where each $\pm$ should be replaced with $+$ or $-$, and the maximum is taken over all combinations of the signs containing an even number of $-$'s.

Analogously,

$$\mathsf{s}_{odd}\left(a_i : i \in \{1,\ldots,n\}\right) = \mathsf{s}_{odd}\left(a_1,\ldots,a_n\right) = \max_{\text{odd number of } -\text{'s}} \sum_{i=1}^{n} \left(\pm a_i\right). \tag{38}$$

Lemma 30. *For any finite sequence of real numbers* $(a_i : i \in \{1,\ldots,n\})$,

$$\mathsf{s}_{even}\left(a_1,\ldots,a_n\right) = \sum |a_i| - 2[a_1 \ldots a_k < 0] \min\left(|a_1|,\ldots,|a_n|\right), \tag{39}$$

$$\mathsf{s}_{odd}\left(a_1,\ldots,a_n\right) = \sum |a_i| - 2[a_1 \ldots a_k > 0] \min\left(|a_1|,\ldots,|a_n|\right). \tag{40}$$

4 KCBS-systems

4.1 Main Theorem

Theorem 31 (contextuality measure and criterion for KCBS-systems)**.** *In a KCBS-system* $(\mathfrak{S},\mathfrak{C})$*, with*

$$\mathfrak{S} = \left\{(V_i, W_{i\oplus_5 1}) : i \in \{1,\ldots,5\}\right\}, \mathfrak{C} = \left\{(V_i, W_i) : i \in \{1,\ldots,5\}\right\}, \tag{41}$$

where $\oplus_5$ *stands for circular addition of* 1 *on* $\{1,2,3,4,5\}$*,*

$$\Delta_0\left(\mathfrak{C}\right) = \frac{1}{2}\sum_{i=1}^{5} \left|\langle V_i\rangle - \langle W_i\rangle\right|, \tag{42}$$

$$\Delta_{\min}\left(\mathfrak{S},\mathfrak{C}\right) = \frac{1}{2}\max\left(2\Delta_0\left(\mathfrak{C}\right),\ \mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_5 1}\rangle : i \in \{1,\ldots,5\}\right) - 3\right). \tag{43}$$

Consequently, the degree of contextuality in the KCBS-system is

$$\mathsf{CNTX}\left(\mathfrak{S},\mathfrak{C}\right) = \frac{1}{2}\max\left(0,\ \mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_5 1}\rangle : i \in \{1,\ldots,5\}\right) - 3 - \sum_{i=1}^{5}\left|\langle V_i\rangle - \langle W_i\rangle\right|\right), \tag{44}$$

and the system has a noncontextual description if and only if

$$\mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_5 1}\rangle : i \in \{1,\ldots,5\}\right) \leq 3 + \sum_{i=1}^{5}\left|\langle V_i\rangle - \langle W_i\rangle\right|. \tag{45}$$

Proof. The computer-assisted proof is based on Lemma 32 below, and its details, omitted here, are analogous to those in the proofs of Theorems 3-6 in Appendix of Ref. [18]. □

Lemma 32. *The necessary and sufficient condition for the connection couplings* $\{(V_i, W_i) : i \in \{1,\ldots,5\}\}$ *to be compatible with the observed pairs* $\{(V_i, W_{i\oplus_5 1}) : i \in \{1,\ldots,5\}\}$ *is*

$$\mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_5 1}\rangle, \langle V_i W_i\rangle : i \in \{1,\ldots,5\}\right) \leq 8, \tag{46}$$

which can be equivalently written as

$$\mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_4 1}\rangle : i \in \{1,\ldots,5\}\right) + \mathsf{s}_{even}\left(\langle V_i W_i\rangle : i \in \{1,\ldots,5\}\right) \leq 8,$$

$$\mathsf{s}_{even}\left(\langle V_i W_{i\oplus_4 1}\rangle : i \in \{1,\ldots,5\}\right) + \mathsf{s}_{odd}\left(\langle V_i W_i\rangle : i \in \{1,\ldots,5\}\right) \leq 8.$$

Remark 33. The compatibility in the formulation of the lemma means the existence of a coupling S for $\mathfrak{S}$ with given marginals $\{(V_i, W_i) : i \in \{1, \ldots, 5\}\}$ and given (coupled) connections $\{(V_i, W_{i\oplus_5 1}) : i \in \{1, \ldots, 5\}\}$.

4.2 Special cases

In a CC KCBS-system, $\Delta_0(\mathfrak{C}) = 0$, the criterion for noncontextuality acquires the form

$$\mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_5 1}\rangle : i \in \{1, \ldots, 5\}\right) \leq 3. \tag{47}$$

If, in addition, the KCBS-exclusion is satisfied,[2] i.e., in every $\langle V_i W_{i\oplus_5 1}\rangle$,

$$\Pr\left[V_i = 1, W_{i\oplus_5 1} = 1\right] = 0, \tag{48}$$

then we have

$$\langle V_i W_{i\oplus_5 1}\rangle = 1 - 2\left(p_i + p_{i\oplus_5 1}\right), \tag{49}$$

where

$$p_i = \Pr\left[V_i = 1\right] = \Pr\left[W_i = 1\right], \quad i \in \{1, \ldots, 5\}. \tag{50}$$

It follows that

$$\sum_{i=1}^{5} p_i \leq 2. \tag{51}$$

This is the KCBS inequality, that has been derived in Ref. [1] as a necessary condition for noncontextuality. As it turns out (we omit the simple proof), this condition is also necessary.

Theorem 34. *In a CC KCBS-system with KCBS exclusion,* (47) *is equivalent to* (51).

5 EPRB-systems

5.1 Main Theorem

Theorem 35 (contextuality measure and criterion for EPRB-systems)**.** *In an EPRB-system* $(\mathfrak{S}, \mathfrak{C})$*, with*

$$\mathfrak{S} = \left\{(V_i, W_{i\oplus_4 1}) : i \in \{1, \ldots, 4\}\right\}, \mathfrak{C} = \left\{(V_i, W_i) : i \in \{1, \ldots, 4\}\right\}, \tag{52}$$

where $\oplus_4$ *stands for circular addition of* 1 *on* $\{1, 2, 3, 4\}$*,*

$$\Delta_0(\mathfrak{C}) = \frac{1}{2}\sum_{i=1}^{4} \left|\langle V_i\rangle - \langle W_i\rangle\right|, \tag{53}$$

and

$$\Delta_{\min}(\mathfrak{S}, \mathfrak{C}) = \frac{1}{2}\max\left(2\Delta_0(\mathfrak{C}),\ \mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_4 1}\rangle : i \in \{1, \ldots, 4\}\right) - 2\right). \tag{54}$$

[2] This means that the directions in the 3D real Hilbert space are chosen strictly in accordance with [1], with no experimental errors, signaling, or contextual biases involved. In this case, the values $(+1, +1)$ for paired measurements $(V_i, W_{i\oplus_5 1})$ are excluded by quantum theory.

Consequently, the degree of contextuality in the EPRB-system is

$$\mathsf{CNTX}\left(\mathfrak{S},\mathfrak{C}\right)=\frac{1}{2}\max\left(0,\mathsf{s}_{odd}\left(\left\langle V_iW_{i\oplus_4 1}\right\rangle : i\in\{1,\ldots,4\}\right)-2-\sum_{i=1}^{4}\left|\left\langle V_i\right\rangle-\left\langle W_i\right\rangle\right|\right), \tag{55}$$

and the system has a noncontextual description if and only if

$$\mathsf{s}_{odd}\left(\left\langle V_iW_{i\oplus_4 1}\right\rangle : i\in\{1,\ldots,4\}\right)\leq 2+\sum_{i=1}^{4}\left|\left\langle V_i\right\rangle-\left\langle W_i\right\rangle\right|. \tag{56}$$

Proof. The computer-assisted proof is based on Lemma 36 below, and its details, omitted here, can be found in Ref. [18], Appendix and Theorems 3 and 5. □

Lemma 36. *The necessary and sufficient condition for the connections* $\{(V_i,W_i):i\in\{1,\ldots,4\}\}$ *to be compatible with the observed pairs* $\{(V_1,W_2),(V_2,W_3),(V_3,W_4),(V_4,W_1)\}$ *is*

$$\mathsf{s}_{odd}\left(\left\langle V_iW_{i\oplus_5 1}\right\rangle,\left\langle V_iW_i\right\rangle : i\in\{1,\ldots,4\}\right)\leq 6, \tag{57}$$

which can be equivalently written as

$$\begin{aligned}&\mathsf{s}_{odd}\left(\left\langle V_iW_{i\oplus_4 1}\right\rangle : i\in\{1,\ldots,4\}\right)+\mathsf{s}_{even}\left(\left\langle V_iW_i\right\rangle : i\in\{1,\ldots,4\}\right)\leq 6,\\&\mathsf{s}_{even}\left(\left\langle V_iW_{i\oplus_4 1}\right\rangle : i\in\{1,\ldots,4\}\right)+\mathsf{s}_{odd}\left(\left\langle V_iW_i\right\rangle : i\in\{1,\ldots,4\}\right)\leq 6.\end{aligned} \tag{58}$$

5.2 Special case

In a CC EPRB-system, $\Delta_0(\mathfrak{C})=0$, the criterion for noncontextuality acquires the form

$$\mathsf{s}_{odd}\left(\left\langle V_iW_{i\oplus_4 1}\right\rangle : i\in\{1,\ldots,4\}\right)\leq 2, \tag{59}$$

which is the standard CHSH inequalities [4–6], presentable in a more familiar way as

$$\begin{array}{c}-2\leq\left\langle V_1W_2\right\rangle+\left\langle V_2W_3\right\rangle+\left\langle V_3W_4\right\rangle-\left\langle V_4W_1\right\rangle\leq 2,\\-2\leq\left\langle V_1W_2\right\rangle+\left\langle V_2W_3\right\rangle-\left\langle V_3W_4\right\rangle+\left\langle V_4W_1\right\rangle\leq 2,\\-2\leq\left\langle V_1W_2\right\rangle-\left\langle V_2W_3\right\rangle+\left\langle V_3W_4\right\rangle+\left\langle V_4W_1\right\rangle\leq 2,\\-2\leq-\left\langle V_1W_2\right\rangle+\left\langle V_2W_3\right\rangle+\left\langle V_3W_4\right\rangle+\left\langle V_4W_1\right\rangle\leq 2.\end{array} \tag{60}$$

6 SZLG-systems

6.1 Main Theorem

Theorem 37 (contextuality measure and criterion for SZLG-systems)**.** *In an SZLG-system* $(\mathfrak{S},\mathfrak{C})$, *with*

$$\mathfrak{S}=\left\{\left(V_i,W_{i\oplus_3 1}\right):i\in\{1,2,3\}\right\},\mathfrak{C}=\left\{\left(V_i,W_i\right):i\in\{1,2,3\}\right\}, \tag{61}$$

where $\oplus_3$ *stands for circular addition of* 1 *on* $\{1,2,3\}$,

$$\Delta_0\left(\mathfrak{C}\right)=\frac{1}{2}\sum_{i=1}^{3}\left|\left\langle V_i\right\rangle-\left\langle W_i\right\rangle\right|, \tag{62}$$

$$\Delta_{\min}(\mathfrak{S}, \mathfrak{C}) = \frac{1}{2}\max\left(2\Delta_0(\mathfrak{C}), \mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_3 1}\rangle : i \in \{1,2,3\}\right) - 1\right). \tag{63}$$

Consequently, the degree of contextuality in the SZLG-system is

$$\mathsf{CNTX}(\mathfrak{S}, \mathfrak{C}) = \frac{1}{2}\max\left(0, \mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_3 1}\rangle : i \in \{1,2,3\}\right) - 1 - \sum_{i=1}^{3}\left|\langle V_i\rangle - \langle W_i\rangle\right|\right), \tag{64}$$

and the system has a noncontextual description if and only if

$$\mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_3 1}\rangle : i \in \{1,2,3\}\right) \leq 1 + \sum_{i=1}^{3}\left|\langle V_i\rangle - \langle W_i\rangle\right|. \tag{65}$$

Proof. The computer-assisted proof is based on Lemma 38 below, and its details, omitted here, can be found in Ref. [18], Appendix and Theorems 4 and 6. □

Lemma 38. *The necessary and sufficient condition for the connections* $\{(V_i, W_i) : i \in \{1,2,3\}\}$ *to be compatible with the observed pairs* $\{(V_i, W_{i\oplus_3 1}) : i \in \{1,2,3\}\}$ *is*

$$\mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_3 1}\rangle, \langle V_i W_i\rangle : i \in \{1,2,3\}\right) \leq 4, \tag{66}$$

which can be equivalently written as

$$\begin{aligned}
&\mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_4 1}\rangle : i \in \{1,2,3\}\right) + \mathsf{s}_{even}\left(\langle V_i W_i\rangle : i \in \{1,2,3\}\right) \leq 4,\\
&\mathsf{s}_{even}\left(\langle V_i W_{i\oplus_4 1}\rangle : i \in \{1,2,3\}\right) + \mathsf{s}_{odd}\left(\langle V_i W_i\rangle : i \in \{1,2,3\}\right) \leq 4.
\end{aligned} \tag{67}$$

6.2 Special cases

If the SZLG-system is a CC-system, $\Delta_0(\mathfrak{C}) = 0$, the criterion for noncontextuality acquires the form

$$\mathsf{s}_{odd}\left(\langle V_i W_{i\oplus_3 1}\rangle : i \in \{1,2,3\}\right) \leq 1. \tag{68}$$

This can be written in the more familiar (Suppes-Zanotti's) form [8] as

$$-1 \leq \langle V_1 W_2\rangle + \langle V_2 W_3\rangle + \langle V_3 W_1\rangle \leq 1 + 2\max\left(\langle V_1 W_2\rangle, \langle V_2 W_3\rangle, \langle V_3 W_1\rangle\right). \tag{69}$$

Remark 39. In the temporal version of SZLG-systems (the Leggett-Garg paradigm proper), V_1 and W_1 are the results of the first measurement in both (V_1, W_2) and (V_3, W_1). They therefore cannot be influenced by later measurements (no signaling back in time). Consequently, if there are no contextual measurement biases, $V_1 \sim W_1$, and

$$\Delta_0(\mathfrak{C}) = \frac{1}{2}\sum_{i=2}^{3}\left|\langle V_i\rangle - \langle W_i\rangle\right|. \tag{70}$$

Including $|\langle V_1\rangle - \langle W_1\rangle|$, however, does not hurt, and it allows one to accommodate cases with non-temporal measurements and biased measurements (when knowing whether variable 1 will be paired with 2 or with 3 changes the way one measures 1).

7 Comparing the Systems

The main theorems regarding our three systems, Theorems 31, 35, and 37, strongly suggest the following generalization, which we formulate as a conjecture.

Remark 40. As of February 2015 we have a proof of this conjecture. A proof of the supporting Lemma 42 is given in Ref. [9].

Conjecture 41. *Let* $(\mathfrak{S}, \mathfrak{C})$ *be a system with*

$$\mathfrak{S} = \left\{ \left(V_i, W_{\pi(i)} \right) : i \in \{1, \ldots, n\} \right\}, \tag{71}$$

where $\pi(\{1, \ldots, n\})$ *is a circular (having a single cycle) permutation of* $\{1, \ldots, n\}$*, and*

$$\mathfrak{C} = \{(V_i, W_i) : i \in \{1, \ldots, n\}\}. \tag{72}$$

Then

$$\Delta_0(\mathfrak{C}) = \frac{1}{2} \sum_{i=1}^{n} |\langle V_i \rangle - \langle W_i \rangle|, \tag{73}$$

and

$$\Delta_{\min}(\mathfrak{S}, \mathfrak{C}) = \frac{1}{2} \max\left(2\Delta_0(\mathfrak{C}),\ \mathsf{s}_{odd}\left(\left\langle V_i W_{\pi(i)} \right\rangle : i \in \{1, \ldots, n\}\right) - n + 2\right). \tag{74}$$

Consequently, the degree of contextuality in this system is

$$\mathsf{CNTX}(\mathfrak{S}, \mathfrak{C}) = \frac{1}{2} \max\left(0, \mathsf{s}_{odd}\left(\left\langle V_i W_{\pi(i)} \right\rangle : i \in \{1, \ldots, n\}\right) - n + 2 - \sum_{i=1}^{n} |\langle V_i \rangle - \langle W_i \rangle|\right), \tag{75}$$

and the system has a noncontextual description if and only if

$$\mathsf{s}_{odd}\left(\left\langle V_i W_{\pi(i)} \right\rangle : i \in \{1, \ldots, n\}\right) \leq n - 2 + \sum_{i=1}^{n} |\langle V_i \rangle - \langle W_i \rangle|. \tag{76}$$

The corresponding generalization of the supporting lemmas is

Lemma 42. *The necessary and sufficient condition for the connections* $\{(V_i, W_i) : i \in \{1, \ldots, n\}\}$ *to be compatible with the observed pairs* $\left\{\left\langle V_i, W_{\pi(i)} \right\rangle : i \in \{1, \ldots, n\}\right\}$ *is*

$$\mathsf{s}_{odd}\left(\left\langle V_i W_{\pi(i)} \right\rangle, \langle V_i W_i \rangle : i \in \{1, \ldots, n\}\right) \leq 2n - 2, \tag{77}$$

which can be equivalently written as

$$\begin{aligned} &\mathsf{s}_{odd}\left(\left\langle V_i W_{\pi(i)} \right\rangle : i \in \{1, \ldots, n\}\right) + \mathsf{s}_{even}\left(\langle V_i W_i \rangle : i \in \{1, \ldots, n\}\right) \leq 2n - 2, \\ &\mathsf{s}_{even}\left(\left\langle V_i W_{\pi(i)} \right\rangle : i \in \{1, \ldots, n\}\right) + \mathsf{s}_{odd}\left(\langle V_i W_i \rangle : i \in \{1, \ldots, n\}\right) \leq 2n - 2. \end{aligned} \tag{78}$$

It is easy to see that the criterion and measure for the SZLG-system (Theorem 35) is a special case of those for the EPRB-system (Theorem 37) which in turn is a special case of those for the KCBS system. Specifically, by putting $\langle V_5 W_1 \rangle = 1$ in the KCBS-system, and assuming in addition that $W_5 \sim V_5$, so that $\langle V_5 W_5 \rangle = 1$ in the maximal coupling, we can replace W_5 in $\langle V_4 W_5 \rangle$ with W_1

and obtain the EPRB-system. By putting then $\langle V_4 W_1\rangle = 1$ in the EPRB-system, and $W_4 \sim V_4$, so that $\langle V_4 W_4\rangle = 1$ in the maximal coupling, can replace W_4 in $\langle V_3 W_4\rangle$ with W_1 and obtain the SZLG-system.

It is easy to see how this pattern generalizes. First of all, any circular permutation π can be replaced, by appropriate renaming, with circular addition of 1 on $\{1,\ldots,n\}$, making the observed pairs $(V_1, W_2), \ldots, (V_{n-1}, W_n), (V_n, W_1)$. Let the SZLG, EPRB, and KCBS systems be designated as systems of order 3, 4, and 5, respectively, and the system in Conjecture 41 as an (n)-system. By putting $\langle V_n W_1\rangle = 1$ in the (n)-system, and assuming that $W_n \sim V_n$, so that $\langle V_n W_n\rangle = 1$ in the maximal coupling, we replace W_n in $\langle V_{n-1} W_n\rangle$ with W_1 and obtain an $(n-1)$-system.

Remark 43. In Ref. [9] we have proved the following theorem: the system $(\mathfrak{S}, \mathfrak{C})$ in Conjecture (41) has a noncontextual description if and only if

$$\mathsf{s}_{odd}\left(\left\langle V_i W_{\pi(i)}\right\rangle, 1 - |\langle V_i\rangle - \langle W_i\rangle| : i = 1, \ldots, n\right) \leq 2n - 2. \tag{79}$$

It is easy to show that the conjectured criterion (76) follows from (79), i.e., its violation is a sufficient condition for contextuality. The conjecture is that it is also true that (79) follows from (76). See Remark 40.

8 Conclusion

We have presented a theory of (non)contextuality in purely probabilistic terms abstracted away from physical meaning. The computational aspects of the theory are confined to finite systems with binary components, but it is easily generalizable to deal with components attaining arbitrary finite numbers of values. As the components of the bunches get more complex and/or connections get longer than pairs, the generalizations become less unique.

The basis for the theory is the principle of Contextuality-by-Default, which has philosophical and mathematical consequences. Mathematically, it leads to revamping (while remaining within its confines) of the Kolmogorovian probability theory, with more prominent than usual emphasis on stochastic unrelatedness. A reformulated theory may even avoid the notion of a sample space altogether [21].

Philosophically, the principle of Contextuality-by-Default may elucidate the difference between "ontic" and "epistemic" aspects of contextuality (perhaps even of probability theory generally).

The theory also offers pragmatic advantages: it allows for non-CC-systems (whether due to signaling or due to context-dependent measurement biases) and for experimental/computational errors in the analysis (including statistical analysis) of experimental data.

The theory has predecessors in the literature. The idea of labeling differently random variables in different contexts and considering the probability with which they can be equal to each other if coupled has been prominently used in Refs. [30, 31].

Acknowledgments

This work is supported by NSF grant SES-1155956 and AFOSR grant FA9550-14-1-0318. We have benefited from collaboration with J. Acacio de Barros, and Gary Oas, as well as from discussions with Samson Abramsky, Guido Bacciagaluppi, and Andrei Khrennikov. An abridged version of this paper was presented at the Purdue Winer Memorial Lectures in November 2014.

References

[1] A.A. Klyachko, M.A. Can, S. Binicioglu, and A.S. Shumovsky (2008). A simple test for hidden variables in spin-1 system. Phys. Rev. Lett. 101, 020403.

[2] J. Bell (1964). On the Einstein-Podolsky-Rosen paradox. Physics 1: 195-200.

[3] J. Bell (1966). On the problem of hidden variables in quantum mechanics. Rev. Modern Phys. 38: 447-453.

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

[5] J.F. Clauser, M.A. Horne (1974). Experimental consequences of objective local theories. Phys. Rev. D 10: 526–535.

[6] A. Fine (1982). Hidden variables, joint probability, and the Bell inequalities. Phys. Rev. Lett. 48: 291–295.

[7] A.J. Leggett & A. Garg (1985). Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? Phys. Rev. Lett. 54: 857–860.

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

[9] J.V. Kujala, E.N. Dzhafarov, & J.-Å. Larsson (2014). Necessary and sufficient conditions for maximal noncontextuality in a broad class of quantum mechanical systems. arXiv:1412.4724.

[10] E. Specker (1960). Die Logik Nicht Gleichzeitig Entscheidbarer Aussagen. Dialectica 14: 239–246 (English translation by M.P. Seevinck available as arXiv:1103.4537.)

[11] E.N. Dzhafarov, J.V. Kujala (2013). All-possible-couplings approach to measuring probabilistic context. PLoS ONE 8(5): e61712. doi:10.1371/journal.pone.0061712.

[12] E.N. Dzhafarov, J.V. Kujala (2014). No-Forcing and No-Matching theorems for classical probability applied to quantum mechanics. Found. Phys. 44: 248–265.

[13] E.N. Dzhafarov, J.V. Kujala (2014). Embedding quantum into classical: contextualization vs conditionalization. PLoS One 9(3): e92818. doi:10.1371/journal.pone.0092818.

[14] E.N. Dzhafarov, J.V. Kujala (2014). Contextuality is about identity of random variables. Phys. Scripta T163, 014009.

[15] E.N. Dzhafarov, J.V. Kujala (2015). Random variables recorded under mutually exclusive conditions: Contextuality-by-Default. In H. Liljenström (Ed.) Advances in Cognitive Neurodynamics IV (pp. 405-410).

[16] J.A. de Barros, E.N. Dzhafarov, J.V. Kujala, G. Oas (2014). Measuring observable quantum contextuality. arXiv:1406.3088

[17] E.N. Dzhafarov, J.V. Kujala (2014). Probabilistic contextuality in EPR/Bohm-type systems with signaling allowed. arXiv:1406.0243.

[18] E.N. Dzhafarov, J.V. Kujala (2014). Generalizing Bell-type and Leggett-Garg-type inequalities to systems with signaling. arXiv:1407.2886.

[19] E.N. Dzhafarov, J.V. Kujala (2014). A qualified Kolmogorovian account of probabilistic contextuality. Lect. Not. Comp. Sci. 8369: 201–212.

[20] E.N. Dzhafarov, J.V. Kujala (2013). Order-distance and other metric-like functions on jointly distributed random variables. Proc. Amer. Math. Soc. 141(9): 3291–3301.

[21] E.N. Dzhafarov, J.V. Kujala (2013). Probability, random variables, and selectivity. arXiv:1312.2239.

[22] A. Khrennikov (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.

[23] A. Khrennikov (2008). Bell-Boole inequality: Nonlocality or probabilistic incompatibility of random variables? Entropy 10, 19-32.

[24] A. Khrennikov (2009). Contextual approach to quantum formalism. Springer: Berlin.

[25] E.N. Dzhafarov, J.V. Kujala (2010). The Joint Distribution Criterion and the Distance Tests for selective probabilistic causality. Frontiers in Quantitative Psychology and Measurement. 1:151 doi: 10.3389/fpsyg.2010.00151.

[26] G. Bacciagaluppi (2015). Leggett-Garg inequalities, pilot waves and contextuality. Int. J. Quant. Found. 1, 1-17.

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

[28] Ll. Masanes, A. Acin, & N. Gisin (2006). General properties of nonsignaling theories. Phys. Rev. A 73: 012112.

[29] S. Popescu & D. Rohrlich (1994). Quantum nonlocality as an axiom. Found. Phys. 24: 379–385.

[30] J.-Å. Larsson (2002). A Kochen-Specker inequality. Europhys. Lett., 58: 799–805.

[31] K. Svozil (2012). How much contextuality? Natural Computing 11, 261-265.