Article

Assumption-Free Derivation of the Bell-Type Criteria of Contextuality/Nonlocality

Ehtibar N. Dzhafarov

Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907, USA; ehtibar@purdue.edu

Abstract: Bell-type criteria of contextuality/nonlocality can be derived without any falsifiable assumptions, such as context-independent mapping (or local causality), free choice, or no-fine-tuning. This is achieved by deriving Bell-type criteria for inconsistently connected systems (i.e., those with disturbance/signaling), based on the generalized definition of contextuality in the contextuality-by-default approach, and then specializing these criteria to consistently connected systems.

Keywords: Bell criteria; contextuality; context-independent mapping; free choice; local causality; no-fine-tuning; nonlocality

1. Introduction

The criteria (necessary and sufficient conditions) of contextuality/nonlocality (usually, but not necessarily, in the form of inequalities) are at the heart of the foundations of quantum physics. Since John Bell's pioneering work, researchers have been interested in what assumptions about nature one needs to justify these criteria. Several such assumptions have been proposed. One of them is that measurements cannot be affected by spacelike remote events (*local causality*). Another assumption is that experimenters choose what they measure independent of the background events determining measurement outcomes (*free choice*). It has also been proposed that statistically identical measurement outcomes should have identical ontological models (*no-fine-tuning*). This article shows that the criteria of contextuality/nonlocality can be derived without any such assumptions.

Although our discussion is valid for essentially all possible systems of random variables, we will use a system describing the Bohm's version of the Einstein-Podolsky-Rosen experiment (EPR/B) [1–3] as a throughout example. This allows us to avoid technicalities needed in a more general exposition. The system in question is

$$
\begin{array}{|c|c|c|c||c|}
\hline
R_1^1 & R_2^1 & & & c = 1 \\
\hline
 & R_2^2 & R_3^2 & & c = 2 \\
\hline
 & & R_3^3 & R_4^3 & c = 3 \\
\hline
R_1^4 & & & R_4^4 & c = 4 \\
\hline\hline
q = 1 & q = 2 & q = 3 & q = 4 & \text{system } \mathcal{R}_4 \\
\hline
\end{array}
\tag{1}
$$

The random variables in (1) represent outcomes of spin measurements of two entangled particles along direction q; in every trial Alice chooses between directions $q = 1$ and 3, and Bob chooses between $q = 2$ and 4 (indicated by subscripts of the random variables). We call subscripts q in R_q^c the *contents* of the corresponding random variables. The four combinations of the Alice-Bob choices form *contexts* $c = 1, \ldots, 4$ (indicated by the superscripts of the random variables). If the spin along an axis q (i.e., content q) is measured in context c, we write:

$$
q \prec c.
\tag{2}
$$

For instance $q = 2 \prec c = 1$ but $q = 3 \not\prec c = 1$. The particles are spin-1/2, so all the random variables in the system are dichotomous, say ± 1.

The structure of this article is as follows. In Section 2, we do preliminary work: we present a rigorous version of the traditional account of the criteria for contextuality/nonlocality of systems of random variables with no disturbance (we call them consistently connected systems), and then we stipulate the traditional assumptions involved in the formulation of these criteria. The main point of this paper is presented in Section 3. There, we generalize the definition of contextuality/nonlocality to inconsistently connected systems (those with disturbance or signaling), and show that the generalized Bell-type criteria in this generalized conceptual setting can be derived with no assumptions. Instead they are derived from a classificatory definition of (non)contextuality based on the notion of the difference between random variables. The traditional Bell-type criteria then immediately follow by specializing the generalized criteria to consistently connected systems, again, with no assumptions about nature.

The term "nonlocality" has been criticized as misleading, unnecessary, or even outright contradicting the tenets of quantum physics in several recent publications [4–6]. In this paper, however, the term is used to simply designate a special case of contextuality, for systems where contexts are formed by spacelike separated components. Barring experimental biases, such a system is consistently connected, and its mathematical analysis does not differ from that of other consistently connected systems [7,8]. Singling nonlocal systems out, however, is justified due to their special importance in quantum physics.

2. Contextuality for Consistently Connected Systems

Consider a random variable Λ and a measurable function F mapping Λ and q into $\{-1, 1\}$. In contextuality/nonlocality analysis of a system we are asking whether one can choose (Λ, F) so that, for every context c,

$$(F(q, \Lambda) : q \prec c) \overset{d}{=} \left(R_q^c : q \prec c \right), \tag{3}$$

where $\overset{d}{=}$ stands for "has the same distribution as". In other words, the joint distribution of all $F(q, \Lambda)$ for a given context c is the same as the joint distribution of all R_q^c in this context. Thus, for system $\mathcal{R}_4$ in (1),

$$
\begin{aligned}
(F(q = 1, \Lambda), F(q = 2, \Lambda)) &\overset{d}{=} \left(R_1^1, R_2^1 \right), \\
(F(q = 2, \Lambda), F(q = 3, \Lambda)) &\overset{d}{=} \left(R_2^2, R_3^2 \right), \\
&etc.
\end{aligned}
\tag{4}
$$

Note that the random variables $\left(R_q^c : q \prec c \right)$ in a given context c are jointly distributed: e.g., the event $[R_1^1 = x, R_2^1 = y]$ is well-defined for any $x, y \in \{-1, 1\}$, and has a probability assigned to it. At the same time, random variables in different contexts, e.g., R_1^1 and R_1^2, or R_1^1 and R_2^2, do not have a joint distribution, as different contexts are mutually exclusive [9]. However, all random variables $F(q, \Lambda)$ are jointly distributed, because Λ is one and the same for all q and c. Therefore, the random variables S_q^c below form a (probabilisitic) *coupling* of system $\mathcal{R}_4$:

$$
\begin{array}{|c|c|c|c|}
\hline
S_1^1 = F(1, \Lambda) & S_2^1 = F(2, \Lambda) & & \\
\hline
& S_2^2 = F(2, \Lambda) & S_3^2 = F(3, \Lambda) & \\
\hline
& & S_3^3 = F(3, \Lambda) & S_4^3 = F(4, \Lambda) \\
\hline
S_1^4 = F(1, \Lambda) & & & S_4^4 = F(4, \Lambda) \\
\hline
\end{array}
\tag{5}
$$

Generally, a coupling of a system of random variables $\mathcal{R} = \left\{ R_q^c : c \in C, q \in Q, q \prec c \right\}$ is a set of jointly distributed $S = \left\{ S_q^c : c \in C, q \in Q, q \prec c \right\}$ such that, for any $c \in C$,

$$S^c = \left\{ S_q^c : q \in Q, q \prec c \right\} \overset{d}{=} \left\{ R_q^c : q \in Q, q \prec c \right\} = R^c. \tag{6}$$

Note that any set of jointly distributed random variables is a random variable. Therefore, S, S^c, and R^c are random variables, but $\mathcal{R}$ is not (hence the difference in notation).

In the coupling (5), the octuple of the S_q^c-variables is jointly distributed, and the within-context joint distributions of S_q^c variables are the same as the joint distributions of the corresponding R_q^c variables. In addition, since $F(q, \Lambda)$ does not depend on c, we have (with Pr standing for probability):

$$\Pr\left[S_q^c = S_q^{c'}\right] = 1, \tag{7}$$

for any $q \prec c, c'$. In other words, the probabilities

$$\Pr\left[S_1^1 = s_1^1, S_2^1 = s_2^1, \ldots, S_4^4 = s_4^4, S_1^4 = s_1^4\right] \tag{8}$$

are well-defined for all 2^8 octuples of $s_q^c \in \{-1, 1\}$. These probabilities are subject to the following constraints:

1. For every fixed (c, q, q'), such that $q, q' \prec c$, and any fixed values $s_q^c, s_{q'}^c \in \{-1, 1\}$,

$$\sum \iota_{\left[S_q^c = s_q^c, S_{q'}^c = s_{q'}^c\right]} \Pr\left[S_1^1 = s_1^1, \ldots, S_1^4 = s_1^4\right] = \Pr\left[R_q^c = s_q^c, R_{q'}^c = s_{q'}^c\right], \tag{9}$$

where the coefficient $\iota_{[expression]}$ is the Boolean indicator of whether the event $S_1^1 = s_1^1, \ldots, S_1^4 = s_1^4$ contains *expression*;

2. For any fixed (c, c', q) such that $q \prec c, c'$, and any fixed values $s_q^c, s_q^{c'} \in \{-1, 1\}$,

$$\sum \iota_{\left[S_q^c = s_q^c, S_q^{c'} = s_q^{c'}\right]} \Pr\left[S_1^1 = s_1^1, \ldots, S_1^4 = s_1^4\right] = \begin{cases} 0 & \text{if } s_q^c \neq s_q^{c'} \\ \Pr\left[R_q^c = s_q^c\right] = \Pr\left[R_q^{c'} = s_q^c\right] & \text{if } s_q^c = s_q^{c'} \end{cases}. \tag{10}$$

This can be compactly presented as a linear programing (LP) task [7,10]:

$$\mathbf{MX} = \mathbf{P}, \mathbf{X} \geq 0, \tag{11}$$

where $\mathbf{X}$ is the vector of the 2^8 probabilities (8), $\mathbf{P}$ is the vector of the probabilities in the right-hand sides of (9) and (10), and $\mathbf{M}$ is a Boolean incidence matrix, whose entries are the ι-coefficients in (9) and (10). The condition $\mathbf{X} \geq 0$ (componentwise) ensures that the solution $\mathbf{X}$, if it exists, consists of numbers interpretable as probabilities (the summation of these values to 1 in (11) is ensured).

Denoting by $\mathsf{LP}[\mathbf{M}, \mathbf{P}]$ a Boolean function that equals 1 if and only if a solution $\mathbf{X}$ exists, it is well-known that $\mathsf{LP}[\mathbf{M}, \mathbf{P}]$ is computable in polynomial time [11]. Consequently,

$$\mathsf{LP}[\mathbf{M}, \mathbf{P}] = 1 \tag{12}$$

can be taken as a criterion of noncontextuality/locality. As an optional step, by a facet enumeration algorithm, this criterion can be presented in the form of inequalities and equations involving moments of the distributions within contexts. For system $\mathcal{R}_4$, this yields [3,12,13]:

$$\left|\left\langle R_1^1 R_2^1 \right\rangle + \left\langle R_2^2 R_3^2 \right\rangle + \left\langle R_3^3 R_4^3 \right\rangle + \left\langle R_4^4 R_1^4 \right\rangle - 2\min\left(\left\langle R_1^1 R_2^1 \right\rangle, \left\langle R_2^2 R_3^2 \right\rangle, \left\langle R_3^3 R_4^3 \right\rangle, \left\langle R_4^4 R_1^4 \right\rangle\right)\right| \leq 2, \tag{13}$$

$$\left\langle R_1^1 \right\rangle = \left\langle R_1^4 \right\rangle, \left\langle R_2^1 \right\rangle = \left\langle R_2^2 \right\rangle, \left\langle R_3^2 \right\rangle = \left\langle R_3^3 \right\rangle, \left\langle R_4^3 \right\rangle = \left\langle R_4^4 \right\rangle. \tag{14}$$

Equalities (14) present the condition of consistent connectedness.

What ontological assumptions have we made in the foregoing? Let us note first that we need no assumptions to derive the criterion $\mathsf{LP}[\mathbf{M}, \mathbf{P}] = 1$ or its equivalents from (3). The derivation of the criterion is by straightforward linear programming, optionally complemented by facet enumeration. However, we need certain assumptions about nature to justify the plausibility of using the function $F(q, \Lambda)$, which is the starting point of the

derivation. Namely, we must have assumed that the mapping does not contain context c among its arguments. For the EPR/Bohm experiment, following Bell [14], this assumption can be called *local causality*, because dependence of the mapping on c can be interpreted as dependence of measurement outcomes on spacelike-remote settings. More generally, however, this can be called *context-independent mapping*, in order to also include the Kochen-Specker type contextuality [15], when measurements in the same context are not spacelike separated. We must have also assumed that Λ in $F(q, \Lambda)$ is one and the same for all q and for all c. This is called the assumption of *free choice*, or *statistical independence* (of measurements and settings). In relation to system $\mathcal{R}_4$, the necessity of adding the free choice assumption to the local causality assumption was pointed out to John Bell by Shimony, Horne, and Clauser in their 1985 interchange [16,17]. The relationship between free choice and local causality (more generally, context-independent mapping) is an interesting issue, but it is discussed elsewhere [18].

One can replace both the assumption of context-independent mapping and the free choice assumption with the assumption proposed by Cavalcanti, called *no-fine-tuning* [19,20]. For our purposes it can be formulated thus: if two random variables sharing a content in different contexts have the same distribution, their representations in the form

$$F(\text{some parameters}, \text{some random variables}) \tag{15}$$

should be identical. This is an attractive alternative to context-independent mapping, because the latter is not especially compelling in the cases of Kochen-Specker-type contextuality, when contexts are not defined by spacelike remote settings. Note that the no-fine-tuning can also be considered a principle of theory construction, essentially a conceptual parsimony principle, rather than an ontological assumption.

As it turns out, however, by generalizing the notion of (non)contextuality to include inconsistently connected systems, one can avoid the necessity of making any of these, or other, falsifiable assumptions. The no-fine-tuning assumption (or parsimony principle) is a consequence of specializing this general definition to consistently connected systems.

3. Contextuality in Inconsistently Connected Systems

The generalization follows the broadening of the class of systems of random variables amenable to contextuality analysis. As one can see in (14), the criterion $\mathsf{LP}[\mathbf{M}, \mathbf{P}] = 1$ can be satisfied only if the system of random variables is consistently connected: random variables measuring the same property in different contexts have the same distribution. We will now drop this constraint, and allow for inconsistent connectedness. In particular, we allow for signaling between Alice and Bob if their measurements are timelike separated.

We begin with the maximally lax representation

$$R_q^c = \gamma\left(q, c, \lambda_q^c\right), \tag{16}$$

in which both the outcome of the measurement and the background random variable are allowed to depend on both q and c. This representation obviously holds for any system of random variables. Now, because all R_q^c in a given context c are jointly distributed, it follows that all λ_q^c in this context are jointly distributed. Therefore, there is a random variable λ^c of which λ_q^c for all $q \prec c$ are functions. As a result, we get a seemingly more restrictive but still universally applicable representation:

$$R_q^c = g(q, c, \lambda^c). \tag{17}$$

The remaining dependence of λ^c on context c can also be eliminated, by the following reasoning. Let us form an arbitrary coupling

$$\Lambda \overset{d}{=} (\lambda^c : \text{all } c), \tag{18}$$

e.g., couple all λ^c independently. Then:

$$R_q^c \stackrel{d}{=} g(q, c, \mathrm{Proj}_c(\Lambda)) = G(q, c, \Lambda), \tag{19}$$

where Proj_c is the cth projection function. This representation (with Λ one and the same for all q and c) is also universally applicable. Note that in (19) we only have $R_q^c \stackrel{d}{=} G(q, c, \Lambda)$, rather than $R_q^c = G(q, c, \Lambda)$, because the latter would make R_q^c jointly distributed across mutually exclusive contexts c (which is nonsensical).

It remains to define (non)contextuality in this generalized conceptual setting. In the contextuality-by-default approach (CbD) [8–10,21,22], the system is considered noncontextual if (Λ, G) in (19) can be chosen so that the probability:

$$\Pr\left[G(q, c, \Lambda) = G(q, c', \Lambda)\right] \tag{20}$$

is the maximal possible, for all (q, c, c') such that $q \prec c, c'$. For dichotomous random variables, this means

$$\Pr\left[G(q, c, \Lambda) = G(q, c', \Lambda) = 1\right] = \min \begin{cases} \Pr\left[R_q^c = 1\right] \\ \\ \Pr\left[R_q^{c'} = 1\right] \end{cases}. \tag{21}$$

The rationale for this definition is as follows. The maximal probability with which $G(q, c, \Lambda)$ and $G(q, c', \Lambda)$ can be made to coincide shows how similar the random variables R_q^c and $R_q^{c'}$ are if taken as an isolated pair, "out of their contexts." Denoting by p_q^c the value of $\Pr[F(q, c, \Lambda) = 1] = \Pr\left[R_q^c = 1\right]$, the maximum probability of $F(q, c, \Lambda) = F(q, c', \Lambda)$ is $1 - \left|p_q^c - p_q^{c'}\right|$, and it is achieved if and only if the joint distribution of $S_q^c = F(q, c, \Lambda)$ and $S_q^{c'} = F(q, c', \Lambda)$ is (assuming $p_q^c \leq p_q^{c'}$):

		$S_q^c = 1$	$S_q^c = -1$		
$S_q^{c'} = 1$		p_q^c	0		p_q^c
$S_q^{c'} = -1$		$p_q^{c'} - p_q^c$	$1 - p_q^{c'}$		$1 - p_q^c$
		$p_q^{c'}$	$1 - p_q^{c'}$		

$$\tag{22}$$

The existence of a coupling of the system in which all pairs $\left(S_q^c, S_q^{c'}\right)$ have this joint distribution indicates that the way R_q^c and $R_q^{c'}$ are related to other random variables in the corresponding contexts leaves their dissimilarity intact. Otherwise, the contexts of the system "force" some of the pairs of R_q^c and $R_q^{c'}$ to be more dissimilar than they are when taken in isolation.

The Bell-type criterion for this generalized definition is also determined by the LP problem (11). The only difference is that the part of vector **P** determined by (10) is replaced with the probabilities given in (22):

$$\sum \iota_{\left[S_q^c = s_q^c, S_q^{c'} = s_q^{c'}\right]} \Pr\left[S_1^1 = s_1^1, \ldots, S_1^4 = s_1^4\right] = \begin{cases} p_q^c & \text{if} & s_q^c = s_q^{c'} = 1 \\ 1 - p_q^{c'} & \text{if} & s_q^c = s_q^{c'} = -1 \\ & \text{etc.} \end{cases}. \tag{23}$$

Note that this definition of (non)contextuality is purely classificatory, it does not involve any assumptions about nature.

Suppose now that the system is consistently connected. Then the maximal probability in (20) is 1, i.e., (22) holds with $p_q^c = p_q^{c'}$, and the system is noncontextual if and only if (Λ, G) can be chosen so that:

$$\Pr\left[G(q, c, \Lambda) = G(q, c', \Lambda)\right] = 1, \tag{24}$$

whenever $q \prec c, c'$. But the latter means that

$$G(q, c, \Lambda) \equiv F(q, \Lambda). \tag{25}$$

We have thus arrived at the same representation as in the previous section, but without assuming context-independent mapping, free choice, or no-fine-tuning.

4. Conclusions

We have seen that if one does not make any constraining assumptions about how a measurement outcome in a system depends on settings, local or remote, and if one adopts the CbD definition of generalized (non)contextuality, the derivation of the traditional criteria of (non)contextuality follows with no ontological assumptions, by simply specializing the definition in question to consistently connected systems.

Let us examine possible doubts about the validity of our analysis.

(1) Is not the property of consistent connectedness from which we derive (24) and (25) an assumption? If the distributions of the random variables are only known to us on a sample level, then it is indeed an assumption, as any other statistical hypothesis. However, the discussion in this article deals with the systems known to us precisely, as random variables rather than samples. For example, a standard quantum mechanical computation can yield the precise joint probabilities for the pairs of variables in each context of system $\mathcal{R}_4$.

(2) By writing $R_q^c = \gamma\left(q, c, \lambda_q^c\right)$, have we not made the assumption of "outcome determinism"? The latter means that the values of the relevant arguments uniquely determine the value of R_q^c. What if q, c, and λ_q^c only determine the distribution of the random variable R_q^c rather than its value? This possibility, however, amounts to introducing yet another random variable among the arguments of the function determining the value of R_q^c:

$$R_q^c = \gamma\left(q, c, \lambda_q^c, \mu_q^c\right). \tag{26}$$

This obviously reduces to the initial representation on renaming $\left(\lambda_q^c, \mu_q^c\right)$ into λ_q^c. This observation is, in fact, a short (and generalized) version of a theorem established as early as 1982 by Arthur Fine [12].

(3) Is not the generalized CbD definition of contextuality a form of the no-fine-tuning assumption? This is clearly not true for the original no-fine-tuning assumption [19,20], as the latter is confined to consistently connected systems. However, it has been shown by M. Jones [23] that it is possible to generalize the no-fine-tuning assumption to become a principle that forbids "hidden influences" in ontological models of a system. "Hidden influences" mean dependence of measurement outcomes on factors that do not influence the distributions of these outcomes. Maximizing the probability of $G(q, c, \Lambda) = G(q, c', \Lambda)$ ensures that the entire difference between the influences of c and c' on the measurement of q is reflected in the difference of their distributions. However, in CbD, this is a consequence of defining (non)contextuality of a system in terms of differences between content-sharing random variables, rather than an assumption about nature or a principle for constructing plausible ontological theories.

(4) Finally, let us address the question about the distinction we make between ontological assumptions and a classificatory definition. Is it defensible? Is not any assumption convertible into a definition and vice versa? Specifically, the traditional analysis of con-

textuality can be presented as a definition according to which a (consistently connected) system is noncontextual if a representation

$$\left(R_q^c : q \prec c \right) \overset{d}{=} \left(F(q, \Lambda) : q \prec c \right) \tag{27}$$

exists, and it is noncontextual otherwise. Not denying this obvious possibility, it is nevertheless reasonable to ask how such a definition is motivated. The assumptions of context-independent mapping, free choice, and no-fine-tuning serve to provide this motivation. (Let us note in passing that most of the traditional analyses of contextuality confuse the distributional equation (27) with the equality $R_q^c = F(q, \Lambda)$. The ensuing logical problems are discussed in [9,21].) In our analysis, we begin with a hidden variable model that cannot be empirically false, $R_q^c = \gamma\left(q, c, \lambda_q^c\right)$, and reduce it to $R_q^c \overset{d}{=} G(q, c, \Lambda)$, that cannot be false either. Then we pose the question about the difference between random variables R_q^c and $R_q^{c'}$, measuring the same content in different contexts. Asked about the motivation for this question, the obvious answer is that we are interested in the dependence of measurements on their contexts. The maximal probability of $S_q^c = S_q^{c'}$ in the coupling

$$\left(S_q^c = F(q, c, \Lambda), S_q^{c'} = F(q, c', \Lambda) \right) \tag{28}$$

provides the answer for R_q^c and $R_q^{c'}$ taken in isolation from their contexts. Finally, we ask the question if thus measured differences between R_q^c and $R_q^{c'}$ are compatible with their respective contexts, by finding out if the maximal probability of $S_q^c = S_q^{c'}$ remains the same or decreases within the overall couplings of the system. If asked about the motivation for this question, the answer is that the situations when these maximal probabilities decrease (for some R_q^c and $R_q^{c'}$) indicate a special form of dependence of measurements on their contexts, beyond the difference of their distributions. There seems to be no assumptions about nature that would pertain to this distinction (see point 3 above). To summarize, the point we make in this paper is not based on the admittedly blurry distinction between definitions and assumptions. Rather, it is based on dealing with a representation, $\gamma\left(q, c, \lambda_q^c\right)$ or $G(q, c, \Lambda)$, that is universally applicable, and does not therefore involve any ontological assumptions.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Bohm, D.; Aharonov, Y. Discussion of experimental proof for the paradox of Einstein, Rosen and Podolski. *Phys. Rev.* **1957**, *108*, 1070–1076.
2. Bell, J. On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.* **1966**, *38*, 447.
3. Clauser, J.F.; Horne, M.A.; Shimony, A.; Holt, R.A. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **1969**, *23*, 880.
4. Khrennikov, A. Two faced Janus of quantum nonlocality. *Entropy* **2020**, *22*, 303.
5. Kupczynski, M. Closing the door on quantum nonlocality. *Entropy* **2018**, *20*, 877.
6. Kupczynski, M. Contextuality-by-Default description of bell tests: Contextuality as the rule and not as an exception. *Entropy* **2021**, *23*, 1104.
7. Abramsky, S.; Brandenburger, A. The sheaf-theoretic structure of non-locality and contextuality. *New J. Phys.* **2011**, *13*, 113036–113075.
8. Kujala, J.V.; Dzhafarov, E.N.; Larsson, J.-A. Necessary and sufficient conditions for extended noncontextuality in a broad class of quantum mechanical systems. *Phys. Rev. Lett.* **2015**, *115*, 150401.
9. Dzhafarov, E.N.; Kujala, J.V. Probabilistic foundations of contextuality. *Fortsch. Phys. Prog. Phys.* **2017**, *65*, 1600040.

10. Dzhafarov, E.N.; Cervantes, V.H.; Kujala, J.V. Contextuality in canonical systems of random variables. *Philos. Trans. R. Soc. A* **2017**, *375*, 20160389.
11. Karmarkar, N. A new polynomial-time algorithm for linear programming. *Combin* **1984**, *4*, 373–395.
12. Fine, A. Joint distributions, quantum correlations, and commuting observables. *J. Math. Phys.* **1982**, *23*, 1306.
13. Dzhafarov, E.N.; Kujala, J.V. All-possible-couplings approach to measuring probabilistic context. *PLoS ONE* **2013**, *8*, e61712. doi:10.1371/journal.pone.0061712.
14. Bell, J.S. La nouvelle cuisine. In *Between Science and Technology*; Sarlemijn, A., Kroes, P., Eds.; Elsevier: Amsterdam, The Netherlands, 1990; pp. 97–115.
15. Kochen, S.; Specker, E.P. The problem of hidden variables in quantum mechanics. *J. Math. Mech.* **1967**, *17*, 59–87.
16. Bell, J.S.; Shimony, A.; Horne, M.A.; Clauser, J.F. An Exchange on Local Beables. *Dialectica* **1985**, *39*, 85–110.
17. Norsen, T. J. S. Bell's Concept of Local Causality. *Am. J. Phys.* **2011**, *79*, 1261.
18. Dzhafarov, E.N. Context-independent mapping and free choice are equivalent. *arXiv* **2021**, arXiv:2110.15910.
19. Cavalcanti, E.G. Classical causal models for Bell and Kochen-Specker inequality violations require fine-tuning. *Phys. Rev. X* **2018**, *8*, 021018.
20. Pearl, J.C.; Cavalcanti, E.G. Classical causal models cannot faithfully explain Bell nonlocality or Kochen-Specker contextuality in arbitrary scenarios. *arXiv* **2021**, arXiv:1909.05434.
21. Dzhafarov, E.N. On joint distributions, counterfactual values, and hidden variables in understanding contextuality. *Philos. Trans. R. Soc. A* **2019**, *377*, 20190144.
22. Dzhafarov, E.N.; Kujala, J.V.; Cervantes, V.H. Contextuality and noncontextuality measures and generalized Bell inequalities for cyclic systems. *Phys. Rev. A* **2020**, *101*, 042119; Erratum in *Phys. Rev. A* **2020**, *101*, 069902; *Phys. Rev. A* **2021**, *103*, 059901.
23. Jones, M. Relating causal and probabilistic approaches to contextuality. *Philos. Trans. R. Soc. A* **2019**, *377*, 20190133.