Contextuality-by-Default: A Brief Overview of Ideas, Concepts, and Terminology

Ehtibar N. Dzhafarov[1], Janne V. Kujala[2], and Victor H. Cervantes[1]

[1]Purdue University
ehtibar@purdue.edu

[2]University of Jyväskylä
jvk@iki.fi

Abstract. This paper is a brief overview of the concepts involved in measuring the degree of contextuality and detecting contextuality in systems of binary measurements of a finite number of objects. We discuss and clarify the main concepts and terminology of the theory called "contextuality-by-default," and then discuss generalizations of the theory to arbitrary systems of arbitrary random variables.

Keywords: contextuality, contextuality-by-default, connection, coupling, cyclic system, inconsistent connectedness, measurements.

1 Introduction

1.1 On the name of the theory

The name "contextuality-by-default" should not be understood as suggesting that any system of measurements is contextual, or contextual unless proven otherwise. The systems are contextual or noncontextual depending on certain criteria, to be described. The name of the theory reflects a philosophical position according to which every random variable's identity is inherently contextual, i.e., it depends on all conditions under which it is recorded, whether or not there is a way in which these conditions could affect the random variable physically. Thus, in the well-known EPR-Bell paradigm, Alice and Bob are separated by a space-like interval that prevents Bob's measurements from being affected by Alice's settings; nevertheless, Bob's measurements should be labeled by both his own setting and by Alice's setting; and as the latter changes with the former fixed, the identity of the random variable representing Bob's measurement changes "by default." One does not have to ask "why." Bob's measurements under two different settings by Alice have no probabilisitic relation to each other; they possess no joint distribution. Therefore one cannot even meaningfully ask the question of whether the two may be "always equal to each other." The questions one can ask meaningfully are all about what joint distributions can be *imposed* (in a well-defined sense) on the system in a way consistent with observations. A system is contextual or noncontextual (or, as we say, has a noncontextual description) depending on the answers to these questions. Thus, the famous Kochen-Specker demonstration of contextuality is, from this point of view, a reductio ad absurdum proof that measurements of a spin in a fixed direction made under different conditions cannot be imposed a joint distribution upon, in which these measurements would be equal to each other with probability 1.

1.2 Notation

We use capital letters $A, B, \ldots, Q$ to denote sets of "objects" (properties, quantities) being measured, and the script letter $\mathscr{C}$ to denote a collection of such sets. We use capital letters R, S, and T to denote the measurements (random variables), and the Gothic letter $\mathfrak{R}$ to denote sets of random variables that do not possess a joint distribution.

2 Contextuality-by-Default at a Glance

The following is an overview of the main concepts and definitions of the contextuality-by-default theory. This is not done at a very high level of generality, in part in order not to be too abstract, and in part because the criterion and measure of contextuality have been developed in detail only for a broad subclass of finite sets of binary measurements. Thus, the notion of a context given below in terms of subsets of measured objects is limited, but it is intuitive, and a way to generalize it is clear (Section 5). The definition of maximally connected couplings is given for binary (± 1) measurements only, and reasonable generalizations here may not be unique. In Section 5 we discuss one, arguably most straightforward way of doing this.

2.1 Measurements are labeled contextually

There is a set Q of "*objects*" we want to measure. For whatever reason, we cannot measure them "*all at once*" (the meaning of this is not necessarily chronological, as explained in Section 4). Instead we define a collection of subsets of Q,

$$\mathscr{C} = \{A \subset Q, B \subset Q, \ldots\}, \tag{1}$$

and measure the objects "*one subset at a time*." We call these subsets of objects *contexts*. Different contexts may overlap. This definition has limited applicability, and we discuss a general definition in Section 5. For now we consider only finite sets Q (hence finite collections of finite contexts).

The *measurement outcome* of each object q (from the set Q) in each context C (from the collection $\mathscr{C}$) is a random variable, and we denote it R_q^C (with $q \in C \in \mathscr{C}$). This is called *contextual labeling* of the measurement outcomes. It ensures that the collection

$$\left\{R_q^A\right\}_{q\in A}, \left\{R_q^B\right\}_{q\in B}, \ldots \tag{2}$$

for all $A, B, \ldots$ comprising $\mathscr{C}$, are pairwise disjoint, as no two random variables taken from two different members of the collection have the same superscript (whether or not they have the same subscript).

2.2 Measurements in different contexts are stochastically unrelated

We call $R^C = \left\{R_q^C\right\}_{q\in C}$ for every $C \in \mathscr{C}$ a *bunch* (of random variables). The random variables within a bunch are *jointly distributed*, because of which we can consider each bunch as a single (multicomponent) random variable. If $q, q' \in C \in \mathscr{C}$, we can answer questions like "what is the correlation between R_q^C and $R_{q'}^C$?". However, if $q \in C$, $q' \in C'$, and $C \neq C'$, then we cannot answer such questions: R_q^C and $R_{q'}^{C'}$ belong to different bunches and do not have a joint distribution. We say that they are *stochastically unrelated*.

2.3 All possible couplings for all measurements

Consider now the (necessarily disjoint) union of all bunches

$$\mathfrak{R} = \bigcup_{C\in\mathscr{C}} R^C = \bigcup_{C\in\mathscr{C}} \left\{R_q^C\right\}_{q\in C}, \tag{3}$$

i.e., the set of all measurements contextually labeled. The use of the Gothic font is to emphasize that this set is *not* a multicomponent random variable: except within bunches, its components are not jointly distributed. We call $\mathfrak{R}$ a *system* (of measurements).

Now, we can be interested in whether and how one could *impose a joint distribution* on $\mathfrak{R}$. To impose a joint distribution on $\mathfrak{R}$ means to find a set of jointly distributed random variables $S = \left\{S_q^C\right\}_{q\in C\in\mathscr{C}}$ such that, for every $C \in \mathscr{C}$,

$$S^C = \left\{S_q^C\right\}_{q\in C} \sim \left\{R_q^C\right\}_{q\in C} = R^C. \tag{4}$$

The symbol $\sim$ means "has the same distribution as." Note that

$$S = \left\{S^C\right\}_{C\in\mathscr{C}} = \left\{S_q^C\right\}_{q\in C\in\mathscr{C}} \tag{5}$$

is a single (multicomponent) random variable, and in probability theory S is called a *coupling* for (or of) $\mathfrak{R}$. Any subset of the components of S is its *marginal*, and S^C is the marginal of S whose components are labeled in the same way as are the components of the bunch R^C.

If no additional constraints are imposed, one can always find a coupling S for any union of bunches. For instance, one can always use an *independent coupling*: create a *copy* S^C of each bunch R^C (i.e., an identically labelled and identically distributed set of random variables), and join them so that they are stochastically independent. The set $S = \left\{S^C\right\}_{C\in\mathscr{C}}$ is then jointly distributed. The existence of a coupling *per se* therefore is not informative.

2.4 Connections and their couplings

Let us form, for every object q, a set of random variables

$$\mathfrak{R}_q = \left\{R_q^C\right\}_{C\in\mathscr{C}}, \tag{6}$$

i.e., all random variables measuring the object q, across all contexts, We call this set, which is not a random variable, a *connection* (for q). Let us adopt the convention that if a context C does not contain q, then R_q^C is not defined and does not enter in $\mathfrak{R}_q$.

A system is called *consistently connected* if, for every $q \in Q$ and any two contexts C, C' containing q,

$$R_q^C \sim R_q^{C'}. \tag{7}$$

Otherwise a system is called (strictly) *inconsistently connected*. Without the adjective "strictly," inconsistent connectedness means that the equality above is not assumed, but it is then not excluded either: consistent connectedness is a special case of inconsistent connectedness.

One possible interpretation of strictly inconsistent connectedness is that the conditions under which a context is recorded may physically influence (in some cases one could say, "signal to") the measurements of the context members. Another possibility is that a choice of context may introduce biases in how the objects are measured and recorded.

2.5 Maximally connected couplings for binary measurements

Every coupling S for $\mathfrak{R}$ has a marginal $S_q = \left\{S_q^C\right\}_{C\in\mathscr{C}}$ that forms a coupling for the connection $\mathfrak{R}_q$. We can also take $\mathfrak{R}_q$ for a given q in isolation, and consider all its couplings $T_q = \left\{T_q^C\right\}_{C\in\mathscr{C}}$. Clearly, the set of all S_q extracted from all possible couplings S for $\mathfrak{R}$ is a subset of all possible couplings T_q for $\mathfrak{R}_q$.

Let us now confine the consideration to *binary measurements*: each random variable in the system has value $+1$ or -1. In Section 5 we will discuss possible generalizations.

A coupling T_q for a connection $\mathfrak{R}_q$ is called *maximal* if, given the expected values $\left\langle R_q^C\right\rangle$ for all C, the value of

$$\mathsf{eq}\left(T_q\right) = \Pr\left[T_q^C = 1 : C \in \mathscr{C}\right] + \Pr\left[T_q^C = -1 : C \in \mathscr{C}\right] \tag{8}$$

is the largest possible among all couplings for $\mathfrak{R}_q$ (again, R_q^C and T_q^C are not defined and are not considered if C does not contain q).

Let us denote

$$\max \mathsf{eq}\left(\mathfrak{R}_q\right) = \max_{\substack{\text{all possible}\\ \text{couplings } T_q \text{for } \mathfrak{R}_q}} \mathsf{eq}\left(T_q\right). \tag{9}$$

It follows from a general theorem mentioned in Section 5 that this quantity is well-defined for all systems, i.e., that the supremum of $\mathsf{eq}\left(T_q\right)$ is attained in some coupling T_q. Clearly, for consistently connected systems $\max \mathsf{eq}\left(\mathfrak{R}_q\right) = 1$ (the measurements can be made "perfectly correlated"). For (strictly) inconsistently connected systems, $\max \mathsf{eq}\left(\mathfrak{R}_q\right)$ is always well-defined, and it is less than 1 for some q. It may even be zero: for ± 1 variables this happens when the $\mathfrak{R}_q$ contains two measurements R_q^A and R_q^B such that $\Pr\left[R_q^A = 1\right] = 1$ and $\Pr\left[R_q^B = 1\right] = 0$.

2.6 Definition of contextuality

Consider again a coupling S for the entire system $\mathfrak{R}$, and for every $q \in Q$, extract from S the marginal S_q that forms a coupling for the connection $\mathfrak{R}_q$.

Central Concept. If, for every $q \in Q$,

$$\mathsf{eq}\left(S_q\right) = \max \mathsf{eq}\left(\mathfrak{R}_q\right), \tag{10}$$

(i.e., if every marginal S_q in S is a maximal coupling for $\mathfrak{R}_q$) then the coupling S for $\mathfrak{R}$ is said to be *maximally connected*.

Intuitively, in this case the measurements can be imposed a joint distribution upon in which the measurements R_q^C of every object q in different contexts C are maximally "correlated," i.e., attain one and the same value with the maximal probability allowed by their observed individual distributions (expectations).

Main Definition. A system $\mathfrak{R}$ is said to be *contextual* if no coupling S of this system is maximally connected. Otherwise, a maximally connected coupling of the system (it need not be unique if it exists) is said to be this system's *noncontextual description* (or, as a terminological variant, *maximally noncontextual description*).

For consistently connected systems this definition is equivalent to the traditional understanding of (non)contextuality. According to the latter, a system has a noncontextual description if and only if there is a coupling for the measurements *labeled noncontextually*. The latter means that all random variables R_q^C within a connection are treated as being equal to each other with probability 1.

2.7 Measure and criterion of contextuality

If (and only if) a system $\mathfrak{R}$ is contextual, then for every coupling S there is at least one $q \in Q$ such that $\mathsf{eq}\left(S_q\right) < \max \mathsf{eq}\left(\mathfrak{R}_q\right)$. This is equivalent to saying that a system is contextual if and only if for every coupling S of it,

$$\sum_{q \in Q} \mathsf{eq}\left(S_q\right) < \sum_{q \in Q} \max \mathsf{eq}\left(\mathfrak{R}_q\right). \tag{11}$$

Define

$$\max \mathsf{eq}\left(\mathfrak{R}\right) = \max_{\substack{\text{all couplings} \\ S \text{ for } \mathfrak{R}}} \left(\sum_{q \in Q} \mathsf{eq}\left(S_q\right) \right). \tag{12}$$

In this definition we assume that this maximum exists, i.e., the supremum of the sum on the right is attained in some coupling S. (This is likely to be true for all systems with finite Q and binary measurements, but we only have a formal proof of this for the cyclic systems considered below.) Then (11) is equivalent to

$$\max \mathsf{eq}\left(\mathfrak{R}\right) < \sum_{q \in Q} \max \mathsf{eq}\left(\mathfrak{R}_q\right), \tag{13}$$

which is a *criterion of contextuality* (necessary and sufficient condition for it). Moreover, it immediately leads to a natural *measure of contextuality*:

$$\mathsf{cntx}\left(\mathfrak{R}\right)=\sum_{q\in Q}\max\mathsf{eq}\left(\mathfrak{R}_{q}\right)-\max\mathsf{eq}\left(\mathfrak{R}\right). \tag{14}$$

Written *in extenso* using (9) and (12),

$$\mathsf{cntx}\left(\mathfrak{R}\right)=\sum_{q\in Q}\max_{\substack{\text{all possible}\\ \text{couplings } T_{q} \text{for } \mathfrak{R}_{q}}}\mathsf{eq}\left(T_{q}\right)-\max_{\substack{\text{all couplings}\\ S \text{ for } \mathfrak{R}}}\sum_{q\in Q}\mathsf{eq}\left(S_{q}\right), \tag{15}$$

where, one should recall, S_q is the marginal of S that forms a coupling for $\mathfrak{R}_q$. We can see that the minuend and subtrahend in the definition of $\mathsf{cntx}\left(\mathfrak{R}\right)$ differ in order of the operations max and $\sum_{q\in Q}$; and while in the minuend the choice of couplings T_q for $\mathfrak{R}_q$ is unconstrained, in the subtrahend the choice of couplings S_q for $\mathfrak{R}_q$ is constrained by the requirement that it is a marginal of the coupling for the entire system $\mathfrak{R}$.

3 The history of the contextuality-by-default approach

A systematic realization of the idea of contextually labeling a system of measurements $\mathfrak{R}$, considering all possible couplings S for it, and characterizing it by the marginals S_q that form couplings for the connections $\mathfrak{R}_q$ of the system was developing through a series of publications [3–7]. The idea of maximally connected couplings as the central concept for contextuality in consistently connected systems was proposed in Refs. [2,8,11] and then generalized to inconsistently connected systems [9,10].

In the latter two references the measure of contextuality (14) and the criterion of contextuality (13) were defined and computed for simple QM systems (cyclic systems of rank 3 and 4, as defined below). Later we added to this list cyclic systems of rank 5, and formulated a conjecture for the measure and criterion formulas for cyclic systems of arbitrary rank [12].

In Refs. [12,18] the contextuality-by-default theory is presented in its current form. The conjecture formulated in Ref. [12] was proved in Ref. [19].

A cyclic system (with binary measurements) is defined as one involving n "objects" (n being called the *rank* of the system) measured two at a time,

$$\left(q_{1},q_{2}\right),\left(q_{2},q_{3}\right),\ldots,\left(q_{n-1},q_{n}\right),\left(q_{n},q_{1}\right). \tag{16}$$

For $i=1,\ldots,n$, the pair $\left(q_{i},q_{i\oplus1}\right)$ forms the context C_i ($\oplus$ standing for circular shift by 1). Each object q_i enters in precisely two consecutive contexts, $C_{i\ominus1}$ and C_i. Denoting the measurement of q_i in context C_j by R_i^j, we have the system represented by bunches $R^{i}=\left(R_{i}^{i},R_{i\oplus1}^{i}\right)$ and connections $\mathfrak{R}_{i}=\left(R_{i}^{i\ominus1},R_{i}^{i}\right)$.

The formula for the measure of contextuality conjectured in Ref. [12] and proved in Ref. [19] is

$$\mathsf{cntx}\left(\mathfrak{R}\right)=\frac{1}{2}\max\begin{cases}\mathrm{s_{odd}}\left(\left\langle R_{i}^{i}R_{i\oplus1}^{i}\right\rangle :i=1,\ldots,n\right)-\sum_{i=1}^{n}\left|\left\langle R_{i}^{i}\right\rangle -\left\langle R_{i}^{i\ominus1}\right\rangle \right|-(n-2)\\ 0\end{cases}. \tag{17}$$

The function $\mathrm{s_{odd}}$ is defined for an arbitrary set of argument $x_1, \dots, x_k$ as

$$\mathrm{s_{odd}}\left(x_1, \dots, x_k\right) = \max\left(\pm x_1 \pm \dots \pm x_k\right), \tag{18}$$

where the maximum is taken over all assignments of $+$ and $-$ signs with an odd number of $-$'s. The criterion of contextuality readily derived from (17) is: the system is contextual if ands only if

$$\mathrm{s_{odd}}\left(\left\langle R_i^i R_{i\oplus 1}^i \right\rangle : i = 1, \dots, n\right) > (n-2) + \sum_{i=1}^{n} \left|\left\langle R_i^i \right\rangle - \left\langle R_i^{i\ominus 1} \right\rangle\right|. \tag{19}$$

For consistently connected systems, the sum on the right vanishes, and we can derive the traditional formulas for Legget-Garg ($n = 3$), EPR/Bell ($n = 4$), and Klyachko-Can-Binicioglu-Shumovsky-type (KCBS) systems ($n = 5$). But the formula also allows us to deal with the same experimental paradigm when they create inconsistently connected systems, due to signaling or contextual biases in experimental design (see Refs. [9, 10, 12, 18] for details).

The contextuality-by-default theory does have precursors in the literature. The idea that random variables in different contexts are stochastically unrelated was prominently considered in Refs. [13, 15–17]. Probabilities of the eq-type with the contextual labeling of random variables, as defined in (8), were introduced in Refs. [20, 21, 23, 25]. The distinguishing feature of the contextuality-by-default theory is the notion of a maximally connected coupling, which in turn is based on the idea of comparing maximal couplings for the connections taken in isolation and those extracted as marginals from the couplings of the entire system. Contextuality-by-default is a more systematic and more general theory of contextuality than those proposed previously, also more readily applicable to experimental data [1, 18].

4 Conceptual and Terminological Clarifications

4.1 Contextuality and quantum mechanics (QM)

The notion of (non)contextuality has its origins in logic [22], but since the publication of Ref. [14] it has been widely considered a QM notion. QM indeed provides the only known to us theoretically justified examples of contextual systems. (Non)contextuality *per se*, however, is a purely probabilisitic concept, squarely within the classical, Kolmogorovian probability theory (that includes the notion of stochastic unrelatedness and that of couplings) [7, 8, 11]. When contextuality is present in a QM system, QM is relevant to answering the question of exactly how the noncontextuality conditions in the system are violated, but it is not relevant to the question of what these conditions are.

4.2 Contexts and QM observables

In particular, the "objects" being measured need not be QM observables. They may very well be questions asked in a poll of public opinion, and the binary measurements then may be Yes/No answers to these questions. It is especially important not to confuse being "measured together" in the definition of a context with being represented by compatible (commutative) observables. Thus, in the theory of contextuality in cyclic systems, $n = 4$ is exemplified by the EPR-Bell paradigm, with Alice's "objects" (spins) being q_1, q_3 and Bob's q_2, q_4. In each of the contexts $(q_1, q_2), \ldots, (q_4, q_1)$, the two objects are compatible in the trivial sense: any observable in Alice's Hilbert space H_A is compatible with any observable in Bob's Hilbert space H_B because the joint space is the tensor product $H_A \otimes H_B$. The case $n = 5$ is exemplified by the KCBS paradigm, where the spins $q_1, \ldots, q_5$ are represented by observables in three-dimensional Hilbert space. In each of the five contexts $(q_1, q_2), \ldots, (q_5, q_1)$ the observables are compatible in the narrow QM sense: they are commuting Hermitian operators. The case $n = 3$ is exemplified by the Leggett-Garg paradigm, where three measurements are made at three distinct time moments, two measurements at a time. The QM representations of the observables in each of the contexts $(q_1, q_2), (q_2, q_3), (q_3, q_1)$ are generally incompatible (noncommuting) operators. In spite of the profound differences in the QM structure of these three cyclic systems, their contextually analysis is precisely the same mathematically, given by (17) and (19).

4.3 The meaning of being measured "together"

It should be clear from the discussion of the Leggett-Garg paradigm that "measuring objects one context at a time" does not necessarily have the meaning of chronological simultaneity. Rather one should think of measurements being grouped and recorded in accordance with some fixed coupling scheme: if q and q' belong to the same context C, there is an empirical procedure by which observations of R_q^C are paired with observations of R_{q}^C. Thus, if the objects being measured are tests taken by students, and the measurements are their test scores, the tests are grouped into contexts by the student who takes them, however they are distributed in time. The grouping of (potential) observations is in essence what couplings discussed in Sections 2.3 and 2.4 do for a set of stochastically unrelated random variables, except that these couplings do not provide a uniquely (empirically) defined joint distribution. Rather the probabilistic couplings imposed on different bunches are part of a purely mathematical procedure that generally yields an infinity of different joint distributions.

4.4 The meaning of a noncontextual description

In the traditional approach to (non)contextuality, where the measurements are labeled by objects but not by contexts, one can define a noncontextual description as simply a coupling imposed on the system. For instance, in the Leggett-Garg paradigm the noncontextual labeling yields three random variables, R_1, R_2, R_3, with $(R_1, R_2), (R_2, R_3), (R_3, R_1)$ jointly observed. A noncontextual description here is any three-component random variable $S = (S_1, S_2, S_3)$ with $(S_1, S_2) \sim (R_1, R_2)$, $(S_2, S_3) \sim (R_2, R_3)$, and $(S_3, S_1) \sim (R_3, R_1)$. The system is contextual if no such description exists.

The situation is different with contextually labeled measurements. For the Leggett-Garg paradigm we now have six variables grouped into three stochastically unrelated contexts, $\left(R_1^1, R_2^1\right), \left(R_2^2, R_3^2\right), \left(R_3^3, R_1^3\right)$. As explained in Section 2.3, such a system always has a coupling, in this case a sextuple S with $\left(S_1^1, S_2^1\right) \sim \left(R_1^1, R_2^1\right)$, $\left(S_2^2, S_3^2\right) \sim \left(R_2^2, R_3^2\right)$, and $\left(S_3^3, S_1^3\right) \sim \left(R_3^3, R_1^3\right)$. One can call any of these couplings a noncontextual description of the system. To characterize (non)contextuality then one can use the term "maximally noncontextual description" for any maximally connected coupling [18]. Alternatively, one can confine the term "noncontextual description" of a system only to maximally connected couplings for it. With this terminology the definition of a contextual system in our theory is the same as in the traditional approach: a system is contextual if it does not have a noncontextual description. The choice between the two terminological variants will be ultimately determined by whether couplings other than maximally connected ones will be found a useful role to play.

5 Instead of a Conclusion: Generalizations

5.1 Beyond objects and subsets

Defining a context as a subset of objects measured together [12,18] is less general than defining it by conditions under which certain objects are measured [2,9,10]. For instance, by the first of these definitions $\{q_1, q_2\}$ for a given pair of objects is a single context, while the second definition allows one to speak of the same pair of objects q_1, q_2 forming several different contexts. Thus, if q_1, q_2 are two tests, they can be given in one order or the other, (q_1, q_2) or (q_1, q_2). In fact, in all our previous discussion of cyclic systems we used the notation for ordered pairs, $(q_i, q_{i\oplus 1})$ rather than $\{q_i, q_{i\oplus 1}\}$. This is inconsequential for cyclic systems of rank $n \geq 3$. For $n = 2$, however, the difference between (q_1, q_2) and (q_2, q_1) is critical if $n = 2$ is to be a nontrivial system (with the distributions of the two bunches not identical). It can be shown that the system can be nontrivial, and $n = 2$ is a legitimate value for (17) and (19).

Being formal and mathematically rigorous here makes things simpler. A context is merely a label (say, superscript) at a random variable with the convention that identically superscripted variables are "bunched together," i.e., they are jointly distributed. An object is merely another label (in our notation, a subscript) that makes all the elements of a bunch different and indicates which elements from different bunches should be put together to form a a connection. So if there are six random variables grouped into three distinct bunches $\left(R_1^1, R_2^1\right)$, $\left(R_1^2, R_2^2\right)$, and $\left(R_1^3, R_2^3\right)$ and into two connections $\left(R_1^1, R_1^2, R_1^3\right)$ and $\left(R_2^1, R_2^2, R_2^3\right)$, we can (but do not have to) interpret this as three different contexts involving the same two objects. From mathematical (and perhaps also philosophical) point of view, measurements grouped into bunches and connections are more fundamental than objects being measured within contexts.

5.2 Beyond binary measurements

How could the definition of a maximally connected coupling be generalized to arbitrary random variables? A straightforward way to do this is to extend definition (8) for a coupling T_q of a connection $\mathfrak{R}_q$ as

$$\mathsf{eq}\left(T_q\right) = \Pr\left[T_q^C = T_q^{C'} \text{ for any two } C, C' \in \mathscr{C}\right]. \tag{20}$$

This is an approach adopted in [2, 12, 18]. It is based on the following mathematical considerations, derived from the discussion of maximal couplings in Thorisson's monograph [24] (Section 7 of Chapter 3).

Given two sigma-additive measures μ and ν on the same sigma algebra, let us write $\mu \leq \nu$ if $\mu(E) \leq \nu(E)$ for every measurable set E. Let μ_q^C be the probability measure associated with R_q^C. Let X_q and Σ_q be the set of values and sigma algebra associated with R_q^C (they are assumed the same for all C, because otherwise one should not consider R_q^C measurements of one and the same object). For every object q, define μ_q as the largest sigma-additive measure such that $\mu_q \leq \mu_q^C$ for all contexts C. The measure μ_q is the largest in the sense that $\mu'_q \leq \mu_q$ for any other measure μ'_q such that $\mu_q \leq \mu_q^C$ for all contexts C. A theorem proved in Ref. [24] (Theorem 7.1) guarantees the existence and uniqueness of μ_q, for any set of probability measures $\left\{\mu_q^C\right\}_{C\in\mathscr{C}}$, whatever the indexing set $\mathscr{C}$. That is, μ_q is uniquely defined for any connection $\mathfrak{R}_q$. Note that μ_q is not generally a probability measure, so $\mu_q\left(X_q\right)$ can be any number in $[0, 1]$. Let us denote

$$\max \mathsf{eq}\left(\mathfrak{R}_q\right) = \mu_q\left(X_q\right). \tag{21}$$

For ± 1-measurements R_q^C this definition specializes to (8)-(9).

Consider now a coupling T_q for $\mathfrak{R}_q$. It is defined on the product sigma-algebra $\bigotimes_{\mathscr{C}} \Sigma_q$ on the product set $\prod_{\mathscr{C}} X_q$. An event $E_q \in \bigotimes_{\mathscr{C}} \Sigma_q$ is called a *coupling event* if $S_q \in E_q$ implies $T_q^C = T_q^{C'}$ for any two $C, C' \in \mathscr{C}$ (assuming, as always, that both C and C' involve q). It follows from Theorem 7.2 in Ref. [24] that

$$\Pr\left[T_q \in E_q\right] \leq \max \mathsf{eq}\left(\mathfrak{R}_q\right), \tag{22}$$

for any q and any choice of E_q. Now, it is natural to define a *maximal coupling* for $\mathfrak{R}_q$ as a coupling T_q for which E_q can be chosen so that

$$\Pr\left[T_q \in E_q\right] = \max \mathsf{eq}\left(\mathfrak{R}_q\right). \tag{23}$$

Theorem 7.3 in Ref. [24] says that such a maximal coupling always exists. Note that E_q in a maximal coupling can always be thought of as the largest measurable subset of the diagonal of the set $\prod_{\mathscr{C}} X_q$.

Having established this generalized notion of a maximal coupling, the theory of contextuality can now be generalized in a straightforward fashion. Consider a coupling S for the entire system $\mathfrak{R}$. The definition of a *maximally connected coupling* remains unchanged: every marginal S_q of a maximally connected coupling S is a maximal coupling for the corresponding connection $\mathfrak{R}_q$. Our Main

Definition could remain unchanged too: a system $\mathfrak{R}$ is *contextual* if and only if no coupling S of this system is maximally connected. This can be equivalently presented as follows. For any set P of probability values, let $f(P)$ be a bounded smooth nonnegative function strictly increasing in all components of P. Thus, for finite systems of random variables f can be chosen as a sum or average, as in (11). Define

$$\max \mathsf{eq}\left(\mathfrak{R}\right) = \max_{\substack{\text{all couplings}\\ S \text{ for } \mathfrak{R}}} f\left(\Pr\left[S_q \in E_q\right] : q \in Q\right), \tag{24}$$

and, if this value exists,

$$\mathsf{cntx}\left(\mathfrak{R}\right) = f\left(\max \mathsf{eq}\left(\mathfrak{R}_q\right) : q \in Q\right) - \max \mathsf{eq}\left(\mathfrak{R}\right). \tag{25}$$

The system is defined as contextual if and only if $\mathsf{cntx}\left(\mathfrak{R}\right) > 0$. We do not know whether $\max \mathsf{eq}\left(\mathfrak{R}\right)$ exists for all possible systems of random variables. If it does not, however, the definition can be extended by replacing max with sup.

This generalization has to be further explored to determine whether it is a good generalization, i.e., whether it provides valuable insights, leads to interesting mathematical developments, and does not yield non-interpretable results when applied to specific systems of measurements.

Acknowledgments. This research has been supported by NSF grant SES-1155956 and AFOSR grant FA9550-14-1-0318.

References

1. Bacciagaluppi, G. (2015). Leggett-Garg inequalities, pilot waves and contextuality. International Journal of Quantum Foundations 1, 1-17.
2. de Barros, J.A. , Dzhafarov, E.N., Kujala, J.V., Oas, G. (2014). Measuring observable quantum contextuality. arXiv:1406.3088.
3. Dzhafarov, E.N., & Kujala, J.V. (2013). All-possible-couplings approach to measuring probabilistic context. PLoS One 8(5): e61712. doi:10.1371/journal.pone.0061712.
4. Dzhafarov, E.N., & Kujala, J.V. (2013). Order-distance and other metric-like functions on jointly distributed random variables. Proceedings of the American Mathematical Society 141, 3291-3301.
5. Dzhafarov, E.N., & Kujala, J.V. (2014). No-Forcing and No-Matching theorems for classical probability applied to quantum mechanics. Foundations of Physics 44, 248-265.
6. Dzhafarov, E.N., & Kujala, J.V. (2014). Embedding quantum into classical: contextualization vs conditionalization. PLoS One 9(3): e92818. doi:10.1371/journal.pone.0092818.
7. Dzhafarov, E.N., & Kujala, J.V. (2014). A qualified Kolmogorovian account of probabilistic contextuality. Lecture Notes in Computer Science 8369, 201-212.
8. Dzhafarov, E.N., & Kujala, J.V. (2014). Contextuality is about identity of random variables. Physica Scripta T163, 014009.

9. Dzhafarov, E.N., & Kujala, J.V. (2014). Probabilistic contextuality in EPR/Bohm-type systems with signaling allowed. arXiv:1406.0243
10. Dzhafarov, E.N., & Kujala, J.V. (2014). Generalizing Bell-type and Leggett-Garg-type inequalities to systems with signaling. arXiv:1407.2886.
11. Dzhafarov, E.N., & Kujala, J.V. (2015). Random variables recorded under mutually exclusive conditions: Contextuality-by-Default. Advances in Cognitive Neurodynamics IV, 405-410.
12. Dzhafarov, E.N., Kujala, J.V., & Larsson, J.-A. (2105). Contextuality in three types of quantum-mechanical systems. Foundations of Physics. DOI:10.1007/s10701-015-9882-9
13. Dzhafarov, E.N., & Kujala, J.V. (2015). Probability, random variables, and selectivity. In W.Batchelder et al. (Eds). The New Handbook of Mathematical Psychology. To be published by Cambridge University Press.
14. Kochen, S., & Specker, E. P. (1967). The problem of hidden variables in quantum mechanics. Journal of Mathematics and Mechanics, 17:59–87.
15. 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. Foundations of Physics 35, 1655-1693.
16. Khrennikov, A. (2008). Bell-Boole inequality: Nonlocality or probabilistic incompatibility of random variables? Entropy 10, 19-32.
17. Khrennikov, A. (2009). Contextual approach to quantum formalism. Springer: Berlin.
18. Kujala, J.V. , & Dzhafarov, E.N., & Larsson, J.-Å. (2015). Necessary and sufficient conditions for maximal contextuality in a broad class of quantum mechanical systems. arXiv:1412.4724.
19. Kujala, J.V., & Dzhafarov, E.N. (2015). Proof of a conjecture on contextuality in cyclic systems with binary variables. arXiv:1503.02181
20. Larsson, J.-Å. (2002). A Kochen-Specker inequality. Europhysics Letters, 58(6):799–805.
21. Simon, C., Brukner, Č., & Zeilinger, A. Hidden-variable theorems for real experiments. Physical Review Letters, 86(20):4427–4430, 2001.
22. Specker, E. P. (1960). Die Logik nicht gleichzeitig entscheidbarer Aussagen. Dialectica 14: 239–246 (English translation by M.P. Seevinck available as arXiv:1103.4537.)
23. Svozil, K. (2012). How much contextuality? Natural Computing 11, 261-265.
24. Thorisson, H. (2000). *Coupling, Stationarity, and Regeneration*. New York: Springer.
25. Winter, A. (2014). What does an experimental test of quantum contextuality prove or disprove? Journal of Physics A: Mathematical and Theoretical, 47(42):42403.