Epistemic Odds of Contextuality in Cyclic Systems

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

[1]Purdue University, [2]University of Turku

Abstract

Beginning with the Bell theorem, cyclic systems of dichotomous random variables have been the object of many foundational findings in quantum mechanics. Here, we ask the question: if one chooses a cyclic system "at random" (uniformly within the hyperbox of all possible systems with given marginals), what are the odds that it will be contextual? We show that the odds of contextuality rapidly tend to zero as the size of the system increases. The result is based on the Contextuality-by-Default theory, in which we do not have to assume that the systems are subject to the no-disturbance/no-signaling constraints.

1 Introduction

Cyclic systems of dichotomous random variables have played a prominent role in contextuality research. Suffice it to say that they are the object of the celebrated Bell theorem [1–4], as well as the Leggett-Garg theorem [5–8], Klyachko-Can-Binicioğlu-Shumovsky theorem [9–11], and many other results. In this paper we present a simple proof of the following proposition: the epistemic (Bayesian) probability that a randomly chosen cyclic system of dichotomous random variables is contextual tends to zero as its rank n increases. The terms in this statement are to be rigorously defined later, but the gist is as follows. Systems of random variables representing measurements or hypothetical physical events can be classified into contextual and noncontextual. If a system is of a special kind, called cyclic, it is represented by a point within an n-dimensional hyperbox $\mathbb{B}$ whose edges are determined by the individual (marginal) distributions of the random variables the system contains. We consider these distributions fixed, and different meanings of a "randomly chosen" cyclic system correspond to different distributions of points within the hyperbox $\mathbb{B}$. Here, we assume this distribution to be uniform. A part of the hyperbox $\mathbb{B}$ forms a noncontextuality polytope $\mathbb{K}$ consisting of all points representing noncontextual cyclic systems.

The *epistemic probability* of choosing a contextual system is then

$$\epsilon = 1 - \frac{\mathsf{vol}\,(\mathbb{K})}{\mathsf{vol}\,(\mathbb{B})}, \tag{1}$$

where vol stands for Euclidean hypervolume. Termed differently, ϵ is the Bayesian probability of contextuality, with uniform prior. The precise value of ϵ depends on the marginal distributions that define the hyperbox $\mathbb{B}$, but we show that $\epsilon \leq 2^{n-1}/n!$, which tends to zero as n increases. The paper draws on the recent detailed analysis of cyclic systems given in [12].

2 Definitions

Our analysis is based on the Contextuality-by-Default (CbD) theory [11,13–15]. A *cyclic system of rank* $n = 2, 3, \ldots$, is a system

$$\mathcal{R}_n = \left\{ \left\{ R_i^i, R_{i\oplus 1}^i \right\} : i = 1, \ldots, n \right\}, \tag{2}$$

where $\oplus 1$ is the cyclic shift $1 \mapsto 2, \ldots, n-1 \mapsto n, n \mapsto 1$, and $\left\{ R_i^i, R_{i\oplus 1}^i \right\}$ are pairs of jointly distributed random variables. We will assume here that all R_j^i are dichotomous, 0/1-variables (although any other labeling will be equally acceptable). The matrix below presents an example of a cyclic system:

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

This is a cyclic system of rank 4, describing, e.g., the object of the best known version of the Bell theorem [2, 3]. The columns of the matrix correspond to properties q being measured, denoted by the subscripts of the variables. Thus, in the target application of the Bell theorem, $q = 1$ and $q = 3$ represent Alice's settings, while $q = 2$ and $q = 4$ represent Bob's settings. We generically refer to $q = j$ in R_j^i as the *content* of this random variable. The rows of the matrix correspond to *contexts* in which the random variables are pairwise recorded, denoted by their superscripts. So, a random variable in a system is uniquely identified by its content and its context.

If any two content-sharing (i.e., equally subscripted, measuring the same property) random variables are identically distributed, i.e., if $\left\langle R_j^i \right\rangle = \left\langle R_j^{i'} \right\rangle$ for any i, i', j for which R_j^i and $R_j^{i'}$ exist, the system is said to be *consistently connected*. This is the CbD term for compliance with the no-disturbance/no-signaling constraint. Unlike most approaches to contextuality (an exception being [16]), CbD does not need this assumption, so the cyclic systems here are generally *inconsistently connected*.

The definition of (non)contextuality is based on the notion of a *coupling*. In system (2), any two context-sharing random variables are jointly distributed, but any two random variables belonging to different contexts are *stochastically unrelated*. The system $\mathcal{R}_n$ as a whole therefore is not jointly distributed. A coupling of $\mathcal{R}_n$ is a set of jointly distributed random variables (hence, a random variable in its own right)

$$S = \left\{S_j^i : j = i, i \oplus 1; i = 1, \ldots, n\right\} \tag{4}$$

such that

$$\left[\begin{array}{c} \left\langle S_i^i \right\rangle = \left\langle R_i^i \right\rangle = p_i^i \\ \left\langle S_{i\oplus 1}^i \right\rangle = \left\langle R_{i\oplus 1}^i \right\rangle = p_{i\oplus 1}^i \\ \left\langle S_i^i S_{i\oplus 1}^i \right\rangle = \left\langle R_i^i R_{i\oplus 1}^i \right\rangle = p_{i,i\oplus 1} \end{array}\right], i = 1, \ldots, n. \tag{5}$$

The system $\mathcal{R}_n$ is *noncontextual* if it has a coupling S in which any two content-sharing random variables coincide with maximal possible probability. It is easily seen that this means

$$\left\langle S_i^i S_i^{i\ominus 1} \right\rangle = \min\left(p_i^i, p_i^{i\ominus 1}\right) = p^{i,i\ominus 1}, i = 1, \ldots, n, \tag{6}$$

where $\ominus 1$ is the inverse of $\oplus 1$. If such a coupling does not exist, the system is *contextual*. The intuition is that the contexts in this case "force" the content-sharing variables to be more dissimilar than they can be if taken in isolation. In the particular case of consistently connected systems, $\min\left(p_i^i, p_i^{i\ominus 1}\right) = 1$, and one can say, with a slight abuse of language, that in contextual systems the contexts prevent the content-sharing random variables from being "the same." This is, essentially, the traditional understanding of contextuality [15,17]. There is a simple closed-form criterion of (non)contextuality proved in [18] (and reduced to one proved in [19] in the special case of consistently connected systems). In this paper, however, we make no use of this criterion.

3 Main Result

To define epistemic probabilities, one needs a principled way of placing a system within a space of systems. Here, we follow the scheme we used in [12] to define a noncontextuality polytope and measures of (non)contextuality. Due to the prominent role of [12] in our reasoning, we explain notation correspondences with that paper in several subsequent footnotes. We begin by introducing three probability vectors.[1] With reference to (5), denote

$$\mathbf{a} = \left(1, p_1^1, p_2^1, \ldots, p_n^n, p_1^n\right)^\intercal, \tag{7}$$

and

$$\mathbf{b} = \left(p_{12}, p_{23}, \ldots, p_{n-1,n}, p_{n1}\right)^\intercal. \tag{8}$$

With reference to (6), denote

$$\mathbf{c} = \left(p^{1n}, p^{21}, \ldots, p^{n-1,n-2}, p^{n,n-1}\right)^\intercal. \tag{9}$$

[1]The probability vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ below correspond, respectively, to $\mathbf{p_l}, \mathbf{p_b}, \mathbf{p_c}$ in [12].

A system represented by $(\mathbf{a}, \mathbf{b}, \mathbf{c})^{\intercal}$ is noncontextual if and only if there is a vector $\mathbf{h} \geq 0$ (componentwise) such that

$$\mathbf{Mh} = (\mathbf{a}, \mathbf{b}, \mathbf{c})^{\intercal}, \tag{10}$$

where $\mathbf{M}$ is an incidence (0/1) matrix whose detailed description is given in [12–14]. For any set of fixed 1-marginals (which means fixed vectors $\mathbf{a}$ and $\mathbf{c}$), we call the convex polytope

$$\mathbb{K} = \{\, \mathbf{b} |\, \exists \mathbf{h} \geq 0 : \mathbf{Mh} = (\mathbf{a}, \mathbf{b}, \mathbf{c})^{\intercal}\} \tag{11}$$

the *noncontextuality polytope*. We remark in passing that the L_1-distance between a point $\mathbf{b}$ and the surface of the polytope $\mathbb{K}$ is a natural measure of contextuality (if $\mathbf{b}$ is outside $\mathbb{K}$) or noncontextuality (if $\mathbf{b}$ is in $\mathbb{K}$) of the system represented by $\mathbf{b}$ [12, 14].

It is easily seen that

$$\max\left(0, p_i^i + p_{i\oplus 1}^i - 1\right) \leq p_{i,i\oplus 1} \leq \min\left(p_i^i, p_{i\oplus 1}^i\right), i = 1, \dots, n, \tag{12}$$

whence $\mathbb{K}$ is inscribed in the hyperbox

$$\mathbb{B} = \prod_{i=1}^{n} \left[\max\left(0, p_i^i + p_{i\oplus 1}^i - 1\right), \min\left(p_i^i, p_{i\oplus 1}^i\right)\right]. \tag{13}$$

Let us agree to call the vertices of $\mathbb{B}$ *odd* or *even* depending on whether its coordinates contain, respectively, an odd or even number of left endpoints of the intervals in (13). We need the following two results from [12].

Lemma 1 (Lemma 13 in [12])**.** *Every even vertex of $\mathbb{B}$ belongs to $\mathbb{K}$ (i.e., represents a noncontextual system).*[2]

Lemma 2 (Lemma 10 in [12])**.** *If all 1-marginal probabilities are* $1/2$*, then $\mathbb{K}$ is the n-demicube whose vertices are even vertices of $\mathbb{B}$.*[3]

We are ready now to make our main observation. We define the epistemic probability of a system falling within a Lebesque-measurable subset $\mathbb{S}$ of $\mathbb{B}$ as the ratio of their Lebesque measures (in our case, Euclidean volumes of polytopes).

Theorem 3. *For any cyclic system of rank n,*

$$\epsilon = 1 - \frac{\mathsf{vol}\mathbb{K}}{\mathsf{vol}\mathbb{B}} \leq \frac{2^{n-1}}{n!}. \tag{14}$$

This upper bound is tight.

[2] In [12], where we also consider other polytopes, the noncontextuality polytope $\mathbb{K}$ is denoted by $\mathbb{P}_{\mathbf{b}}$ or, if the values of the random variables are encoded as $-1/+1$ rather than $0/1$, by $\mathbb{E}_{\mathbf{b}}$. The hyperbox $\mathbb{B}$, with the $-1/+1$ encoding of values, is denoted by $\mathbb{R}_{\mathbf{b}}$.

[3] Note that the system in this lemma is a special case of a consistently connected system. In reference to (13), $\mathbb{B}$ in this case is hypercube $[0, 1/2]^n$. In [12], it corresponds to hypercube $[-1, 1]^n$ denoted by $\mathbb{C}_{\mathbf{b}}$.

Proof. By Lemma 1, since $\mathbb{K}$ is convex, it contains the polytope $\mathbb{D}$ formed by the even vertices of $\mathbb{B}$. This polytope is an n-demibox.[4] Within $\mathbb{B}$, it is separated from any odd vertex of $\mathbb{B}$ by the corner formed by this vertex and the endpoints of the sides of $\mathbb{B}$ emanating from this vertex. The volume of this corner is $L_1 \dots L_n/n!$, where L_i is the length of the ith side in (13). There are 2^{n-1} such corners, whence

$$\mathsf{vol}\mathbb{D} = \mathsf{vol}\mathbb{B} - 2^{n-1}\frac{L_1 \dots L_n}{n!}. \tag{*}$$

Since $\mathsf{vol}\mathbb{B} = L_1 \dots L_n$, the epistemic probability for a point randomly chosen within $\mathbb{B}$ to fall within $\mathbb{D}$ is

$$\frac{\mathsf{vol}\mathbb{D}}{\mathsf{vol}\mathbb{B}} = 1 - \frac{2^{n-1}}{n!}. \tag{**}$$

The upper bound stated in the theorem now follows from $\mathsf{vol}\mathbb{K} \geq \mathsf{vol}\mathbb{D}$. By Lemma 2, in the case of uniform 1-marginals $\mathbb{K}$ coincides with $\mathbb{D}$. This proves that the upper bound is tight. □

Although we fixed the 1-marginals in the proof, we see that the epistemic probability is bounded by an expression that only depends on n. The convergence of this upper bound to zero is quite fast:

n	2	3	4	5	10	15	20	50	...
$\epsilon \leq$	1	6.67 $\times 10^{-1}$	3.34 $\times 10^{-1}$	1.34 $\times 10^{-1}$	1.42 $\times 10^{-4}$	1.26 $\times 10^{-8}$	2.16 $\times 10^{-13}$	1.86 $\times 10^{-50}$	...

(15)

Depending on the 1-marginals, ϵ can be much smaller than $2^{n-1}/n!$.

One could also consider the epistemic probability $\widetilde{\epsilon}$ of choosing a contextual system in the $3n$-dimensional space formed by all vectors

$$\widetilde{\mathbf{p}} = \left(\left\langle R_i^i \right\rangle, \left\langle R_i^i R_{i\oplus 1}^i \right\rangle, \left\langle R_{i\oplus 1}^i \right\rangle : i = 1, \dots, n\right). \tag{16}$$

Because the upper bound for ϵ in Theorem 3 does not depend on 1-marginals, this upper bound also bounds $\widetilde{\epsilon}$, irrespective of the epistemic distribution of choices of the 1-marginals (provided the conditional probability of contextuality, given the 1-marginals, is defined as above).

4 Conclusion

Our finding is surprising, as it shows that complexity and contextuality may very well be antagonists. Whether this has deeper interpretational consequences depends on how much it can be generalized beyond the class of cyclic systems. One problem is that size of a system is not a well-defined concept outside specially defined classes of systems. For cyclic systems, their rank n determines simultaneously the number of contexts (n), the number of contents (n), and the

[4] We use this term as a straightforward generalization of n-demicube.

number of random variables ($2n$). Generally, however, the number of contents and contexts can be incremented independently, and it is easy to see that our result will not always hold. Consider, e.g., a system with two contents and increasing number n of contexts. It can be shown that the epistemic probability with which such a system is contextual generally does not decrease with increasing n (e.g., within the class of consistently connected systems this epistemic probability is 1 for all $n \geq 2$). Even if we define the size of a system as the rank of its largest cyclic subsystem, our result still will not be generalized automatically: as shown in [12], a system whose cyclic subsystems are all noncontextual may very well be contextual (although the epistemic probability of this has not been investigated). Further work is needed.

What can be said about cyclic systems of random variables that are not dichotomous? CbD requires that each random variables in an initial system be replaced with a set of jointly distributed dichotomizations thereof, and only then subjected to contextuality analysis [13]. However, a cyclic system thus dichotomized is no longer cyclic. In the case of categorical random variables with unordered sets of values, we form all possible dichotomizations, and then we have a simple necessary condition for noncontextuality, given by the *nominal dominance theorem* [13]. Using this condition, our computations show that for cyclic systems the epistemic probability of contextuality *increases* with the number of unordered values of the random variables. This, however, is not an easily interpretable result, because as the set of possible values of random variables increases in cardinality, it is progressively less feasible to treat it as completely unordered, and it becomes impossible when the cardinality is infinite. For ordered/structured sets of values the idea of all possible dichotomizations is no longer justifiable, and the nominal dominance theorem no longer applies (see the discussion in the concluding section of [13]). Further work is needed.

One implication is obvious, however: insofar as one is concerned with cyclic systems of dichotomous random variables, unless one is guided by a predictive theory, one is unlikely to stumble upon a contextual system of a sufficiently large size. Of course, quantum mechanics is such a predictive theory, which is why we know of the existence of contextual cyclic systems (although even there, most of experimental work is confined to cyclic systems of ranks not exceeding 5). However, our finding poses a serious problem for attempts to seek contextuality outside quantum mechanics, where such a predictive theory may not exist.

References

[1] J. Bell, Phys. 1, 195 (1964).

[2] J. Bell, Rev. Modern Phys. 38, 447 (1966).

[3] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. 23, 880 (1969).

[4] A. Fine, J. Math. Phys. 23, 1306 (1982).

[5] A. J. Leggett and A. Garg, Phys. Rev. Lett. 54, 857 (1985).

[6] P. Suppes and M. Zanotti, Synth. 48, 191 (1981).

[7] J. Kofler and C. Brukner, Phys. Rev. A 87, 052115 (2013).

[8] G. Bacciagaluppi, Int. J. Quant. Found. 1, 1 (2015).

[9] A. A. Klyachko, M. A. Can, S. Binicioğlu, and A. S. Shumovsky, Phys. Rev. Lett. 101, 020403 (2008).

[10] R. Lapkiewicz, P. Li, C. Schaeff, N. K. Langford, S. Ramelow, M. Wieśniak, and A. Zeilinger, Nature 474, 490 (2011).

[11] J. V. Kujala, E. N. Dzhafarov, and J.-Å. Larsson, Phys. Rev. Lett. 115, 150401 (2015).

[12] E. N. Dzhafarov, J. V. Kujala, and V. H. Cervantes, Phys. Rev. A 101:042119 (+Erratum: Phys. Rev. A 101:069902.)

[13] E. N. Dzhafarov, V. H. Cervantes, and J. V. Kujala, Phil. Trans. Roy. Soc. A 375, 20160389 (2017).

[14] J. V. Kujala and E. N. Dzhafarov, Phil. Trans. Roy. Soc. A 377, 20190149 (2019).

[15] E. N. Dzhafarov and J. V. Kujala, Fortsch. Phys. - Prog. Phys. 65, 1600040 (1-11) (2017).

[16] B. Amaral, C. Duarte, and R. I. Oliveira, J. Math. Phys. 59, 072202 (2018).

[17] E. N. Dzhafarov, Phil. Trans. Roy. Soc. A 377, 20190144 (2019).

[18] J. V. Kujala and E. N. Dzhafarov, Found. Phys. 46, 282 (2016).

[19] M. Araújo, M. T. Quintino, C. Budroni, M. T. Cunha, and A. Cabello, Phys. Rev. A 88, 022118 (2013).