Is contextuality about the identity of random variables?

Mojtaba Aliakbarzadeh
School of Chemistry, Physics and Mechanical Engineering
Queensland University of Technology
Brisbane, 4000, Australia

Kirsty Kitto
Connected Intelligence Centre
University of Technology Sydney
PO Box 123, Broadway, 2007, Australia

Peter Bruza
School of Information Systems
Queensland University of Technology
Brisbane, 4000, Australia

(Dated: October 1, 2019)

Recent years have seen new general notions of contextuality emerge. Most of these employ context-independent symbols to represent random variables in different contexts. As an example, the operational theory of Spekkens [1] treats an observable being measured in two different contexts identically. Non-contextuality in this approach is the impossibility of drawing ontological distinctions between identical elements of the operational theory. However, a recent collection of work seeks to exploit context-dependent symbols of random variables to interpret contextuality [2, 3]. This approach associates contextuality with the possibility of imposing a particular joint distribution on random variables recorded under different experimental contexts. This paper compares these two different treatments of random variables and highlights the limitations of the context-dependent approach as a physical theory.

I. INTRODUCTION

Contextuality in quantum mechanics (QM) refers to the dependence of measurement results for specific observables upon the experimental arrangement being used to measure that observable [4]. Although contextuality has been part of the conceptual framework of QM for decades, recent literature has attempted to arrive at a deeper understanding of this subtle concept. For example, Abramsky and Brandenburger [5] unify the concepts of nonlocality and contextuality using sheaf theory, Cabello et al. [6] use a graph theoretical approach to model contextuality, and Acín et al. [7] construct a general contextuality model using the combinatorics of hypergraphs which generalises both the sheaf and graph theoretical approaches. Importantly to the argument mounted here, in 2005, Spekkens [1] generalized the standard treatment of contextuality in QM to arbitrary operational theories, which allows for the identification of contextuality in theory-independent frameworks.

Thus, a wide range of results are providing insights about this key physical phenomenon. However, each uses different mathematical structures and notations, and it is essential that we start to construct connections between them to improve our understanding of contextuality. Here we formally compare Spekkens’ generalized notion of contextuality with a recent competing generalized notion of contextuality called Contextuality-by-Default (CbD) [2, 3], which exploits context-dependent symbols of random variables to interpret contextuality.

In Spekkens’ approach, contextuality is defined as the non-existence of a statistically equivalent description at the ontological level for operationally equivalent procedures. However, in an actual experiment, it is not possible to attain exact operational equivalence [8]. To solve this problem, Mazurek et al. [8] suggest a general method which considers equivalences not between the procedures, but certain convex combinations of them. Interestingly, an inexact operational equivalence can also be achieved using the CbD notation. We show here that this is a result of the context-dependent symbols of random variables. However, we will show that this different realization of random variables does not provide a clear definition of ontic states. We also indicate that for a system satisfying non-signaling and no-disturbance conditions, CbD should revert back to the normal representation of random variables to satisfy the expected behaviour of that system.

In Section II we will briefly introduce the operational approach of Spekkens and the CbD approach. This is followed by comparison of these two approaches in Section III. Section IIIA demonstrates how a different definitions of probability space and random variables in the CbD approach lead to an inexact operational equivalence. Section IIIB investigates the differences between these two
approaches for concepts like parameter independence and non-signaling conditions. And finally, Section III.C evaluates the CbD approach using a cyclic example of contextuality.

II. PRELIMINARIES

A. Spekkens’ approach

In Spekkens’ approach [1], two preparation procedures are operationally equivalent \((P \equiv P')\) if:

\[
p(k|M, P) = p(k|M, P') \quad \text{for all } M.
\]

Similarly, two measurement events are operationally equivalent \(([k|M] \equiv [K'|M'])\) if:

\[
p(k|M, P) = p(k'|M', P) \quad \text{for all } P.
\]

Spekkens defines noncontextuality based on the definition of operational equivalence as:

\textbf{Definition 1} An ontological model is preparation noncontextual if we can represent every preparation procedure independent of context:

\[
P \equiv P' \Rightarrow \mu_P(\lambda) = \mu_{P'}(\lambda) \quad \forall \lambda \in \Lambda,
\]

And the model is measurement noncontextual if we can represent every measurement event independent of context:

\[
[k|M] \equiv [K'|M'] \Rightarrow \xi_\lambda(k|M) = \xi_\lambda(k|M') \quad \forall \lambda \in \Lambda.
\]

B. Contextuality-by-default

Contextuality-by-Default (CbD) [2, 3, 9] exploits contextual random variables to formalize contextuality. In this approach, random variables which we call \textit{double random variables} are represented using double indexing (e.g. \(a_{ij}^c\)), where \(q\) represents an observable (a physical property that we measure) and \(c\) indicates a context of that measurement. In this model, a system of random variables comprises \textit{stochastically unrelated} “bunches”, each of which is a set of jointly distributed random variables which have the same context. The term “stochastically unrelated” is used to indicate that there is no joint distribution for the random variables when each random variable belongs to a different bunch.

The CbD approach defines contextuality as the impossibility of imposing a joint distribution on stochastically unrelated bunches. This imposed distribution is named a “maximally connected coupling” [10] which will be explained in more detail in Section III.C.

III. COMPARING THE APPROACHES

We can start to see how the assumptions of the CbD approach subtly differ from those in the physics community with a consideration of previous work in the foundations of quantum physics. For example, Shimony [11] defines a probability distribution \(p_\lambda(a|\mathcal{A})\) (similar to \(\xi_\lambda(a|\mathcal{A})\) in Spekkens’ notation) on the set of ontic states in Bell’s experiment, pointing to the impossibility of constructing a joint probability for non-commuting observables \(\mathcal{A}_1\) and \(\mathcal{A}_2\). This is similar to what CbD defines as \textit{stochastically unrelated} for two random variables \(a_{\mathcal{A}_1, \mathcal{B}_1}\) and \(a_{\mathcal{A}_2, \mathcal{B}_1}\). However, in the CbD method, two random variables \(a_{\mathcal{A}_1, \mathcal{B}_1}\) and \(a_{\mathcal{A}_2, \mathcal{B}_2}\) are defined as \textit{stochastically unrelated} as well, a situation for which there is no counterpart in Shimony’s approach. We will now proceed to compare the CbD approach with that of Spekkens.

A. Merely close to operationally equivalent

Contextuality can emerge from non-commutativity of quantum observables, where the corresponding random variables of the non-commuting observables cannot be treated in a classical probability theory, since they cannot have a value at the same time. However, the CbD approach can be considered as a model within the framework of Kolmogorovian probability theory [9] [12]. As de Barros et al. [13] state, to define the double random variables, we need a separate probability space for each possible context. Thus, we have a random variable \(a_i^e : \Omega_j \rightarrow E_i\) (For simplicity, from now on we will use \(a_i^e\) instead of \(a_{ij}^c\), where subscripts \(i\) indicate different observables and \(j\) indicate different contexts). Here, \(E\) is a certain set of possible values and \(\Omega\) is a probability space. As an example, for the observable \(\mathcal{A}_1\) in Bell’s experiment, \(\Omega\) can be related to one of the two possible contexts \(\{\mathcal{A}_1, \mathcal{B}_1\}\) and \(\{\mathcal{A}_1, \mathcal{B}_2\}\).

This consideration of different probability spaces or different random variables for only one observable in different contexts is not allowed within the definition of measurement contextuality suggested by Spekkens. In his model, the measurement procedures which admit contextuality on the ontological level are operationally context-independent. This was explained by Simmons et al. [14]:

… the same notation is used for the objects in the first place, as a context-independent symbol is all that is needed to calculate probabilities. However there is no formal argument to be made that these elements which are operationally context-independent should also be ontologically context-independent… (p.2)

The double indexing notation associates e.g. two random variables \(a_{\mathcal{A}_1, \mathcal{B}_1}\) and \(a_{\mathcal{A}_1, \mathcal{B}_2}\) with the observable \(\mathcal{A}_1\), where each different random variable is defined based on a different probability space. Substituting
these two random variables instead of the two outcomes in the operational equivalence equation [4], we obtain: 
\[ p(a_{a_1}^{(A_1, B_1)} | (A_1, B_1)), P \] = \[ p(a_{a_1}^{(A_1, B_2)} | (A_1, B_2)), P \].
This does not completely match with the original definition of operational equivalence. Mazurek et al. [5] describe this new equation as: “merely close to operationally equivalent”.

B. Signaling conditions

CbD suggests a measure of contextuality for both the case of signaling and non-signaling [10]. In this section we will focus on constraints for signaling (the Parameter independence (PI) and non-signaling conditions [13,17]), investigating their possible representation using the CbD notation.

First, consider two stochastically unrelated random variables \( a_{A_1}^{(A_1, B_1)} \) and \( A_{B_1}^{(A_1, B_2)} \). At a superficial level, we may assume that the PI condition for each ontic state \( \lambda \) is satisfied if \( \xi_\lambda(a_{A_1}^{(A_1, B_1)} | A_1, B_1) = \xi_\lambda(a_{A_1}^{(A_1, B_2)} | A_1, B_2) \). But we cannot check the validity of this representation since the CbD approach does not have a clear position about the ontic state. Kujala et al. [2] discuss the existence of joint distribution and its relation to a hidden variable \( \lambda \):

The existence of a joint distribution of several random variables is equivalent to the possibility of presenting them as functions of a single, hidden variable \( \lambda \).

But this fails to provide a more specific definition of precisely how the contexts of the double indexed random variables relate to \( \lambda \). In contrast, in the operational approach, a joint distribution can be reproduced by averaging response functions \( \xi_\lambda(k|M) \) for different ontic states \( \lambda \), where each \( \xi_\lambda(k|M) \) is a function of \( \lambda \).

Unlike the PI condition, the non-signaling condition is expressed independently of the ontic state \( \lambda \). This provides an opportunity to precisely explore the meaning of this condition in CbD and compare it with other approaches. In Table I, we represent the joint probability distributions for Bell’s experiment using the CbD notation.

Table I shows that double indexing can provide an opportunity to precisely explore the meaning of this condition in CbD and compare it with other approaches. In Table I, we represent the joint probability distributions for Bell’s experiment using the CbD notation.

The following table shows the joint probability distributions for Bell’s experiment using the double indexing scenario.

| \( A_1, B_1 \) | \( a_{A_1}^{(A_1, B_1)} = +1 \) | \( a_{A_1}^{(A_1, B_1)} = -1 \) |
| \( a_{A_2}^{(A_1, B_1)} = +1 \) | \( p_1 \) | \( p_2 \) |
| \( a_{A_2}^{(A_1, B_1)} = -1 \) | \( p_5 \) | \( p_6 \) |
| \( A_1, B_2 \) | \( a_{A_1}^{(A_1, B_2)} = +1 \) | \( a_{A_1}^{(A_1, B_2)} = -1 \) |
| \( a_{A_2}^{(A_1, B_2)} = +1 \) | \( p_3 \) | \( p_4 \) |
| \( a_{A_2}^{(A_1, B_2)} = -1 \) | \( p_7 \) | \( p_8 \) |

\( j, j' \in \{1, ..., n\} \), this notation means \( a_i \) has the same distribution in both contexts \( j \) and \( j' \). Alternatively this relation is denoted by \( Pr[a_i = a'_i] = 1 \).

Kujala et al. [2] consider CBD as an extended notion of contextuality that allows for inconsistent connectedness (Signaling). There are other approaches that relax the non-signaling condition in the Bell experiment; for example, Brask and Chaves [18] suggest novel causal interpretations of the CHSH violation allowing communication between two sides of an experiment. Their casual structures can simulate quantum and non-signaling correlations. However, it is not possible to compare such approaches with CbD, since it not clear how to translate the double random variables and their possible relation with ontic states into these casual structures.

C. A cyclic contextuality example

Kujala et al. [2] single out a category of contextual systems with binary random variables and denote them as a cyclic class. In this class, each context (or bunch) includes exactly two observables, and each observable is measured in exactly two contexts. The number of observables and the number of contexts are equal to each other and called the rank \( n \) of the system. The cyclic system of rank 2 forms the simplest contextual scenario in the CbD approach, which is equivalent to the order effect of projective measurements in QM.

It is a common belief that we need at least three measurements to derive the simplest scenario of contextuality in QM [19,20]. This scenario is designed based on Specker’s example of contextuality [21], which requires three bivalent measurements \( \{M_1, M_2, M_3\} \) that can be measured jointly in pairs but not all at once (i.e. as a triple). In QM, this constraint on the triplewise joint measurement can be attained using three bivalent non-orthogonal measurements (POVMs), for which joint measurability does not imply commutativity [22].
Here, we focus only on the classical version of Specker’s scenario, since the CbD notation does not discuss scenarios where measurements are non-orthogonal. For this scenario, we assume that the system is consistently connected (See definition [2]). This means that the associated random variable of a measurement such as \( M_j \) must have the same distributions in two contexts \( \{ M_1, M_2 \} \) and \( \{ M_1, M_3 \} \). This can be represented by \( p_1 + p_2 = p_5 + p_6 \) for the probabilities in Table II.

TABLE II. Three bunches in the CbD representation of Specker’s scenario.

| Bunch 1 | \( a_{M_1}^{(M_1, M_2)} = 0 \) | \( a_{M_2}^{(M_1, M_2)} = 1 \) |
|---------|-----------------|-----------------|
|         | \( p_1 = 0 \)   | \( p_2 = 0.5 \) |
| Bunch 2 | \( a_{M_1}^{(M_2, M_3)} = 0 \) | \( a_{M_2}^{(M_2, M_3)} = 1 \) |
|         | \( p_9 = 0 \)   | \( p_{10} = 0.5 \) |
| Bunch 3 | \( a_{M_1}^{(M_1, M_3)} = 0 \) | \( a_{M_2}^{(M_1, M_3)} = 1 \) |
|         | \( p_5 = 0 \)   | \( p_6 = 0.5 \) |
|         | \( p_7 = 0.5 \) | \( p_8 = 0 \) |

Moreover, Specker’s scenario requires that the anti-correlation condition be satisfied. Dzhafarov et al. [23] present this condition as:

\[
Pr[a_{M_i}^{(M_i, M_j)} = -a_{M_j}^{(M_i, M_j)}] = 1, \quad i, j \in \{1, 2, 3\}.
\] (5)

Table III suggests a possible representation of Specker’s scenario using the CbD notation. Here, we assume \( a_i = a_i \) for any measurement \( i \) in two different contexts and \( j \) and \( j' \). Therefore, two random variables \( a_{M_1}^{(M_1, M_2)} \) and \( a_{M_1}^{(M_1, M_3)} \) take the same value (e.g. 1). We represent these valuations with horizontal hatching in their corresponding cells. Because of the anti-correlation condition in each pairwise joint measurement, \( a_{M_2}^{(M_2, M_3)} \) should be 0. Furthermore \( a_{M_2}^{(M_2, M_3)} \) is 0 as well, since it belongs to the same measurement \( M_2 \), which are represented by vertical hatching. Continuing this argument, we will reach a contradiction for the value of \( a_{M_3}^{(M_2, M_3)} \) which is represented by the grid hatching.

TABLE III. Specker’s scenario. The horizontal axis represents three measurements and the vertical axis indicates three contexts. The six random variables which can take values \{0, 1\} are represented by horizontal and vertical hatchings. The raised contradiction is a proof of contextuality which is represented by the grid hatching.

However, this argument seems not to be matched by the CbD approach since the equality \( a_i = a_i \) violates the double indexing assumption. Instead of this argument, Dzhafarov et al. [23] use the concept of coupling to investigate the existence of contextuality in Specker’s scenario:

**Definition 3** A coupling of a set of random variables \( a_1, ..., a_n \) is any jointly set of random variables \( b_1, ..., b_n \) such that \( a_i \sim b_i, ..., a_n \sim b_n \).

They claim that the system is contextual since there is no maximally connected coupling for the system. In their model, connection is defined as a set of random variables (such as: \( a_i, ..., a_j \)) with the same observable \( i \). And the maximality for the coupling of a system of random variables is defined as [9]:

**Definition 4** let \( a_i, ..., a_j \) be a connection of a system of random variables, an associated coupling \( b_i, ..., b_j \) is a maximal coupling if \( Pr(b_i = ... = b_j) \) has the largest value between all possible couplings. If all the couplings related to the connections of that system are maximal couplings, then the main coupling of the system is maximally connected.

We will show that this approach reduces to the above argument (in Table III), since it also requires the equality \( a_i = a_i \). This equality is concluded from the consistently connected (no-disturbance) condition.

The maximal couplings of three possible connections are constructed in Table IV. There is a restriction to construct a maximally connected coupling based on the three maximal couplings. As illustrated in Figure 1, it is not possible to have a coupling in which all probabilities are achieved together and be compatible with the probabilities in the bunches and connections. In this picture, if we associate 0 to the random variable \( S_1 \), then \( S_2 \) should be 1 because of the anti-correlation condition. Moving clockwise we reach the connection 2, in which two random variables \( S_1 \) and \( S_2 \) should take a same value and a same distribution because of the consistently connected condition. By moving further clockwise, we will reach a contradiction for the value of \( S_1 \) in the connection 1. This restriction on construction of the maximally connected coupling is the proof of the contextuality.

This proof considers an equality between two random variables of each connection: \( a_i = a_i \). This is similar to what we described earlier for the case of non-signaling, where the double indexing notation was reduced to the standard noncontextual representation of random variables. Here, if we ignore the double indexing scenario, we can remove the three imaginary connections in Figure 1 and convert the CbD notation to the standard representation of Specker scenario. This demonstrates that ChD adds extra complexity to the modelling of scenarios like Specker’s, without adding new insights to contextuality.

The other cyclic system of rank 3 in the CbD approach is associated with Leggett-Garg (LG) inequality.
CbD considers a similar structure to Specker’s scenario.

[24] CbD considers a similar structure to Specker’s scenario for LG inequality but without the anti correlation condition [9]. As a result, CbD can interpret the violation of this inequality as contextuality. However, further research is required to find the exact interpretation of micro-realism in this approach.

Although the CbD approach can relate bunches to empirical meanings, the coupling itself has no empirical meaning. Dzhafarov and Kujala [9, p. 11] declare that the coupling is merely a mathematical process: “If the bunches are assumed to have links to empirical observations, then the couplings can be said to have no empirical meaning. A coupling forms a base set of its own, consisting of itself”. This makes it impossible to generally compare the meaning of contextuality in the CbD approach with the other approaches of contextuality in physics. However, in this section we provided mathematical comparisons for a cyclic example.

### IV. CONCLUSION

Mazurek et al. [8] suggest an experimental test based on Spekkens operational approach for real situations of inexact operational equivalence. Here, we compared the Spekkens’ approach with the CbD notation which can also lead to an inexact operational equivalence. This comparison provides a new angle to study the operational approach and would help us to unify our understandings of contextuality. This comparison also helps us to evaluate the CbD approach and highlight its limitations. In that regard, we demonstrated that there is no clear relation between the double indexing notation and ontic states. We also highlighted that there is no empirical meaning for the coupling process of these random variables. We mainly explained that the identification of random variables does not add anything to the meaning of contextuality for the systems satisfying non-signaling and no-disturbance conditions (e.g., Specker scenario).

### TABLE IV. Three maximal couplings of the three connections in the CbD representation of Specker’s scenario.

| Coupling | $T_{M_1}^{(M_1, M_2)}$ | $T_{M_2}^{(M_1, M_3)}$ | $T_{M_3}^{(M_1, M_2)}$ |
|----------|-----------------|-----------------|-----------------|
| Coupling 1 | 0.5 | 0 | 0 |
| Coupling 2 | 0 | 0.5 | 0 |
| Coupling 3 | 0.5 | 0 | 0.5 |

### FIG. 1. S is a possible maximally connected coupling for the Specker scenario.

---

[1] R. W. Spekkens, Physical Review A 71, 52108 (2005).
[2] J. V. Kujala, E. N. Dzhafarov, and J.-A. Larsson, Physical Review Letters 115, 150401 (2015).
[3] E. N. Dzhafarov and J. V. Kujala, Physica Scripta T163, 014009 (2014).
[4] S. Kochen and E. P. Specker, In The Logico-Algebraic Approach to Quantum Mechanics, 293 (1967).
[5] S. Abramsky and A. Brandenburger, New Journal of Physics 13, 113036 (2011).
[6] A. Cabello, S. Severini, and A. Winter, Physical Review Letters 112, 040401 (2014).
[7] A. Acín, T. Fritz, A. Leverrier, and A. B. Sainz, Communications in Mathematical Physics 334, 533 (2015).
[8] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, Nature Communications 7, ncomms11780 (2016).
[9] E. N. Dzhafarov and J. V. Kujala, Journal of Mathematical Psychology 74, 41 (2016).
[10] E. N. Dzhafarov, J. V. Kujala, and J.-A. Larsson, in Quantum Interaction: 9th International Conference, QI 2015, Filzbach, Switzerland, July 15-17, 2015, Revised Selected Papers, edited by H. Atmanspacher, T. Fink, and E. Pothos (Springer International Publishing, 2016) pp. 12–23.
[11] A. Shimony, The British Journal for the Philosophy of Science 35, 25 (1984).
[12] E. N. Dzhafarov and M. Kon, Journal of Mathematical Psychology 85, 17 (2018).
[13] J. A. de Barros, J. V. Kujala, and G. Oas, Journal of Mathematical Psychology 74, 34 (2016).
[14] A. W. Simmons, J. J. Wallman, H. Pashayan, S. D. Bartlett, and T. Rudolph, New Journal of Physics 19, 033030 (2017).
[15] J. P. Jarrett, Nois 18, 569 (1984).
[16] A. Shimony, Foundations of quantum mechanics in the light of new technology , 225 (1984).
[17] T. Maudlin, Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics (John Wiley & Sons, 2011).
[18] J. B. Brask and R. Chaves, Journal of Physics A: Math-
[19] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, Physics Reports 506, 1 (2011).

[20] R. Kunjwal, Contextuality beyond the Kochen-Specker theorem, Ph.D. thesis, The Institute of Mathematical Sciences, Chennai (2016).

[21] B. Ernst Specker, (2011), arXiv:arXiv:1103.4537v3.

[22] R. Kunjwal and S. Ghosh, Physical Review A 89, 042118 (2014).

[23] E. N. Dzhafarov, J. V. Kujala, and J.-Å. Larsson, Foundations of Physics 45, 762 (2015).

[24] A. J. Leggett and A. Garg, Physical Review Letters 54, 857 (1985).