Proof of a Conjecture on Contextuality in Cyclic Systems with Binary Variables

Janne V. Kujala[1] and Ehtibar N. Dzhafarov[2]

[1]University of Jyväskylä, jvk@iki.fi
[2]Purdue University, ehtibar@purdue.edu

Abstract

We present a proof for a conjecture previously formulated by Dzhafarov, Kujala, and Larsson (Found. Phys. 7, 762-782, 2015). The conjecture specifies a measure for the degree of contextuality and a criterion (necessary and sufficient condition) for contextuality in a broad class of quantum systems. This class includes Leggett-Garg, EPR/Bell, and Klyachko-Can-Binicioglu-Shumovsky type systems as special cases. In a system of this class certain physical properties $q_1, \ldots, q_n$ are measured in pairs (q_i, q_j); every property enters in precisely two such pairs; and each measurement outcome is a binary random variable. Denoting the measurement outcomes for a property q_i in the two pairs it enters by V_i and W_i, the pair of measurement outcomes for (q_i, q_j) is (V_i, W_j). Contextuality is defined as follows: one computes the minimal possible value Δ_0 for the sum of $\Pr[V_i \neq W_i]$ (over $i = 1, \ldots, n$) that is allowed by the individual distributions of V_i and W_i; one computes the minimal possible value $\Delta_{\min}$ for the sum of $\Pr[V_i \neq W_i]$ across all possible couplings of (i.e., joint distributions imposed on) the entire set of random variables $V_1, W_1, \ldots, V_n, W_n$ in the system; and the system is considered contextual if $\Delta_{\min} > \Delta_0$ (otherwise $\Delta_{\min} = \Delta_0$). This definition has its justification in the general approach dubbed Contextuality-by-Default, and it allows for measurement errors and signaling among the measured properties. The conjecture proved in this paper specifies the value of $\Delta_{\min} - \Delta_0$ in terms of the distributions of the measurement outcomes (V_i, W_j).

Keywords: CHSH inequalities; contextuality; criterion for contextuality; Klyachko-Can-Binicioglu-Shumvosky inequalities; Leggett-Garg inequalities; measurement bias; measurement errors; probabilistic couplings; signaling.

1 Introduction

According to Acín et al. [1], with only few exceptions, literature on contextuality mostly concerns particular examples and lacks general theory. Perhaps adding to the list of exceptions, two recent papers [2, 3] present a theory of contextuality that, although not entirely general, applies to a very broad class of quantum systems. Defining context as the set of physical properties that are measured conjointly, the novelty of this theory is in that it applies non-trivially also in the presence of context-dependent measurement biases (e.g., due to interactions/signaling, or imperfections in the measurement procedure). When such context-dependent measurement biases are present, the distribution of the measurement of a given physical property may vary over different contexts. A system is considered noncontextual if there exists a joint distribution of the measurements of all contexts such that the measurements of a given physical property over different contexts are (in a well-defined sense) maximally correlated.

This definition, formulated in Refs. [2, 3], applies to all systems where each measurement has a finite number of possible outcomes. Analytic and computational results, however, are confined to the subclass of so-called *cyclic systems*, where each physical property appears in two different contexts, each context consists of two different physical properties, and each measurement has two possible outcomes. This class includes Leggett-Garg, EPR/Bell, and Klyachko-Can-Binicioglu-Shumovsky type systems as special cases. As the main result of Ref. [3], a criterion (necessary and sufficient condition) was derived for a system to be contextual given the joint distributions of the measurements in each context. In Ref. [2] a measure of the degree of contextuality was defined based on how far the measurements of each physical property over different contexts are from being maximally correlated. This measure has a theoretically justified formulation (see below for details) and it can be used to define a criterion of contextuality: a system is contextual if and only if the degree of contextuality is positive. Computer-assisted calculations were used in Ref. [2] to derive an expression for the measure of contextuality for systems of 3, 4, or 5 physical properties. Based on these, a general expression was conjectured for any number $n \geq 2$ of physical properties. Interestingly, the criterion of contextuality implied by the expression conjectured in Ref. [2] was somewhat simpler and inherently different in form from the one derived analytically in Ref. [3], although both must be equivalent if the conjecture is true. In this paper, we show that the conjecture is indeed true, and thereby make the results complete for the special class of cyclic systems.

1.1 Terminology and notation

A *cyclic* (single cycle) system (see Figure 1) is defined as a system of measured properties $q_1, \ldots, q_n$ ($n \geq 2$) and measurement results (random variables) satisfying the following conditions [2, 3]:

1. the properties are measured in pairs $(q_1, q_2), \ldots, (q_{n-1}, q_n), (q_n, q_1)$, called *contexts*, so that each property enters in precisely two contexts;

2. the result of measuring $(q_i, q_{i\oplus 1})$, $i = 1, \ldots, n$, is a pair of jointly distributed ± 1 random variables $(V_i, W_{i\oplus 1})$, called a *bunch* (with $\oplus$ denoting circular addition: $i \oplus 1 = i + 1$ for $i < n$, and $n \oplus 1 = 1$).

The Leggett-Garg system [4, 5], the EPR/Bell system [6–10], and Klyachko-Can-Binicioglu-Shumovsky system [11] are cyclic systems with n equal, respectively, 3, 4, and 5. See [2, 3] for details.

The distribution of every bunch $(V_i, W_{i\oplus 1})$ is uniquely determined by the expectations $\langle V_i \rangle$, $\langle W_{i\oplus 1} \rangle$, and $\langle V_i W_{i\oplus 1} \rangle$. A pair $\{V_i, W_i\}$ of random variables, representing the same property q_i in different contexts, is called a *connection*: V_i, W_i are not jointly distributed, so the expectation $\langle V_i W_i \rangle$ is undefined.

1.2 Contextuality

A (probabilistic) coupling for random variables $X, Y, Z, \ldots$ is defined as any random variable $(X^*, Y^*, Z^*, \ldots)$ such that $X^*, Y^*, Z^*, \ldots$ are distributed as, respectively $X, Y, Z, \ldots$. By defi-

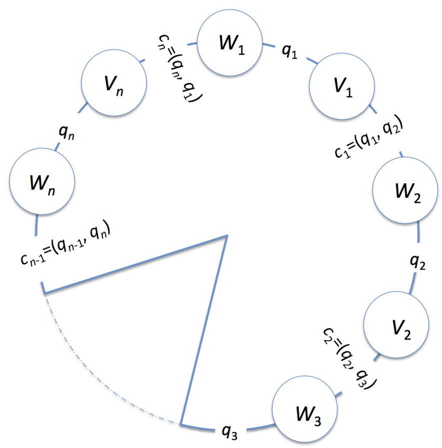

Figure 1: A schematic representation of a cyclic system. For $i = 1, \ldots, n$, the measurement of q_i is denoted V_i if q_i is measured together with $q_{i\oplus 1}$; and the measurement of q_i is denoted W_i if q_i is measured together with $q_{i\ominus 1}$. As a result, the observed pairs of random variables ("bunches") are (V_1, W_2), (V_2, W_3), ..., (V_n, W_1).

nition of a random variable (in the broad sense of the term, including vectors and processes), the components of $(X^*, Y^*, Z^*, \ldots)$ are jointly distributed. For simplicity, we omit asterisks and speak of a coupling $(X, Y, Z, \ldots)$ for $X, Y, Z, \ldots$ (imposing thereby, non-uniquely, a joint distribution on $X, Y, Z, \ldots$ that otherwise may not have one).

In relation to cyclic systems, we are interested in two types of couplings: (1) couplings (V_i, W_i) for the connections $\{V_i, W_i\}$ $(i = 1, \ldots, n)$; and (2) couplings $((V_1, W_2), (V_2, W_3), \ldots, (V_n, W_1))$ for the set of the observed bunches $(V_1, W_2), (V_2, W_3), \ldots, (V_n, W_1)$. The latter coupling can be written as a random variable $(V_1, W_2, V_2, W_3, \ldots, V_n, W_1)$, with the proviso that its 2-marginals $(V_i, W_{i\oplus 1})$ are distributed as the corresponding bunches. Put differently, it is the coupling for the *entire system of random variables* that agrees with the observed bunches.

For every $i = 1, \ldots, n$, among all couplings for the connection $\{V_i, W_i\}$, we consider one in which $\langle V_i W_i \rangle = 1 - |\langle V_i \rangle - \langle W_i \rangle|$. This coupling is called *maximal*, because $1 - |\langle V_i \rangle - \langle W_i \rangle|$ is the maximum possible value for $\langle V_i W_i \rangle$ (with given $\langle V_i \rangle$ and $\langle W_i \rangle$). Equivalently, in the maximal coupling for $\{V_i, W_i\}$ the probability $\Pr[V_i \neq W_i]$ attains its minimum possible value: $\frac{1}{2} |\langle V_i \rangle - \langle W_i \rangle|$. It is clear that all connections have maximal couplings if and only if

$$\Delta = \sum_{i=1}^{n} \Pr[V_i \neq W_i] \tag{1}$$

is at its minimal possible value. We denote this value

$$\Delta_0 = \frac{1}{2} \sum_{i=1}^{n} |\langle V_i \rangle - \langle W_i \rangle| \,. \tag{2}$$

Clearly, in every coupling $(V_1, W_2, V_2, W_3, \ldots, V_n, W_1)$ for the entire system the value of Δ is uniquely defined and cannot fall below Δ_0. This leads to the following definition of a *measure* *(degree) of contextuality*: it is

$$\mathsf{CNTX} = \Delta_{\min} - \Delta_0 \geq 0, \tag{3}$$

where $\Delta_{\min}$ is the smallest possible value of Δ across all couplings for the entire system. The system is *contextual* if $\mathsf{CNTX} > 0$, and it is not if $\mathsf{CNTX} = 0$. In Ref. [3], if a system is not contextual, it is said to have a *maximally noncontextual description*.

1.3 Conjecture

In Ref. [2], it was conjectured that

$$\Delta_{\min} = \frac{1}{2} \max \begin{cases} \mathsf{s}_1 \left(\langle V_i W_{i\oplus 1} \rangle : i = 1, \ldots, n \right) - (n-2), \\ \sum_{i=1}^{n} \left| \langle V_i \rangle - \langle W_i \rangle \right|, \end{cases} \tag{4}$$

where the function $\mathsf{s}_1 (x_1, \ldots, x_k)$ for any $k > 1$ real-valued arguments is defined as

$$\mathsf{s}_1 (x_1, \ldots, x_k) = \max \sum_{i=1}^{k} m_i x_i, \tag{5}$$

with the maximum taken over all $m_1 \ldots, m_k \in \{-1, 1\}$ such that $\prod_{i=1}^{k} m_i = -1$. It follows then that the measure of contextuality is

$$\mathsf{CNTX} = \frac{1}{2} \max \begin{cases} \mathsf{s}_1 \left(\langle V_i W_{i\oplus 1} \rangle : i = 1, \ldots, n \right) - \sum_{i=1}^{n} \left| \langle V_i \rangle - \langle W_i \rangle \right| - (n-2), \\ 0, \end{cases} \tag{6}$$

and a system is contextual if and only if

$$\mathsf{s}_1 \left(\langle V_i W_{i\oplus 1} \rangle : i = 1, \ldots, n \right) > \sum_{i=1}^{n} \left| \langle V_i \rangle + \langle W_i \rangle \right| + (n-2). \tag{7}$$

In Ref. [2] this was shown to be true for $n = 3, 4, 5$, but not generally.

In Ref. [3], a different criterion for contextuality in cyclic systems was derived: a system is contextual if and only if

$$\mathsf{s}_1 \left(\langle V_i W_{i\oplus 1} \rangle, 1 - \left| \langle V_i \rangle - \langle W_i \rangle \right| : i = 1, \ldots, n \right) > 2n - 2. \tag{8}$$

Of the two criteria, the conjectured (7) and the proved (8), the former is more specific, as it is easy to see that (8) follows from it (which means that it is known to be a necessary condition for contextuality). However, this is not the main reason why one should be interested in (7). The main reason is that (7) follows from a conjectured formula for the fundamental theoretical quantity $\Delta_{\min}$ whose excess (3) over the minimum possible value Δ_0 is used as a measure of contextuality, whereas the degree of violation of (8) does not have any theoretically motivated interpretation as a measure of contextuality.

1.4 What we do in this paper

It is easy to see that (8) and (7) are equivalent for *consistently-connected* systems, i.e., those with $\Delta_0 = 0$. Due to the results obtained in Ref. [2], the two criteria should also be equivalent for $n = 3, 4, 5$. However, it is easy to show by examples that the two inequalities, (7) and (8), are not algebraic variants of each other: in particular, the expressions

$$\mathsf{s}_1 \left(\langle V_i W_{i\oplus 1}\rangle : i = 1, \ldots, n\right) - \sum_{i=1}^{n} |\langle V_i\rangle - \langle W_i\rangle| - (n-2)$$

and

$$\mathsf{s}_1 \left(\langle V_i W_{i\oplus 1}\rangle, 1 - |\langle V_i\rangle - \langle W_i\rangle| : i = 1, \ldots, n\right)$$

are not equal to each other.

Nevertheless, the two inequalities are equivalent, as we prove in Section 3 (Theorem 13). That is, (8) is indeed a criterion as conjectured in Ref. [2]. In Section 4 we prove the general formula (6) for the contextuality measure (Theorem 14). Note that the criterion (7) is merely a consequence of the contextuality measure formula (6), i.e., the logical derivability diagram is

$$\begin{array}{c}
\text{Definition of Contextuality (3)} \\
\downarrow \\
\text{Contextuality Measure Formula (6)} \\
\downarrow \\
\text{Criterion of Contextuality (7)}
\end{array}$$

However, in this paper we arrive at the formulas for the measure and criterion in a more circuitous way. We first prove that the criterion (7) is equivalent to the previously derived criterion (8), and then we use (7) and the definition of contextuality to derive the measure formula (6):

$$\begin{tikzcd}
\text{Definition of Contextuality (3)} \arrow[r, "\text{Ref. [2]}"] \arrow[d, "\text{Theorem 14}"'] & \text{Criterion of Contextuality (8)} \arrow[d, leftrightarrow, "\text{Theorem 13}"] \\
\text{Contextuality Measure Formula (6)} & \text{Criterion of Contextuality (7)} \arrow[l, "\text{Theorem 14}"]
\end{tikzcd}$$

2 Results we need for the proofs

Here we list some results from Ref. [3]. We make use of the function s_1 defined in (5) and the function

$$\mathsf{s}_0 \left(x_1, \ldots, x_k\right) = \max \sum_{i=1}^{k} m_i x_i, \tag{9}$$

with $k > 1$ real-valued arguments, where the maximum is taken over all $m_1 \ldots, m_k \in \{-1, 1\}$ such that $\prod_{i=1}^{k} m_i = 1$.

Lemma 1. *For any* $a_1, \ldots, a_n, b_1, \ldots, b_m \in \mathbb{R}$,

$$\begin{aligned} \mathsf{s}_1(a_1, \ldots, a_n, b_1, \ldots, b_m) = \max\{ \ & \mathsf{s}_0(a_1, \ldots, a_n) + \mathsf{s}_1(b_1, \ldots, b_m), \\ & \mathsf{s}_1(a_1, \ldots, a_n) + \mathsf{s}_0(b_1, \ldots, b_m) \ \} \end{aligned}$$

and

$$\begin{aligned} \mathsf{s}_0(a_1, \ldots, a_n, b_1, \ldots, b_m) = \max\{ \ & \mathsf{s}_0(a_1, \ldots, a_n) + \mathsf{s}_0(b_1, \ldots, b_m), \\ & \mathsf{s}_1(a_1, \ldots, a_n) + \mathsf{s}_1(b_1, \ldots, b_m) \ \}. \end{aligned}$$

Lemma 2. *Jointly distributed* ± 1*-valued random variables* A *and* B *with given expectations* $\langle A \rangle, \langle B \rangle, \langle AB \rangle$ *exist if and only if*

$$\begin{cases} -1 \leq \langle A \rangle \leq 1, \\ -1 \leq \langle B \rangle \leq 1, \\ |\langle A \rangle + \langle B \rangle| - 1 \leq \langle AB \rangle \leq 1 - |\langle A \rangle - \langle B \rangle| . \end{cases}$$

Lemma 3. *Given jointly distributed random variables* (A, B) *and jointly distributed random variables* (B', C)*, there exists a coupling* (A, B, B', C) *such that* $B = B'$ *if and only if* B *and* B' *have the same distribution.*

Lemma 4. *Jointly distributed* ± 1 *random variables* $A_1, \ldots, A_n$ $(n \geq 2)$ *with given expectations* $\langle A_1 \rangle, \ldots, \langle A_n \rangle, \langle A_1 A_2 \rangle, \ldots, \langle A_{n-1} A_n \rangle$ *exist if and only if* $A = A_i$ *and* $B = A_{i+1}$ *satisfy the condition of Lemma 2 for all* $i = 1, \ldots, n-1$.

Theorem 5. *Jointly distributed* ± 1*-valued random variables* $A_1, \ldots, A_n$ $(n \geq 3)$ *with given expectations*

$$\langle A_1 \rangle, \ldots, \langle A_n \rangle, \quad \langle A_1 A_2 \rangle, \ldots, \langle A_{n-1} A_n \rangle, \langle A_n A_1 \rangle$$

exist if and only if $A = A_i$ *and* $B = A_{i \oplus 1}$ *satisfy the condition of Lemma 2 for all* $i = 1, \ldots, n$ *and*

$$\mathsf{s}_1 \left(\langle A_1 A_2 \rangle, \ldots, \langle A_{n-1} A_n \rangle, \langle A_n A_1 \rangle \right) \leq n - 2.$$

The main result of Ref. [3] is the following theorem.

Theorem 6. *For each* $i = 1, \ldots, n$ $(n \geq 2)$*, let the distribution of a pair* $(V_i, W_{i \oplus 1})$ *of* ± 1*-valued random variables be given. Then, the following two statements are equivalent:*

1. *there exists a joint distribution (coupling) of the pairs such that for all* $i = 1, \ldots, n$ *the probability*

   $$\Pr[V_i \neq W_i]$$

   in the joint is the minimum possible allowed by the marginal distributions of V_i *and* W_i *(i.e., there exists a maximally noncontextual description),*

2. *the main criterion*

   $$\mathsf{s}_1 \left(\langle V_i W_{i \oplus 1} \rangle, 1 - |\langle V_i \rangle - \langle W_i \rangle| : i = 1, \ldots, n \right) \leq 2n - 2$$

   holds true.

3 The equivalence result

In Ref. [2], the inequality

$$\mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle : i = 1,\dots,n\right) \le n - 2 + \sum_{i=1}^{n} \left|\langle V_i\rangle - \langle W_i\rangle\right|$$

was derived by computer-assisted calculations as a criterion for the existence of a maximally non-contextual description for $n = 3$, $n = 4$, and $n = 5$, and it was conjectured that the same pattern would hold for all $n \ge 2$. In this section, we prove (in Theorem 13 below) this conjecture by showing that the inequality shown above is equivalent to the main criterion of Theorem 6.

Lemma 7. *For any numbers $a_1,\dots,a_n \in \mathbb{R}$ ($n \ge 2$), exactly one of the following conditions hold:*

1. *for some index k, the inequality $a_i \ge |a_k|$ holds for all $i \ne k$.*
2. *for some distinct indices j and k, the inequality $a_j + a_k < 0$ holds.*

Lemma 8. *Suppose that for some index k the inequalities $a_i \ge a_k$ and $a_i \ge 0$ hold for all $i \ne k$ (this is implied in particular by condition 1 of Lemma 7). Then,*

$$\mathsf{s}_1\left(a_1,\dots,a_n\right) = \sum_{i\ne k} a_i - a_k.$$

If condition 1 of Lemma 7 holds, then

$$\mathsf{s}_0\left(a_1,\dots,a_n\right) = \sum_{i} a_i.$$

Lemma 9. *For each $i = 1,\dots,n$ ($n \ge 2$), let $(V_i, W_{i\oplus 1})$ be a pair of jointly distributed random variables. For any combination of signs $m_1,\dots,m_n \in \{-1,+1\}$, there exist another set of random variables*

$$\begin{aligned} \hat{V}_i :&= m_i V_i,\\ \hat{W}_i :&= m_i W_i \end{aligned}$$

(with each pair $\left(\hat{V}_i, \hat{W}_{i\oplus 1}\right)$ jointly distributed) preserving the values of s_0 and s_1 of the product expectations:

$$\begin{aligned} \mathsf{s}_0\left(\langle V_i W_{i\oplus 1}\rangle : i = 1,\dots,n\right) &= \mathsf{s}_0\left(\langle \hat{V}_i \hat{W}_{i\oplus 1}\rangle : i = 1,\dots,n\right),\\ \mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle : i = 1,\dots,n\right) &= \mathsf{s}_1\left(\langle \hat{V}_i \hat{W}_{i\oplus 1}\rangle : i = 1,\dots,n\right), \end{aligned}$$

and preserving the upper and lower bounds

$$\begin{aligned} 1 - \left|\langle \hat{V}_i\rangle - \langle \hat{W}_i\rangle\right| &= 1 - \left|\langle V_i\rangle - \langle W_i\rangle\right|,\\ \left|\langle \hat{V}_i\rangle + \langle \hat{W}_i\rangle\right| - 1 &= \left|\langle V_i\rangle + \langle W_i\rangle\right| - 1 \end{aligned}$$

of the hypothetical connection expectation $\langle V_i W_i\rangle$ for all $i = 1,\dots,n$. This implies in particular that the conditions (12) and (13) as well as the measure (26) to be defined later are all insensitive to negations of the physical properties.

Lemma 10. *One can always find such a configuration of signs $m_1, \dots, m_n \in \{-1, +1\}$ in Lemma 9 that $\langle \hat{V}_i \hat{W}_{i\oplus 1}\rangle$, $i = 1, \dots, n$, satisfy condition 1 of Lemma 7 and since the conditions (12) and (13) as well as the measure (26) are all symmetrical w.r.t. permutations of the indices, one can generally assume that condition 1 of Lemma 7 is satisfied with $k = n$.*

Proof. Let k be the index that minimizes $|\langle V_i W_{i\oplus 1}\rangle|$ and let $m_{k\oplus 1} := 1$. Then, define recursively

$$m_{i\oplus 1} := \begin{cases} +1, & m_i \langle V_i W_{i\oplus 1}\rangle \geq 0, \\ -1, & \text{otherwise}, \end{cases}$$

for all $i = k \oplus 1, k \oplus 2, \dots, k \oplus (n-1)$. This implies $\langle \hat{V}_i \hat{W}_{i\oplus 1}\rangle = m_i m_{i\oplus 1} \langle V_i W_{i\oplus 1}\rangle \geq 0$ for all $i \neq k$ and so $\langle \hat{V}_i \hat{W}_{i\oplus 1}\rangle = |\langle V_i W_{i\oplus 1}\rangle| \geq |\langle V_k W_{k\oplus k}\rangle|$ for all $i \neq k$ and condition 1 of Lemma 7 is satisfied. □

Lemma 11. *For any $a, b, c, d \in \mathbb{R}$, we have*

$$\begin{aligned} -|d+c| + |a-c| + |d-b| - |a-b| &\leq 2\max\{|b|, |d|\}, \\ -|a-b| - |d-c| + ||a-c| - |d-b|| &\leq 0. \end{aligned}$$

Proof. Using the triangle inequality $|a-c| \leq |a-b| + |b-(-d)| + |(-d)-c|$, we obtain

$$\begin{aligned} & -|d+c| + |a-c| + |d-b| - |a-b| \\ & \leq -|d+c| + (|a-b| + |b-(-d)| + |(-d)-c|) + |d-b| - |a-b| \\ & = |b+d| + |d-b| = 2\max\{|b|, |d|\}, \end{aligned}$$

which is the first inequality. The latter inequality follows as the conjunction of the following two triangle inequalities

$$\begin{aligned} |a-c| &\leq |a-b| + |d-c| + |d-b|, \\ |d-b| &\leq |a-b| + |d-c| + |a-c|. \end{aligned}$$

□

Lemma 12. *For all $i = 1, \dots, n$ ($n \geq 2$), let (V_i, W_i) be a pair of jointly distributed ± 1-valued random variables satisfying $\langle V_i W_i\rangle = 1 - |\langle V_i\rangle - \langle W_i\rangle|$ and let $\rho_1, \dots, \rho_n$ satisfy*

$$|\langle V_i\rangle + \langle W_{i\oplus 1}\rangle| - 1 \leq \rho_i \leq 1 - |\langle V_i\rangle - \langle W_{i\oplus 1}\rangle| \tag{10}$$

for $i = 1, \dots, n$. If there exist distinct indices j and k such that

$$\big(1 - |\langle V_j\rangle - \langle W_j\rangle|\big) + \big(1 - |\langle V_k\rangle - \langle W_k\rangle|\big) \leq 0, \tag{11}$$

then, there exists a joint distribution of the pairs (V_i, W_i), $i = 1, \dots, n$, satisfying $\langle V_i W_{i\oplus 1}\rangle = \rho_i$ for $i = 1, \dots, n$.

Proof. We prove the statement by induction. For $n = 2$, denoting $\langle V_i W_{i\oplus 1}\rangle = \rho_i$ and $\langle V_i W_i\rangle = 1 - |\langle V_i\rangle - \langle W_i\rangle|$ for $i = 1, 2$, Theorem 5 implies that a joint exists if

$$\mathsf{s}_1 (\langle V_1 W_2\rangle, \langle V_2 W_1\rangle, \langle V_1 W_1\rangle, \langle V_2 W_2\rangle) \leq 2,$$

where, without loss of generality (by Lemmas 9 and 10), we assume $\langle V_1W_2\rangle \geq |\langle V_2W_1\rangle|$, and condition (11) yields $\langle V_1W_1\rangle + \langle V_2W_2\rangle \leq 0$. Thus, we have by Lemma 1

$$
\begin{aligned}
&\mathsf{s}_1\left(\langle V_1W_2\rangle, \langle V_2W_1\rangle, \langle V_1W_1\rangle, \langle V_2W_2\rangle\right)\\
&= \max\begin{cases}\mathsf{s}_1\left(\langle V_1W_2\rangle, \langle V_2W_1\rangle\right) + \mathsf{s}_0\left(\langle V_1W_1\rangle, \langle V_2W_2\rangle\right)\\ \mathsf{s}_0\left(\langle V_1W_2\rangle, \langle V_2W_1\rangle\right) + \mathsf{s}_1\left(\langle V_1W_1\rangle, \langle V_2W_2\rangle\right)\end{cases}\\
&= \max\begin{cases}\langle V_1W_2\rangle - \langle V_2W_1\rangle - \langle V_1W_1\rangle - \langle V_2W_2\rangle\\ \langle V_1W_2\rangle + \langle V_2W_1\rangle + |\langle V_1W_1\rangle - \langle V_2W_2\rangle|\end{cases}\\
&\leq \max\begin{cases}(1-|\langle V_1\rangle - \langle W_2\rangle|) - (|\langle V_2\rangle + \langle W_1\rangle| - 1)\\ \qquad - (1-|\langle V_1\rangle - \langle W_1\rangle|) - (1-|\langle V_2\rangle - \langle W_2\rangle|)\\ (1-|\langle V_1\rangle - \langle W_2\rangle|) + (1-|\langle V_2\rangle - \langle W_1\rangle|)\\ \qquad + |(1-|\langle V_1\rangle - \langle W_1\rangle|) - (1-|\langle V_2\rangle - \langle W_2\rangle|)|\end{cases}\\
&= \max\begin{cases}-|\langle V_2\rangle + \langle W_1\rangle| + |\langle V_1\rangle - \langle W_1\rangle| + |\langle V_2\rangle - \langle W_2\rangle| - |\langle V_1\rangle - \langle W_2\rangle|\\ 2 - |\langle V_1\rangle - \langle W_2\rangle| - |\langle V_2\rangle - \langle W_1\rangle| + ||\langle V_1\rangle - \langle W_1\rangle| - |\langle V_2\rangle - \langle W_2\rangle||\end{cases}\\
&\leq \max\begin{cases}2\max\left\{|\langle W_2\rangle|, |\langle V_2\rangle|\right\}\\ 2\end{cases} \leq 2,
\end{aligned}
$$

where the two inequalities follow from respectively Lemma 2 and Lemma 11. Thus, the statement holds for $n = 2$.

Assuming then that the statement holds for all systems smaller than n ($n \geq 3$), we prove it for a system of size n. With no loss of generality, assume $j = 1 < k < n$. By Lemma 4, there exists a joint for the chain $(V_k, W_{k+1}, \ldots, V_n, W_1)$ satisfying $\langle V_iW_{i\oplus 1}\rangle = \rho_i$ for $i = k, \ldots, n$ and this joint has a certain marginal of (V_k, W_1) satisfying by Lemma 2 the range

$$|\langle V_k\rangle + \langle W_1\rangle| - 1 \leq \langle V_kW_1\rangle \leq 1 - |\langle V_k\rangle - \langle W_1\rangle|.$$

By the induction assumption, there exists a joint of $(V_1, W_2, \ldots, V_k, W_1)$ satisfying $\langle V_iW_{i\oplus 1}\rangle = \rho_i$ for $i = 1, \ldots, k-1$ and matching the marginal of (V_k, W_1) mentioned above. By Lemma 3, it follows that there exists a joint of all (V_i, W_i) satisfying $\langle V_iW_{i\oplus 1}\rangle = \rho_i$ for $i = 1, \ldots, n$. □

Theorem 13. *For each $i = 1, \ldots, n$ ($n \geq 2$), let the distribution of a pair $(V_i, W_{i\oplus 1})$ of ± 1-valued random variables be given. Then, the main criterion*

$$\mathsf{s}_1\left(\langle V_iW_{i\oplus 1}\rangle, 1 - |\langle V_i\rangle - \langle W_i\rangle| : i = 1, \ldots, n\right) \leq 2n - 2 \tag{12}$$

is equivalent to

$$\mathsf{s}_1\left(\langle V_iW_{i\oplus 1}\rangle : i = 1, \ldots, n\right) \leq n - 2 + \sum_{i=1}^{n} |\langle V_i\rangle - \langle W_i\rangle|. \tag{13}$$

Proof. First, without loss of generality (by Lemmas 9 and 10), we assume $|\langle V_nW_1\rangle| \leq \langle V_iW_{i\oplus 1}\rangle$,

$i = 1, \ldots, n-1$. By Lemma 8, this condition implies

$$\mathsf{s}_0 \left(\langle V_i W_{i\oplus 1} \rangle : i = 1, \ldots, n \right) = \sum_{i=1}^{n} \langle V_i W_{i\oplus 1} \rangle \,, \tag{14}$$

$$\mathsf{s}_1 \left(\langle V_i W_{i\oplus 1} \rangle : i = 1, \ldots, n \right) = \sum_{i=1}^{n-1} \langle V_i W_{i\oplus 1} \rangle - \langle V_n W_1 \rangle \,. \tag{15}$$

By Lemma 1, the criterion (12) is equivalent to the conjunction of

$$\mathsf{s}_1 \left(\langle V_i W_{i\oplus 1} \rangle : i = 1, \ldots, n \right) + \mathsf{s}_0 \left(1 - |\langle V_i \rangle - \langle W_i \rangle| : i = 1, \ldots, n \right) \le 2n - 2, \tag{16}$$

$$\mathsf{s}_0 \left(\langle V_i W_{i\oplus 1} \rangle : i = 1, \ldots, n \right) + \mathsf{s}_1 \left(1 - |\langle V_i \rangle - \langle W_i \rangle| : i = 1, \ldots, n \right) \le 2n - 2. \tag{17}$$

Case 1. Suppose now that that the terms $1 - |\langle V_i \rangle - \langle W_i \rangle|$ satisfy for some k the condition $1 - |\langle V_i \rangle - \langle W_i \rangle| \ge |1 - |\langle V_k \rangle - \langle W_k \rangle||$ for all $i \ne k$. Then, Lemma 8 yields

$$\mathsf{s}_0 \left(1 - |\langle V_i \rangle - \langle W_i \rangle| : i = 1, \ldots, n \right) = n - \sum_{i=1}^{n} |\langle V_i \rangle - \langle W_i \rangle| \,, \tag{18}$$

$$\mathsf{s}_1 \left(1 - |\langle V_i \rangle - \langle W_i \rangle| : i = 1, \ldots, n \right) = n - 2 - \sum_{i \ne k} |\langle V_i \rangle - \langle W_i \rangle| + |\langle V_k \rangle - \langle W_k \rangle| \,. \tag{19}$$

Now (18) implies that (16) is equivalent to (13). Furthermore, using (19) and (14), the left side of (17) becomes

$$\begin{aligned} &\sum_{i=1}^{n} \langle V_i W_{i\oplus 1} \rangle + n - 2 - \sum_{i \ne k} |\langle V_i \rangle - \langle W_i \rangle| + |\langle V_k \rangle - \langle W_k \rangle| \\ &\le 2n - 2 - \sum_{i=1}^{n} |\langle V_i \rangle - \langle W_{i\oplus 1} \rangle| - \sum_{i \ne k} |\langle V_i \rangle - \langle W_i \rangle| + |\langle V_k \rangle - \langle W_k \rangle| \end{aligned}$$

which, by the triangle inequality

$$|\langle V_k \rangle - \langle W_k \rangle| \le \sum_{i=1}^{n} |\langle V_i \rangle - \langle W_{i\oplus 1} \rangle| + \sum_{i \ne k} |\langle V_i \rangle - \langle W_i \rangle| \tag{20}$$

implies (17) (i.e., (17) always holds in Case 1). It follows that the two conditions for maximal noncontextuality are equivalent under the assumption of this case.

Case 2. Suppose then that the assumption of Case 1 does not hold. Then, by Lemma 7, the condition (11) of Lemma 12 holds and so Lemma 12 implies that a joint exists, which, by Theorem 5 implies that (12) holds. However, the condition (11) also yields $\left|\langle V_j \rangle - \langle W_j \rangle\right| + |\langle V_k \rangle - \langle W_k \rangle| \ge 2$ which implies that that the right side of (13) is at least n whereas the left side cannot exceed n and so (13) holds true as well. Thus, the two conditions for the existence of a maximally noncontextual description are equivalent under this case as well. $\square$

4 Proof of the conjecture on the measure of contextuality

In this section, we prove the following theorem, which was shown to be correct for the special cases of $n = 3$, $n = 4$, and $n = 5$ and conjectured to hold for all $n \ge 2$ in Ref. [2]:

Theorem 14. *For each $i = 1, \ldots, n$ ($n \geq 2$), let the distribution of a pair $(V_i, W_{i\oplus 1})$ of ± 1-valued random variables be given. Then, the minimum possible value of*

$$\Delta = \sum_{i=1}^{n} \Pr\left[V_i \neq W_i\right]$$

over all possible joints of the given pairs is

$$\Delta_{\min} = \frac{1}{2} \max \begin{cases} \mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle : i = 1, \ldots, n\right) - (n-2), \\ \sum_{i=1}^{n} \left|\langle V_i\rangle - \langle W_i\rangle\right|. \end{cases}$$

Remark 15. Here the bottom expression for $\Delta_{\min}$ corresponds to the case that a maximally noncontextual description exists, i.e., that all probabilities $\Pr\left[V_i \neq W_i\right]$, $i = 1, \ldots, n$, are at their minimum values allowed by the marginals. In Ref. [2], the excess of $\Delta_{\min}$ over its minimum possible value $\frac{1}{2}\sum_{i=1}^{n} \left|\langle V_i\rangle - \langle W_i\rangle\right|$ given the marginals is defined as a measure of contextuality.

To prove Theorem 14, we will first rewrite it in a more accessible form for our proof. Noting that for a pair of ± 1-valued random variables, the probability $\Pr\left[V_i \neq W_i\right]$ is fully determined by the expectation $\langle V_i W_i\rangle$ through the identity

$$\langle V_i W_i\rangle = 1 - 2\Pr\left[V_i \neq W_i\right], \tag{21}$$

we see that minimizing the probability $\Pr\left[V_i \neq W_i\right]$ is equivalent to maximizing the connection expectation $\langle V_i W_i\rangle$. Hence, a description is maximally noncontextual if and only if the connection expectations satisfy

$$\langle V_i W_i\rangle = 1 - \left|\langle V_i\rangle - \langle W_i\rangle\right|, \quad i = 1, \ldots, n,$$

that is, all the connection expectations are at their maximum values as given by Lemma 2. Using (21), we can rewrite Theorem 14 in the following equivalent form:

Theorem 16. *For each $i = 1, \ldots, n$ ($n \geq 2$), let the distribution of a pair $(V_i, W_{i\oplus 1})$ of ± 1-valued random variables be given. The maximum possible value of*

$$S = \sum_{i=1}^{n} \langle V_i W_i\rangle$$

over all possible joints of the given pairs is

$$M := \min \begin{cases} 2n - 2 - \mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle : i = 1, \ldots, n\right), \\ n - \sum_{i=1}^{n} \left|\langle V_i\rangle - \langle W_i\rangle\right|. \end{cases}$$

The proof needs several lemmas:

Lemma 17. *Suppose $a_1, \ldots, a_n \in [-1, 1]$. Then, $\mathsf{s}_0(a_1, \ldots, a_n) + \mathsf{s}_1(a_1, \ldots, a_n) \leq 2n - 2$.*

Lemma 18. *Jointly distributed ± 1-valued random variables $A_1, \ldots, A_n$ ($n \geq 2$) with expectations $\langle A_1\rangle, \ldots, \langle A_n\rangle, \langle A_1 A_2\rangle, \ldots, \langle A_{n-1} A_n\rangle, \langle A_n A_1\rangle$ exist if and only if $A = A_i$ and $B = A_{i\oplus 1}$ satisfy the condition of Lemma 2 for all $i = 1, \ldots, n$ and*

$$\mathsf{s}_0\left(\langle A_1 A_2\rangle, \ldots, \langle A_{n-1} A_n\rangle\right) - (n-2) \leq \langle A_n A_1\rangle \leq (n-2) - \mathsf{s}_1\left(\langle A_1 A_2\rangle, \ldots, \langle A_{n-1} A_n\rangle\right).$$

Furthermore, the range given by the above inequalities is always nonempty and intersects the range

$$|\langle A_n\rangle + \langle A_1\rangle| - 1 \le \langle A_n A_1\rangle \le 1 - |\langle A_n\rangle - \langle A_1\rangle|$$

given by Lemma 2.

Proof. The first part is a direct corollary of Theorem 5 and Lemma 1. The fact that the two ranges intersect follows from the fact that a joint always exists for the chain with expectations $\langle A_1 A_2\rangle, \ldots, \langle A_{n-1} A_n\rangle$ and this joint yields some 2-marginal for (A_n, A_1) which satisfies Lemma 2. $\square$

Lemma 19. *For each $i = 1, \ldots, n$ ($n \ge 2$), let the distribution of a pair $(V_i, W_{i\oplus 1})$ of ± 1-valued random variables be given. Suppose that the main criterion (12) does not hold and formally define the expressions*

$$S := \sum_{i=1}^{n} \langle V_i W_i\rangle,$$
$$M := 2n - 2 - \mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle : i = 1, \ldots, n\right).$$

If there exists some values $\{\langle V_i W_i\rangle : i = 1, \ldots, n\}$ (called connection vector) satisfying the condition of Lemma 2 for each $i = 1, \ldots, n$, the condition

$$\begin{aligned} &\langle V_k W_k\rangle = 1 - |\langle V_k\rangle - \langle W_k\rangle|, \\ &\langle V_i W_i\rangle \ge |1 - |\langle V_k\rangle - \langle W_k\rangle||, \; i \ne k \end{aligned} \tag{22}$$

for some index k, and the inequality $S \le M$, then there exist another connection vector $\{\langle V_i W_i\rangle : i = 1, \ldots, n\}$ such that a joint of the given observable pairs having these connection expectations exists and satisfies the equation $S = M$.

Proof. Consider the set of all connection vectors $\{\langle V_i W_i\rangle : i = 1, \ldots, n\}$ satisfying the condition of Lemma 2 for each pair (V_i, W_i), $i = 1, \ldots, n$, and the condition (22). Within this set we have the assumed configuration satisfying $S \le M$ and the configuration $\langle V_i W_i\rangle = 1 - |\langle V_i\rangle - \langle W_i\rangle|$, $i = 1, \ldots, n$, which yields $S > M$ due to (13) not being satisfied. Since S is a continuous function of the vector $\{\langle V_i W_i\rangle : i = 1, \ldots, n\}$ and the conjunction of the condition of Lemma 2 and condition (22) defines a connected set of connection vectors, there exist a configuration $\{\langle V_i W_i\rangle : i = 1, \ldots, n\}$ satisfying the condition of Lemma 2 and (22) with equality $S = M$. For this configuration, expanding the definitions of S and M we obtain

$$\mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle : i = 1, \ldots, n\right) + \sum_{i=1}^{n} \langle V_i W_i\rangle = 2n - 2. \tag{23}$$

Also, since (22) implies condition 1 of Lemma 7 we obtain from Lemma 8

$$\mathsf{s}_0\left(\langle V_i W_i\rangle : i = 1, \ldots, n\right) = \sum_{i=1}^{n} \langle V_i W_i\rangle, \tag{24}$$

$$\mathsf{s}_1\left(\langle V_i W_i\rangle : i = 1, \ldots, n\right) = \sum_{i \ne k} \langle V_i W_i\rangle - \left(1 - |\langle V_k\rangle - \langle W_k\rangle|\right). \tag{25}$$

Now (23) and (24) imply

$$\mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle : i=1,\dots,n\right)+\mathsf{s}_0\left(\langle V_i W_i\rangle : i=1,\dots,n\right)=2n-2$$

and Lemma 2 applied to (24) and (25) yield

$$\begin{aligned}
&\mathsf{s}_0\left(\langle V_i W_{i\oplus 1}\rangle : i=1,\dots,n\right)+\mathsf{s}_1\left(\langle V_i W_i\rangle : i=1,\dots,n\right)\\
&\le \sum_{i=1}^{n}\left(1-\left|\langle V_i\rangle-\langle W_{i\oplus 1}\rangle\right|\right)+\sum_{i\neq k}\left(1-\left|\langle V_i\rangle-\langle W_i\rangle\right|\right)-\left(1-\left|\langle V_k\rangle-\langle W_k\rangle\right|\right)\\
&=2n-2+\underbrace{\left|\langle V_k\rangle-\langle W_k\rangle\right|-\sum_{i=1}^{n}\left|\langle V_i\rangle-\langle W_{i\oplus 1}\rangle\right|-\sum_{i\neq k}\left|\langle V_i\rangle-\langle W_i\rangle\right|}_{\le 0\,\text{(by the triangle inequality (20))}}.
\end{aligned}$$

Thus, by By Lemma 1 and Theorem 5, a joint exists. □

Now we are ready to prove the main result. It suffices to prove the following partial statement (Theorem 16 and hence also Theorem 14 follow immediately):

Theorem 20. *For each $i=1,\dots,n$ ($n\ge 2$), let the distribution of a pair $(V_i,W_{i\oplus 1})$ of ± 1-valued random variables be given. If the main criterion (12) does not hold, then the maximum possible value of*

$$S=\sum_{i=1}^{n}\langle V_i W_i\rangle$$

over all possible joints of the given pairs is given by

$$M=2n-2-\mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle : i=1,\dots,n\right)\ge n-2. \tag{26}$$

Proof. A larger value cannot be reached since we have

$$\begin{aligned}
&2n-2\\
\text{(Theorem 5)}&\ge \mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle, \langle V_i W_i\rangle : i=1,\dots,n\right)\\
\text{(Lemma 1)}&\ge \mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle : i=1,\dots,n\right)+\mathsf{s}_0\left(\langle V_i W_i\rangle : i=1,\dots,n\right)\\
&\ge \mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle : i=1,\dots,n\right)+S.
\end{aligned}$$

To show that this value can be reached, without loss of generality (by Lemmas 9 and 10) assume

$$\langle V_i W_{i\oplus 1}\rangle\ge\left|\langle V_n W_1\rangle\right|,\quad i=1,\dots,n-1, \tag{27}$$

and since the criterion (12) is not satisfied, by Lemma 12, the maximal connection expectations $1-\left|\langle V_i\rangle-\langle W_i\rangle\right|$ satisfy for some k the condition

$$1-\left|\langle V_i\rangle-\langle W_i\rangle\right|\ge\left|1-\left|\langle V_k\rangle-\langle W_k\rangle\right|\right|,\quad i\neq k. \tag{28}$$

We will show that a specific configuration $\{\langle V_i W_i\rangle : i=1,\dots,n\}$ exists for which $S=M$ and for which a joint of the pairs $(V_i,W_{i\oplus 1})$, $i=1,\dots,n$, exists. This is shown separately for four cases that exhaust all possible situations under the present assumptions. The four cases are defined as follows:

$$\begin{tikzcd}
\forall i : \left|\langle V_i\rangle + \langle W_i\rangle\right| - 1 \leq \left|1 - \left|\langle V_k\rangle - \langle W_k\rangle\right|\right|? \arrow[d, "No"'] \arrow[dr, "Yes"] & \\
\text{Case 1} & 1 - \left|\langle V_k\rangle - \langle W_k\rangle\right| \geq 0? \arrow[d, "No"] \arrow[dl, "Yes"'] \\
\langle V_n W_1\rangle \leq 1 - \left|\langle V_k\rangle - \langle W_k\rangle\right|? \arrow[d, "No"'] \arrow[dr, "Yes"] & \text{Case 2} \\
\text{Case 3} & \text{Case 4}
\end{tikzcd}$$

Case 1. We have for some index j,

$$\left|\langle V_j\rangle + \langle W_j\rangle\right| - 1 > \left|1 - \left|\langle V_k\rangle - \langle W_k\rangle\right|\right|. \tag{29}$$

In this case, we define $\langle V_i W_i\rangle = 1 - \left|\langle V_i\rangle - \langle W_i\rangle\right|$ for all $i \neq j$ and choose an arbitrary value $\langle V_j W_j\rangle \geq \left|\langle V_j\rangle + \langle W_j\rangle\right| - 1$ from the range given by Lemma 18 (intersected with the range given by Lemma 2). Since a joint exists by Lemma 18 for this configuration, we obtain $S \leq M$. Furthermore, since (28) and (29) imply that condition (22) is satisfied for this configuration, Lemma 19 implies that a joint exists for some configuration with $S = M$.

Case 2. We have

$$\begin{aligned}
&1 - \left|\langle V_k\rangle - \langle W_k\rangle\right| < 0, \\
&\left|\langle V_j\rangle + \langle W_j\rangle\right| - 1 \leq \left|1 - \left|\langle V_k\rangle - \langle W_k\rangle\right|\right|, \quad j \neq k.
\end{aligned} \tag{30}$$

In this case, we choose an index $j \neq k$ and define $\langle V_j W_j\rangle = \left|1 - \left|\langle V_k\rangle - \langle W_k\rangle\right|\right| = -\left(1 - \left|\langle V_k\rangle - \langle W_k\rangle\right|\right)$ and $\langle V_i W_i\rangle = 1 - \left|\langle V_i\rangle - \langle W_i\rangle\right|$ for $i \neq j$. These values satisfy condition (22) by (28) and the condition of Lemma 2 by (28) and (30). Now $\langle V_j W_j\rangle + \langle V_k W_k\rangle = 0$ implying that $S \leq n - 2 \leq M$ and so by Lemma 19 a joint exists for a configuration with $S = M$.

Case 3. We have

$$\begin{aligned}
&0 \leq 1 - \left|\langle V_k\rangle - \langle W_k\rangle\right| < \langle V_n W_1\rangle, \\
&\left|\langle V_i\rangle + \langle W_i\rangle\right| - 1 \leq 1 - \left|\langle V_k\rangle - \langle W_k\rangle\right|, \quad i = 1, \ldots, n.
\end{aligned} \tag{31}$$

In this case, we define $\langle V_i W_i\rangle = 1 - \left|\langle V_k\rangle - \langle W_k\rangle\right| \geq 0$ for all $i = 1, \ldots, n$ and condition (22) is satisfied trivially. These values also satisfy the condition of Lemma 2 by (28) and (31). Furthermore, by (27) and (31), all product expectations $\{\langle V_i W_{i\oplus 1}\rangle, \langle V_i W_i\rangle : i = 1, \ldots, n\}$ are nonnegative and the smallest value among them is $\langle V_i W_i\rangle = 1 - \left|\langle V_k\rangle - \langle W_k\rangle\right|$. So, by Lemma 8, we can expand

$$\begin{aligned}
&\mathsf{s}_1\left(\langle V_i W_{i\oplus 1}\rangle, 1 - \left|\langle V_k\rangle - \langle W_k\rangle\right| : i = 1, \ldots, n\right) \\
&= \sum_{i=1}^{n} \langle V_i W_{i\oplus 1}\rangle + (n-2)\left(1 - \left|\langle V_k\rangle - \langle W_k\rangle\right|\right) \\
&\leq n + (n-2) \leq 2n - 2.
\end{aligned}$$

Thus, a joint exists by Theorem 5 which implies that $S \leq M$ and so by Lemma 19 a joint exists for a configuration with $S = M$.

Case 4. We have

$$\begin{aligned} \langle V_n W_1 \rangle \leq 1 - |\langle V_k \rangle - \langle W_k \rangle| \geq 0, \\ |\langle V_i \rangle + \langle W_i \rangle| - 1 \leq 1 - |\langle V_k \rangle - \langle W_k \rangle|, \quad i = 1, \dots, n. \end{aligned} \tag{32}$$

In this case, we define $\langle V_i W_i \rangle = 1 - |\langle V_k \rangle - \langle W_k \rangle| \geq |\langle V_i \rangle + \langle W_i \rangle| - 1$ for $i \neq k$ and (28) implies that these values satisfy the condition of Lemma 2. Thus, a joint exists by Lemma 18 for

$$\langle V_k W_k \rangle = \min \begin{cases} 2n - 2 - \mathsf{s}_1\big(\langle V_i W_{i\oplus 1} \rangle : i = 1, \dots, n; \langle V_i W_i \rangle : i \neq k\big), \\ 1 - |\langle V_k \rangle - \langle W_k \rangle|. \end{cases} \tag{33}$$

Assume first that the top expression in (33) is the minimum. Since here $\langle V_n W_1 \rangle$ is by (27) and (32) the smallest argument to $\mathsf{s}_1(\dots)$ in (33) and all other arguments are nonnegative Lemma 8 yields

$$\begin{aligned} &\mathsf{s}_1\big(\langle V_i W_{i\oplus 1} \rangle : i = 1, \dots, n; \langle V_i W_i \rangle : i \neq k\big) \\ &= \sum_{i=1}^{n-1} \langle V_i W_{i\oplus 1} \rangle - \langle V_n W_1 \rangle + (n-1)(1 - |\langle V_k \rangle - \langle W_k \rangle|) \\ &= \mathsf{s}_1(\langle V_i W_{i\oplus 1} \rangle : i = 1, \dots, n) + (n-1)(1 - |\langle V_k \rangle - \langle W_k \rangle|) \\ &= \mathsf{s}_1(\langle V_i W_{i\oplus 1} \rangle : i = 1, \dots, n) + S - \langle V_k W_k \rangle. \end{aligned}$$

Substituting this in (33) yields

$$S = 2n - 2 - \mathsf{s}_1(\langle V_i W_{i\oplus 1} \rangle : i = 1, \dots, n) = M.$$

Suppose then that the bottom expression in (33) is the minimum. Then, we have

$$\langle V_i W_i \rangle = 1 - |\langle V_k \rangle - \langle W_k \rangle| \geq 0, \quad i = 1, \dots, n$$

and condition (22) is satisfied and since a joint exists, it follows $S \leq M$. But then by Lemma 19 a joint exists for a configuration with $S = M$. □

5 Conclusion

We have derived the formula (6) for contextuality measure in cyclic systems with binary variables, previously conjectured based on the principles laid out in Refs. [2,3]. The measure is based on the principled comparison of the minimal values Δ_0 and $\Delta_{\min}$ of

$$\Delta = \sum_{i=1}^{n} \Pr[V_i \neq W_i],$$

with Δ_0 computed across all couplings for the individual pairs $\{V_i, W_i\}$ (i.e., by minimizing each $\Pr[V_i \neq W_i]$ separately), and with $\Delta_{\min}$ computed across all couplings for the entire set of the random variables $\{V_i, W_i : i = 1, \dots, n\}$.

A logical consequence of this measure is the criterion (necessary and sufficient condition) for contextuality (7). In this paper, however, we proved (7) by showing its equivalence to the criterion (8) derived in Ref. [3], and used this equivalence to derive (6).

Acknowledgments

This work is supported by NSF grant SES-1155956, AFOSR grant FA9550-14-1-0318, and A. von Humboldt Foundation. The authors benefited from collaboration with Acacio de Barros, Gary Oas, and Jan-Åke Larsson.

References

[1] A. Acín, T. Fritz, A. Leverrier, A.B. Sainz (2015). A Combinatorial Approach to Nonlocality and Contextuality. Commun. Math. Phys. 334, 533–628.

[2] Dzhafarov, E.N., & Kujala, J.V., & Larsson, J-Å (2015). Contextuality in three types of quantum-mechanical systems. Found. Phys. , 762-782.

[3] J.V. Kujala, E.N. Dzhafarov, & J.-Å. Larsson (2015). Necessary and sufficient conditions for extended noncontextuality in a broad class of quantum mechanical systems. Phys. Rev. Lett. 115, 150401.

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

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

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

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

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

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

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

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