

## Article

# Dendrographic Representation of Data: CHSH Violation vs. Nonergodicity

Oded Shor<sup>1,2</sup>, Felix Benninger<sup>1,2,3</sup>  and Andrei Khrennikov<sup>4,\*</sup>

<sup>1</sup> Felsenstein Medical Research Center, Beilinson Hospital, Petach Tikva 4941492, Israel; shor.oded@gmail.com (O.S.); felixbenninger@gmail.com (F.B.)

<sup>2</sup> Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 6997801, Israel

<sup>3</sup> Department of Neurology, Rabin Medical Center, Petach Tikva 4941492, Israel

<sup>4</sup> Department of Mathematics, Faculty of Technology, Linnaeus University, 35195 Växjö, Sweden

\* Correspondence: Andrei.Khrennikov@lnu.se

**Abstract:** This paper is devoted to the foundational problems of *dendrographic holographic theory* (DH theory). We used the ontic–epistemic (implicate–explicate order) methodology. The epistemic counterpart is based on the representation of data by dendrograms constructed with hierarchic clustering algorithms. The ontic universe is described as a p-adic tree; it is zero-dimensional, totally disconnected, disordered, and bounded (in p-adic ultrametric spaces). Classical–quantum interrelations lose their sharpness; generally, simple dendrograms are “more quantum” than complex ones. We used the CHSH inequality as a measure of quantum-likeness. We demonstrate that it can be violated by classical experimental data represented by dendrograms. The seed of this violation is neither nonlocality nor a rejection of realism, but the nonergodicity of dendrographic time series. Generally, the violation of ergodicity is one of the basic features of DH theory. The dendrographic representation leads to the local realistic model that violates the CHSH inequality. We also considered DH theory for Minkowski geometry and monitored the dependence of CHSH violation and nonergodicity on geometry, as well as a Lorentz transformation of data.

**Keywords:** ontic; epistemic; dendrograms; dendrographic holographic theory; clustering algorithms; quantumness; CHSH inequality; nonergodicity



**Citation:** Shor, O.; Benninger, F.; Khrennikov, A. Dendrographic Representation of Data: CHSH Violation vs. Nonergodicity. *Entropy* **2021**, *23*, 971. <https://doi.org/10.3390/e23080971>

Academic Editors: Rosario Lo Franco, Arkady Plotnitsky and Gregg Jaeger

Received: 13 June 2021  
Accepted: 23 July 2021  
Published: 28 July 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

A model for the Universe based on the hierarchic relational representation of its components was suggested in [1]. The present paper is devoted to the foundational problems structured within Primas and Atmanspacher’s [2,3] ontic–epistemic and Herz and Boltzmann’s [4,5] descriptive observational structuring of physical theories (see also [6–8]), as well as Bohm’s [9] implicate–explicate order viewpoint on the Universe. One of the aims of our study was to re-establish local realism in physics (cf. [10–18]).

We describe the representation of experimental data by dendrograms using *hierarchic clustering algorithms*. They lead to the p-adic model at the ontic level. (From this viewpoint, our research is part of p-adic theoretical physics, see, for example, [19–28] and Section 6.) In [1], we named our theory *dendrographic holographic theory* (DH theory). In this paper, we do not discuss holography (see [1]) but proceed with the same name.

The p-adic tree ( $p > 1$  is a natural number) is the infinite tree with a homogeneous structure of branching at each node, one incoming and p-outcoming edges. This tree is endowed with a natural metric (so-called “ultrametric”) and the algebraic structure of a ring (see Appendix A).

We recall that an ontic theory is about reality as it is, and an epistemic theory is about knowledge collected through observations. The ontic universe is unapproachable by observers. Nevertheless, its structure can be theoretically reconstructed by increasing the size of dendrograms; observer  $O$  generates approximations of the ontic universe. The

limiting mathematical structure is uniquely determined by the p-adic tree (or algebraically, the p-adic ring, denoted as  $\mathbf{Z}_p$ ).

One of the main aims of novel mathematical modeling of ontic–epistemic structuring is to better understand the classical–quantum interrelation. Recalling that *quantum mechanics is an epistemic model*, it is about observations’ outcomes (see Bohr [29,30] and Plotnitsky [31–33]). In DH theory, “quantumness” is present only at the level of dendrograms, and the p-adic ontic universe is classical. The heuristic criterium of quantum-likeness is very natural: simple dendrograms are more quantum(-like) than complex dendrograms. As a test of quantum-likeness, we used the *CHSH inequality*, not for the standard observables, but for new observables reflecting the hierarchic relational structure of data. It should be highlighted that the original experimental data used for this paper are classical. Its quantum-like structure becomes visible through dendrographic representation. The seed of the violation of the CHSH inequality is neither a nonlocality nor a rejection of realism, but the *nonergodicity* of dendrographic time series. The dendrographic representation leads to a local realistic model that violates the CHSH inequality.

This is a good place to recall that a few authors have coupled quantum behavior with the violation of ergodicity [34–38]. We also considered DH theory for Minkowski geometry and monitored the dependence of CHSH violation and nonergodicity on geometry as well as a Lorentz transformation of data.

The reader who is not interested in foundations can jump directly to Section 7, which is devoted to the experimental data, correlations, and the violation of the CHSH inequality for hierarchic observables and its correlation with the nonergodicity of data.

## 2. Dendrogram Representation of the Universe

### 2.1. Systems as Clusters of Clicks of Detectors

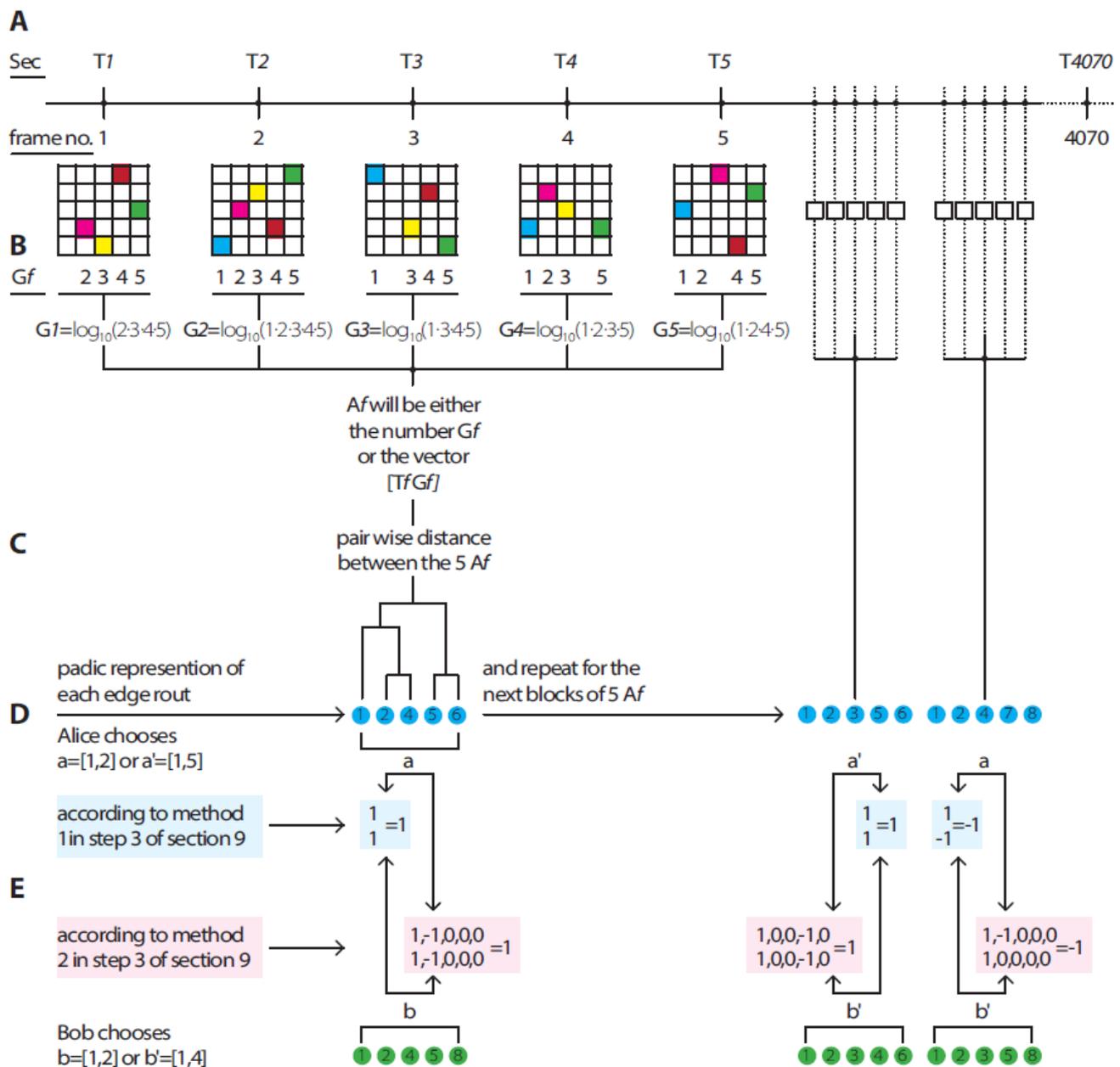
According to Bohr, the outcomes of measurements are not the objective properties of systems [29–33]. They quantitatively represent interrelation between a system  $S$  and an observer  $O$ . An ontic system exists independently of  $O$ . However, it is unapproachable by the observer; in fact,  $O$  constructs its approximate representation by using data.

Consider an observer  $O$  constructing a representation of a surrounding environment—in an extreme case, of the Universe. In DH theory, the observer has the free choice not only to perform or not perform some observation, but even the free choice to decompose the collected data into blocks and to treat these blocks as the (epistemic) systems’ representation. *The fundamental entities of the epistemic theory are events* (cf. [14,39–43]), *not systems*. Systems are composed of a certain number of events (clusters of events), for example, several clicks of detectors (or other detection events). This position is very close to the view of Zeilinger on quantum mechanics (learned through private discussions with Akh).

### 2.2. Representation of Systems by Dendrograms

A simple collection of experimental statistical data does not tell observer  $O$  about the *genuine hierarchic* interrelations between physical processes under observations. Such hierarchic structure can be approximately recovered with the aid of *clustering algorithms*. This hierarchic clustering leads to the treelike geometry. In DH theory, (epistemic) systems are described by *dendrograms*, or finite trees. Any observer  $O$  can reconstruct their environment and, in the limit, the whole  $O$ -Universe, by collecting observational data.

In DH theory, the *decomposition of the  $O$ -Universe into subsystems is done through the decomposition of data into blocks*. Each block of data is represented by a dendrogram—this is a system. The data can be subdivided into blocks in different ways. The  $O$ -Universe can be decomposed into systems of different complexities depending on the block’s size. We illustrate this process in Figure 1 by dividing data into blocks of equal size, where  $d = 5$ .



**Figure 1.** Abstract and numerical examples of the process from real data of a CCD chip (illustrated as a  $5 \times 5$  CCD chip). Double-slit experiment via the construction of a dendrogram (with 5 edges corresponding to 5 frames) with a series of 1 and  $-1$  values, from which CHSH correlations were calculated.

### 2.3. *p*-Adic Metric for Hierarchic Relations

The points at the baselevel of any dendrogram *D* (see Figure 1) represent the relational characteristics of the system expressed by *D*. For the simplest hierarchic algorithms based on the choices of just two alternatives—two branchings at the vertexes, the base points can be described by vectors with coordinates 0/1. The natural distance between these characteristics is given by 2-adic ultrametric  $r_2$ ; the longer the common root of the paths going from the dendrogram’s root to points *a, b*, the shorter the distance between them (see Appendix A). This metric encodes the hierarchic relations inside the dendrographic system and with its surrounding environment. More general hierarchic clustering algorithms with *p*-branching at vertexes generate the representation of systems by dendrograms having more complex topology. (Here,  $p > 1$  is a natural number determining the tree

structure). The branches (and the corresponding endpoints) are encoded by vectors with the coordinates belonging to the set  $\{0, 1, \dots, p - 1\}$ . The corresponding common root distance is denoted by  $r_p$ .

Take two points,  $a, b$ , at the baselevel of dendrogram  $D$  representing a system. The common root distance  $r_p(a, b)$  gives the quantification of the closeness of these two systems' characteristics with respect to the hierarchic order relation. The quantity

$$R_D = \max\{r_p(a, b) : a, b \text{ belong to } D\}$$

(dendrogram's diameter) determines the degree of hierarchic interrelation between the elements of the dendrogram system. If  $R_D$  is small, the system's characteristics are highly relationally connected, if  $R_D$  is large, the system is relationally sparse.

#### 2.4. Epistemic Realism, Hidden Variables

Each dendrogram  $D$  can be represented by the vectors with coordinates belonging to the set  $\{0, 1, \dots, p - 1\}$ . These are (hierarchic relational) *hidden variables*. Observer  $O$  can invent relational observables defined as functions of these hidden variables, for example,  $A = A(u)$ . This is *the realistic model of observations*, which is constructed by the observer  $O$  through hierarchic structuring of the experimental data. This approach can be called *realism without reality*, as opposed to the *reality without realism* viewpoint on quantum physics suggested by Plotnitsky [33,44]. Such hidden variables are constructed on the basis of experimental data; they are, so to say, epistemic variables. However, the mathematical model is realistic in the sense of a set-theoretical representation of states as function of hidden variables and observables.

### 3. From Epistemic (Explicate) to Ontic (Implicate): p-Adic Universe

By considering the data blocks of increasing size, in the limit,  $O$  can reconstruct the ontic description. The dendrogram-based epistemic theory leads to the p-adic geometry of the ontic universe. (The role of the model parameter  $p$  is discussed in Section 5.) In the limit, dendrograms of increasing size generate the infinite p-adic tree endowed with the p-adic ultrametric. Its infinite branches are points of the ring  $\mathbb{Z}_p$ ; in fact,  $\mathbb{Z}_p$  coincides with the unit ball. Thus, with respect to the hierarchic p-metric, *the Universe is bounded*. Topologically the p-adic universe differs crucially from the Universe endowed with the Euclidean or Minkowski geometry. The basic properties of  $\mathbb{Z}_p$  match very well with Bohm's views on implicate order. This space *is disordered, zero-dimensional, and totally disconnected*.

### 4. From Ontic (Implicate) to Epistemic (Explicate)

In Section 3, the mathematical structure of the ontic universe was approached from the special dendrogram representation of experimental data. Now, we proceed with another way around—from the ontic p-adic model to the epistemic dendrographic model. Our previous epistemic-to-ontic pathway (Section 3) centralized the role of an observer  $O$ . This can make the impression that the personal observer perspective plays a crucial role in DH theory (cf. with QBism [45]). Now, by starting with the p-adic ontic model, we diminish the subjective component.

The points of the p-adic tree  $\mathbb{Z}_p$  represent all possible events that can happen in the Universe. Thus, as well as the epistemic universe, the ontic universe is represented as a set of events. However, these are not observational events; they cannot be associated with, say, the clicks of detectors. We consider the p-adic points *absolute events*. An absolute event is the endpoint of the infinitely long path of the p-adic tree; mathematically, the sequence is

$$a = a_1 \dots a_n \dots, \text{ with coordinates } a_j = 0, 1, \dots, p - 1.$$

Each finite cutoff of the infinite path  $a$  is an element of the explicate order, and it can be interpreted as the characteristic of a physical system. The latter is determined in a variety of ways depending on the embedding of the finite path  $a' = a_1 \dots a_n$  into various

dendrograms. We considered the following process of the structuring of the  $Z_p$  universe. Let us select some set of points  $S$  of  $Z_p$  (finite or infinite). These points are geometrically presented as paths starting at the root of the  $p$ -adic tree. Thus,  $S$  is a subtree of this tree. The common roots of these paths determine the relational closeness of corresponding points in  $S$ . As well as in the epistemic model, these points can be considered as characteristics of  $S$ . Observer  $O$  is not involved in this consideration. Roughly speaking, the system  $S$  exists irrespectively, whether somebody looks at it or not.

Now, take the tree of an ontic system  $S$  and cut it at some level, say,  $n$  steps from the root. Such a dendrogram expresses a system  $S_n$  belonging to the explicate level of nature's description. In principle, there is no need to couple it to some observer. This construction provides the realist interpretation, even for systems represented by finite dendrograms. The only subjective element is the place of cutoff (cf. with the problem of the boundary between the quantum and classical worlds).

### 5. The Role of the Clustering Algorithm

Different hierarchic clustering algorithms generate different dendrograms and observers' free will also covers the algorithm's choice. However, in the ontic description, this dependence on the algorithm's choice is washed out. All algorithms with  $p$ -alternative branching generate approximations for the same ontic model of the Universe, given by the  $p$ -adic tree. The rings of  $p$ -adic numbers  $Z_p$  are not isomorphic for a different  $p$ . However, their topological and algebraic properties are the same. Similar problems have been discussed in  $p$ -adic theoretical physics [19–28]: What  $p$  does correspond to the physical reality? There are various suggestions, including such exotic choices as  $p = 127$  [26]. The pragmatic solution was to treat all  $p$ -adic models on equal grounds and consider the  $p$ -adic encoding similarly to the representation of real numbers by using different bases.

Of course, there exist clustering algorithms that generate nonhomogeneous trees, with different branching indexes for different vertexes. Such dendrographic models lead to arbitrary trees (arbitrary ultrametric spaces). One may try to exclude such models by appealing to the relational scale homogeneity of nature. In any event, we keep to  $p$ -adic models. Our original choice of the concrete hierarchic clustering algorithm [1] was motivated by studies, including the data analysis of Murtagh [46–48]. In the present study, we played with a few other algorithms, but their outputs are very similar with slight differences in details. In Section 11.3, we apply seven different linkage clustering algorithms with no apparent differences in CHSH value outcomes (Figure 5 in Section 11.3). We refer the interested reader to the Matlab software website for details on those seven linkage clustering algorithms [49].

### 6. "Traditional $p$ -Adic Theoretical Physics": Emphasis of the Number-Theoretic Structure

On  $p$ -adic trees, it is possible to introduce arithmetic operations; algebraically, these spaces are similar to the real line. However, topologically  $p$ -adic trees differ crucially from the real line.  $p$ -adic theoretical physics (see, e.g., [19–28] and references therein) flourished in the 1990s, stimulated by string theory [19,21,22]. Its main problem, which finally led to an essential dampening of its development during recent years, was the absence of coupling with real experimental data. Roughly speaking, there are no measurement devices that generate  $p$ -adic outputs. However, as was clearly stated by Volovich [26], there are also no measurement devices that generate irrational numbers; due to the finite precision of measurements, all their outputs are rational numbers. Volovich claimed that the only physical numbers are rational numbers. Since rational numbers are dense, both in real and  $p$ -adic fields, Volovich suggested the unification of standard and  $p$ -adic physics on the basis of rational numbers. Although he did not use the ontic–epistemic terminology, we can say that, for him, mathematically, epistemology is based on rational numbers. Beyond this general rational epistemology, one can recover the real and  $p$ -adic ontic models. We suggest another epistemology for  $p$ -adic ontology. This is the hierarchic relational epistemology. It

differs crucially from the real-order epistemology of “standard physics”. We can say that Volovich’s approach was number-theoretical and our approach is hierarchic relational.

The DH framework provides the rigid coupling of theory and experimental data. We hope that it will lead to the renaissance of p-adic modeling in physics. (We guess that its main theoretical results can be adapted to DH theory.)

## 7. Dendrogram Viewpoint on Classical–Quantum Interrelation

In our epistemic model, the sharp classical–quantum separation disappears. *The degree of classicality is based on a system’s complexity*—the size and topological complexity of its dendrogram representation. Quantum systems are characterized by a low complexity of their dendrograms. Thus, electrons and atoms can behave as quantum in some measurements because they have a very simple hierarchic structure of interrelation between their components. In our approach, not only can systems be described by quantum mechanics but even classical physical systems can exhibit quantum(-like) behavior within a hierarchic representation of experimental data.

Bell-type inequalities (and in particular, the CHSH inequality) have been used to show the “quantumness” degree of photons or particle systems [50–55]. We selected four system characteristics (the endpoints of dendrograms representing systems)  $a, a', b, b'$  and calculated the correlations for pairs  $(a, b)$ ,  $(a, b')$ ,  $(a', b)$ , and  $(a', b')$ . Finally, a CHSH combination of correlations was formed and the problem of its exceeding of two was studied via a numerical simulation. For Alice and Bob, the fixed settings are  $a, a'$  and  $b, b'$ , respectively. The experimenter calculated the correlations  $C_{ab}$ ,  $C_{ab'}$ ,  $C_{a'b}$ ,  $C_{a'b'}$ , and then the CHSH combination of them, as follows:

$$C = C_{ab} - C_{ab'} + C_{a'b} + C_{a'b'}$$

*Inequality  $|C| > 2$  means data collected in experimental runs is consistent with predictions of quantum theory and contradicts local realism, that is, it cannot be described by a hidden variable model that is local.*

We considered two time series of experimental data. The data was collected in measurements of one fixed classical physical observable. The outcomes of this observable carry the information about hierarchic relations between the characteristics of a physical system under measurement. (The data was decomposed in the systems by the observer.) These relations are not visible in the straightforward graphic representation of the data (see Figure 1). They were extracted from “real data” with the aid of a clustering algorithm.

The standard viewpoint on the violation of the Bell-type inequalities [50–55] is that it shows that “quantum correlations are stronger than classical ones.”

By trying to explain this difference between quantum and classical correlations, Bell and his followers pointed to a contradiction between quantum theory and local realism. The latter is mathematically formalized as a representation of observables by functions of hidden variables; locality means that, say, Alice’s functions do not depend on the settings selected for Bob’s functions, that is,  $S_a = S_a(u)$ , where  $u$  is the hidden variable.

The interpretation of the violation of the Bell-type inequalities is the subject of hot foundational debates (see, e.g., [56–60]). Plenty of so-called loopholes were pointed out. Two basic ones, the nonlocality and the detection efficiency loopholes [61,62], were successfully closed in 2015 experiments [63–65].

## 8. The Role of Nonergodicity

De Broglie [66] was the first to highlight that experimentally obtained probabilities may deviate from the ontic probability distribution of hidden variables. He justified this possibility within his double solution theory in terms of the pilot wave. It seems that Krennikov’s works [67–69] were the first publications in which the difference between the frequency experimental probabilities and the measure-theoretic distributions of hidden variables was pointed out. Khrennikov appealed to the von Mises frequency probability theory. The latter straightforwardly implied that the Bell-type inequalities can be violated

for local models. However, it seems to be impossible to mathematically formalize the notion of realism within the von Mises frequency theory [70]. Finally, in [37], the interplay between frequency and measure-theoretic probabilities was structured within the *ergodic theory* and it was shown that the violation of ergodicity is a sufficient condition for the violation of the Bell-type inequalities. However, this was shown in a purely theoretical framework. In our further considerations, we show that in DH theory, the Bell-type inequalities are violated because of the nonergodicity of the time series of experimental data restructured with hierarchic clustering algorithms.

Hence, in DH theory, the Bell-type inequalities are not a tool for the rejection of the local realistic description or the classical probability model (Kolmogorov [71], 1933). We use these inequalities to express the quantumness degree of experimental statistical data—in the spirit of the article by [72].

As was mentioned in Section 1, a few authors have pointed out that quantum effects can be derived in a classical statistical framework by rejecting the assumption of ergodicity [34–38]. In particular, in [35], it was guessed that the originally very emphasized the difference between quantum and classical statistical modeling can be eliminated by rejecting the hypothesis of the ergodicity of data; even black body radiation can be modeled classically, but within nonergodic statistical mechanics. In the paper by [37], *nonergodicity was pointed as one of the sufficient conditions for the violation of the Bell-type inequalities*, a condition that is equally important as nonlocality and a rejection of realism (the hidden variable description).

DH theory provides the possibility of introducing hidden variables (of the special type) beyond any experimental data and considering a class of realistic observables, those represented as functions of hidden variables. Hence, straightforwardly, it seems that Bell-type inequalities cannot be violated under the assumption of locality. However, we show the violation of the CHSH inequality for special selection of “settings” that determine the observables. The reason for this is precisely the violation of ergodicity; the measure-theoretic and frequency averages do not coincide [37].

However, the natural question arises: Why do the hierarchically determined observables violate the assumption of ergodicity? The answer to this question is that the hierarchical sub-systems have already encoded in them the hierarchical structure of the “universal” dendrogram, as was demonstrated in our paper [1].

We directly check the hypothesis of ergodicity in Section 10 and we show that for a wide range of “settings”, the hierarchic observables are not ergodic. However, we also found that for some “settings”, the observables are ergodic, but, nevertheless, the CHSH inequality is violated. How does this happen? This situation was also discussed in paper [37]:

“Consider two stationary ergodic processes  $x(t)$  and  $y(t)$ . Is the vector process  $z(t) = (x(t), y(t))$  ergodic? The answer is “no”. We remark that this implies that the product  $x(t)y(t)$  of two ergodic processes need not be ergodic. Thus, in principle, in quantum theory nonergodicity can be generated by measurements on compound systems. At the same time, measurements of each system separately generate ergodic stationary processes. In such a case, measurements on compound systems really have the special feature, nonergodicity.”

In addition, we really found that in the case of the violation of the CHSH inequality, the product of Alice and Bob observables is always nonergodic.

## 9. Scheme for the Calculation of Dendrographic Correlations

For concreteness, we consider  $p = 2$ .

**Step 1.** From a data time series to a dendrogram: a hierarchic clustering of data.

Consider a time series of some data,  $Z_1, Z_2, \dots, Z_n$ . Split it into blocks of length  $d$ . (In Figure 1, these blocks have a length of 5). For each block, we constructed (with a hierarchic clustering algorithm) a dendrogram with  $d$  basepoints. We obtained a dendrogram time series,  $D_1, D_2, \dots, D_s$ . All dendrograms have  $k$  levels, where  $2^{k-1} < d \leq 2^k$ . (In Figure 1 step A,  $d = 5$ , hence,  $k = 3$ .)

### Step 2. From a dendrogram time series to a 2-adic time series.

First, we remark that each dendrogram  $D$  can be represented by vectors of the form  $(\alpha_0, \alpha_1, \dots, \alpha_k)$ , where  $\alpha_j = 0, 1$ . Each vector encodes a path on  $D$  from its root to the corresponding endpoint;  $D$  is the selection of  $d$  vectors from  $2^k$  vectors.

From this vector representation, we moved to representations by natural numbers, which we define as features of the dendrogram, by using the following formula:

$$(\alpha_0 \alpha_1 \dots \alpha_k) = \alpha_0 + \alpha_1 2 + \alpha_2 2^2 + \dots + \alpha_k 2^k.$$

Thus, from a series of dendrograms  $D_1, D_2, \dots, D_s$ , we obtained a series of natural numbers, which were sorted in ascending order in each block of length  $d$ , according to the hierarchical structure of each dendrogram  $D_n$ .

These natural numbers play the role of polarization in the real CHSH experiment. Later, we define new observables depending on such analogs of polarization. These  $d$ -blocks of natural numbers can be treated as hidden variables. New observables will be functions of these "hierarchic hidden variables".

Then, for correlations, we considered two time series,  $Z_1, Z_2, \dots, Z_n$  and  $Z_1', Z_2', \dots, Z_n'$ . By applying the procedures of Steps 1 and 2, we obtained two series of dendrograms  $D_1, D_2, \dots, D_s$  and  $D'_1, D'_2, \dots, D'_s$ , and two corresponding series of natural numbers. They were formed from consecutive blocks of  $d$  natural numbers, which were generated through the  $d$ -decomposition of the original time series:

$x = x_1, x_2, \dots, x_n$ ,  $x_n$  is composed of blocks with  $d$  natural numbers

$y = y_1, y_2, \dots, y_n$ ,  $y_n$  is composed of blocks with  $d$  natural numbers

### Step 3. Defining the observables.

**First method** (see Figure 1 step E).

We constructed two time series of dendrograms, one for **Alice**  $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \dots, \mathbf{A}_n)$  and one for **Bob**  $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \dots, \mathbf{B}_n)$ .

For **Alice**, we selected two pairs of numbers,  $\mathbf{a} = [a_1 \ a_2]$  and  $\mathbf{a}' = [a_1' \ a_2']$ . The two pairs are not identical. The analogs of the two vectors are orientations of polarization beam splitters or Stern–Gerlach magnets.

For **Bob**, we selected two pairs of numbers,  $\mathbf{b} = [b_1 \ b_2]$  and  $\mathbf{b}' = [b_1' \ b_2']$ . The two pairs are not identical.

For each  $\mathbf{A}_i$   $i \in 1, 2, 3, \dots, n$ , we randomly decided between pair  $\mathbf{a}$  or  $\mathbf{a}'$ .

For each  $\mathbf{B}_i$   $i \in 1, 2, 3, \dots, n$ , we randomly decided between pair  $\mathbf{b}$  or  $\mathbf{b}'$ .

If, for  $\mathbf{A}_i$ , we chose pair  $\mathbf{a}$ , we obtained the following:

If  $\mathbf{A}_i$  had both of the numbers in  $\mathbf{a}$ ,  $S_{ai} = 1$ , otherwise,  $S_{ai} = -1$ .

Metaphorically, we can say that if the "polarization" of  $\mathbf{A}_i$  coincides with  $\mathbf{a}$ , the detector with the output is +1 clicks, if not, the detector with the output is -1 clicks.

We proceeded in the same way for the selection of  $\mathbf{a}'$  and  $\mathbf{b}, \mathbf{b}'$ . Then we calculated the correlations, as follows:

$$C_{ab} = (\sum S_{ai} S_{bi}) / \text{length}(\mathbf{a} \text{ and } \mathbf{b} \text{ are selected together})$$

$$C_{ab'} = (\sum S_{ai} S_{b'i}) / \text{length}(\mathbf{a} \text{ and } \mathbf{b}' \text{ are selected together})$$

$$C_{a'b} = (\sum S_{a'i} S_{bi}) / \text{length}(\mathbf{a}' \text{ and } \mathbf{b} \text{ are selected together})$$

$$C_{a'b'} = (\sum S_{a'i} S_{b'i}) / \text{length}(\mathbf{a}' \text{ and } \mathbf{b}' \text{ are selected together})$$

$$C = C_{ab} - C_{ab'} + C_{a'b} + C_{a'b'}$$

**Second method** (see Figure 1 step E).

We constructed two time series of dendrograms, one for **Alice**  $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \dots, \mathbf{A}_n)$  and one for **Bob**  $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \dots, \mathbf{B}_n)$ .

For **Alice**, we selected two pairs of numbers,  $\mathbf{a} = [a_1 \ a_2]$  and  $\mathbf{a}' = [a_1' \ a_2']$ . The two pairs are not identical.

For **Bob** we select two pairs of numbers  $\mathbf{b} = [b_1 \ b_2]$   $\mathbf{b}' = [b_1' \ b_2']$  the two pairs aren't identical. For each  $\mathbf{A}_i$   $i \in 1, 2, 3 \dots, n$ , we randomly decided between pair  $\mathbf{a}$  or  $\mathbf{a}'$

If for  $A_i$  we chose pair  $a$  we have the following  
 We replace the natural number in  $A_i$  that equal  $a_1$  to 1  
 We replace the natural number in  $A_i$  that equal  $a_2$  to  $-1$   
 We replace all other natural numbers in  $A_i$  to zero. We indicate that pair  $a$  was chosen to  $A_i$ .  
 The same we do for other settings,  $a', b, b'$ -  
 We find  $i$  that  $a$  and  $b$  were chosen =  $na_b$ ,  
 We find  $i$  that  $a'$  and  $b$  were chosen =  $na'_b$ ,  
 We find  $i$  that  $a$  and  $b'$  were chosen =  $na_b'$ ,  
 We find  $i$  that  $a'$  and  $b'$  were chosen =  $na'_b'$ ,  
 We calculated  $S_{ab} = (A_i = na_b) * (B_i = na_b)$ .  
 We calculated  $S_{a'b} = (A_i = na'_b) * (B_i = na'_b)$ .  
 We calculated  $S_{ab'} = (A_i = na_b') * (B_i = na_b')$ .  
 We calculated  $S_{a'b'} = (A_i = na'_b') * (B_i = na'_b')$ .  
 For block  $j$  of size  $d$  of  $S_{ab}$ , if there are two values of 1,  $X_{abj} = 1$ , if there is one value of 1,  $X_{abj} = 1$ , if there is one value of  $-1$   $X_{abj} = -1$ , if all values are 0,  $X_{abj} = -1$ .  
 For block  $j$  of size  $d$  of  $S_{a'b}$ , if there are two values of 1,  $X_{a'bj} = 1$ , if there is one value of 1,  $X_{a'bj} = 1$ , if there is one value of  $-1$ ,  $X_{a'bj} = -1$ , if all values are 0,  $X_{a'bj} = -1$ .  
 For block  $j$  of size  $d$  of  $S_{ab'}$ , if there are two values of 1,  $X_{ab'j} = 1$ , if there is one value of 1,  $X_{ab'j} = 1$ , if there is one value of  $-1$ ,  $X_{ab'j} = -1$ , if all values are 0,  $X_{ab'j} = -1$ .  
 For block  $j$  of size  $d$  of  $S_{a'b'}$ , if there are two values of 1,  $X_{a'b'j} = 1$ , if there is one value of 1,  $X_{a'b'j} = 1$ , if there is one value of  $-1$ ,  $X_{a'b'j} = -1$ , if all values are 0,  $X_{a'b'j} = -1$ .

We calculated the following:

$$C_{ab} = (\sum X_{abj}) / \text{length}(X_{abj})$$

$$C_{a'b} = (\sum X_{a'bj}) / \text{length}(X_{a'bj})$$

$$C_{ab'} = (\sum X_{ab'j}) / \text{length}(X_{ab'j})$$

$$C_{a'b'} = (\sum X_{a'b'j}) / \text{length}(X_{a'b'j})$$

$$C = C_{ab} - C_{a'b} + C_{ab'} + C_{a'b'}$$

### 10. Nonergodicity Check Methods

Nonergodicity means that measure-theoretic and frequency averages do not coincide. We claim that this is the main reason for the violation of the CHSH inequality. We have operated with the frequency averages.

$S_{ab} = \langle S_a, S_b \rangle$ , but the standard proofs of the Bell-type inequalities were done for measure-theoretic correlation  $S_{ab, \text{measure}} = \langle \langle S_a, S_b \rangle \rangle$ . In order to check if our data was ergodic, we continued with the first scheme.

#### 10.1. Scheme for Determining if Data Is Nonergodic

We started with one sequence  $A = (A_1, A_2, \dots, A_n)$ , where each  $d$ -vector  $A_j = (u_1, u_2, \dots, u_d)$  represents a dendrogram. We fixed  $a = [a_1 a_2]$  and defined

$$S_a = S_a(u) \tag{1}$$

We continued to calculate  $S_a$  as in the first method of Step 3 in Section 9. We stress that (1) is the condition of realism, and the value of hidden variable  $u$  determines the outcome of the observable  $S_a$ . Thus,

$$\langle S_a \rangle = \left( \sum_{j=1}^n S_a(A_j) \right) / N \tag{2}$$

is the frequency average. Next, we computed the probability distribution of hidden variables. We fixed one  $d$ -vector  $u$ , which is present in our dendrograms, and calculated the proportional number of occurrences of  $u$ , as follows.

$$p(u) = \frac{\text{number of occurrences of } u}{n} \tag{3}$$

Then

$$\ll Sa \gg = \sum_u Sa(u)p(u) \tag{4}$$

where the sum is with respect to all hidden variables  $u$ , which are present in our sequence of dendrograms. In the simplest situation,

$$\ll Sa \gg \neq \langle Sa \rangle$$

The deviation should be visible, not just a very small deviation.

### 10.2. Scheme for Determining if Correlations between Data Are Nonergodic

However, it can be that at the level of one sequence, the ergodicity holds, but it is violated for correlations (A. Khrennikov, one of the current authors, has a paper about this possibility and it seems that violation of CHSH can be of this type).

Then, we proceeded without the random choice and calculated for  $A = (A_1, \dots, A_n)$ ,  $B = (B_1, \dots, B_n)$   $Sa$  and  $Sb$  as in the first method of Step 3 in Section 9. Thus, our frequency correlation is as follows:

$$\langle Sa, Sb \rangle = \left( \sum_{j=1}^n Sa(A_j)Sb(B_j) \right) / N \tag{5}$$

Then, we completed this analysis by the measure-theoretic framework. We considered all possible pairs of hidden variables,  $d$ -vectors,  $u, v$ , which are present in the pairs  $A_j, B_j$ , and found the probability number of occurrences of  $u, v$ , as follows:

$$p(u, v) = \frac{\text{number of occurrences of } u, v}{n} \tag{6}$$

and

$$\ll Sa, Sb \gg = \sum_u Sa(u)Sb(v)p(u, v) \tag{7}$$

and for nonergodic correlations,

$$\ll Sa, Sb \gg \neq \langle Sa, Sb \rangle.$$

## 11. Results

We define the two distance metrics as follows:

Euclidean: *euclidean distance* =  $\sqrt{(x_i - x_j)^2}$ , where  $x \in \text{spatial data point}$ ,  $i, j \in 1, 2 \dots n$ .  $n = d$  block size.

Minkowski: *minkowski distance* =  $-(t_i - t_j)^2 + (x_i - x_j)^2$ , where  $x \in \text{spatial data point}$ ,  $i, j \in 1, 2 \dots n$ .  $n = d$  block size,  $t \in \text{temporal data point}$ ,  $i, j \in 1, 2 \dots n$ .  $n = d$  block size.

We analyzed the same datasets by applying the Section 9 scheme (first and second methods) to the calculation of dendrograms correlations for **three different “views”** of the datasets:

1. We constructed dendrograms with the Euclidean distance metric, which is shown in Section 1 of the results.
2. We constructed dendrograms with the Minkowski distance metric, which is shown in Section 2 of the results.
3. We Lorentz-transformed the datasets and then constructed dendrograms with the Minkowski distance metric.

### 11.1. CHSH Violations for 2-Slit Diffraction Experiment Data

As an example, we reproduced the Bell violations of correlations from a very classical double-slit diffraction experiment [73]. We refer to parts of the working flow process in Figure 1.

The original experiment used a CCD camera with a  $512 \times 512$  detector chip. The detection pattern in each frame in the experiment was represented as a binary matrix of  $512 \times 512$ , where values of 0 represent no detection of photons in the corresponding detector, while values of 1 represent the detection of photons in the corresponding detector. For each frame and its corresponding binary matrix, we found all column positions that had the value 1, multiplied these positions' values, and calculated the  $\log_{10}$  value of the outcome; this resulted in unique numbers representing each frame's detection pattern (Figure 1 steps A and B).

Thus, each frame was encoded with only one unique number  $A_f$  and time number  $T_f$ , in which  $f \in \text{frame } 1, 2, \dots, N, N = 4070$ . From each fixed number of  $d$  consecutive frames,  $d = 4, 5, \dots, 9$  frames, and the pairwise distances between each of the  $d$   $A_f$  (or  $d$  vectors  $(T_f A_f)$ ) were calculated (Figure 1 step C).

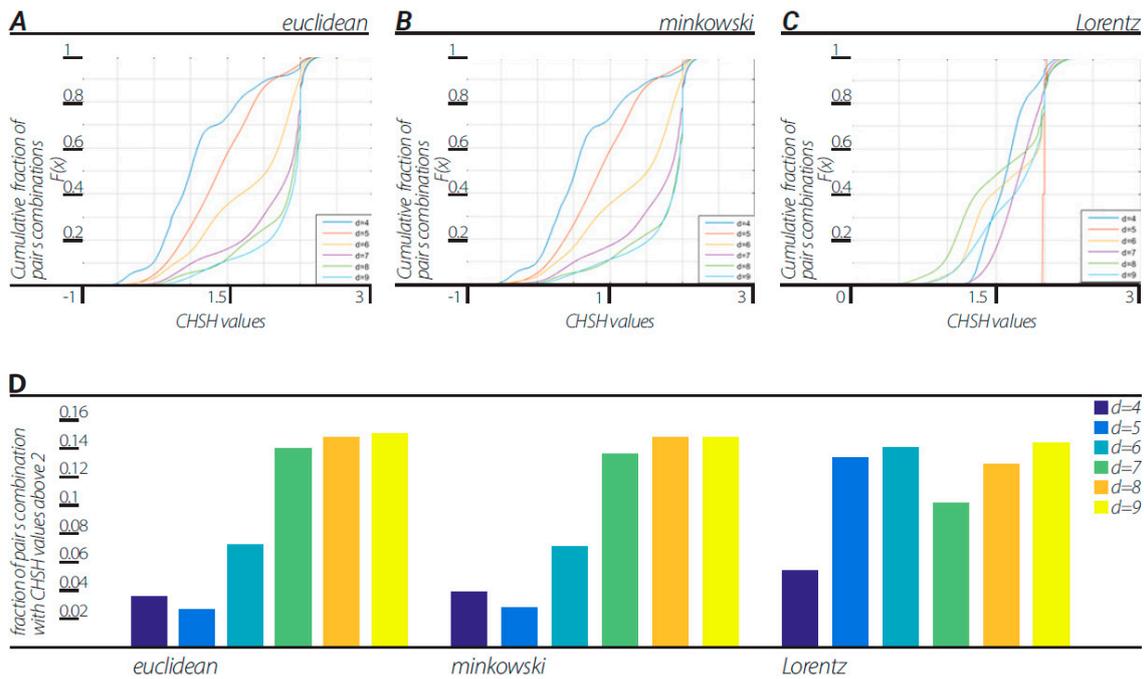
The pairwise distances were calculated according to the above three "views" of the datasets: first, by applying the Euclidean distance metric between each of the  $A_f$  values, secondly, by applying the Minkowski distance metric between each of the two vectors  $(T_f A_f)$ , and thirdly, by applying the Lorentz transformation to each vector  $(T_f A_f)$ , where the reference frame was 0.99 the speed of light, and then each Lorentzian-transformed vector  $[T_f A_f]_{\text{Lorentzian}}$  was compared to another Lorentzian-transformed vector by applying the Minkowski distance metric between such vectors.

Dendrograms were constructed from the pairwise distances (Figure 1 step D) by using a ward or a weighted linkage algorithm. The result was a series of consecutive dendrograms,  $D_1, D_2, \dots, D_n$ , where  $n = \lfloor N/d \rfloor$  with the block length/edge number in a single dendrogram  $d = 4, 5, \dots, 9$ . For each such  $d$ , we split the dendrogram series into two series (the first chronological half and the second chronological half) of dendrograms, each with a length of  $n/2$  (the above process is shown in Figure 1). One series was the observable Alice measures and the second series was the observable Bob measures. As in Step 2 of Section 9, we represented each dendrogram edge  $D_n$  with a natural number (as shown in Figure 1 step D). These features were sorted into ascending order according to the hierarchical structure of the dendrogram. We then evaluated the CHSH correlations in accordance with the first and second methods of Step 3.

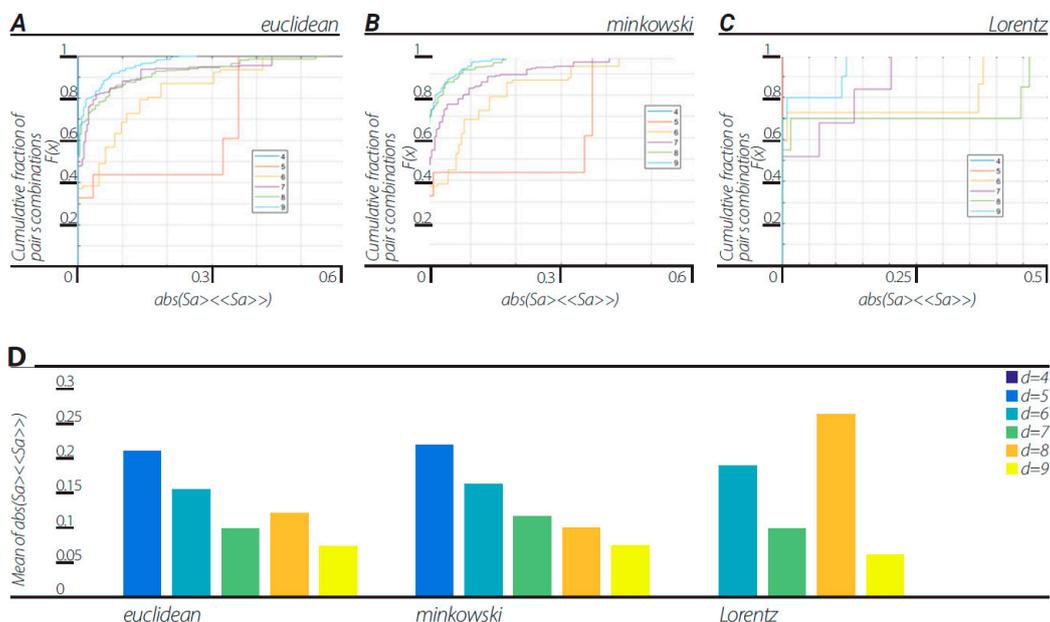
### 11.2. CHSH First Method

In line with the first method of Step 3 in Section 9 (numerical example in Illustration 1E), we calculated the CHSH values for each combination of features  $ab = [ [a \ b] [a' \ b'] [a' \ b] [a \ b'] ]$ , where  $a, a', b,$  and  $b'$  were chosen out of the number of features in the whole two-dendrogram series. Cumulative distribution functions (CDFs) of CHSH values of all pair combinations are shown in Figure 2. Dendrograms were constructed with block sizes of  $d = 4, 5, \dots, 9$  in three different ways corresponding to the three views described above. All views in all block sizes of  $d = 4, 5, \dots, 9$  showed a fraction of the pairs producing CHSH values above 2, resulting in a violation of CHSH inequality.

In order to verify that nonergodicity is the cause of the violation of the CHSH inequality, we tested the amount of ergodicity of the dendrographic data (Figure 3). With the same data that the CHSH values were calculated from, we tested the data for any pair of features  $a = [a_1 \ a_2]$ , as described in Section 10.1. The results show that the dendrographic data for dendrograms of block sizes  $d = 5, 6, \dots, 9$ , the mean ergodicity score  $(|\langle Sa \rangle - \langle\langle Sa \rangle\rangle|)$  shows considerable fractions of all possible  $a = [a_1 \ a_2]$  pairs with a non-zero value, indicating clear nonergodicity. Moreover, cross-correlating the fractions of CHSH values above 2 and the data ergodicity score in each dendrogram of block sizes  $d = 5, 6, \dots, 9$  for the Euclidean metric view, Minkowski view and Lorentzian view resulted in correlation coefficients of  $-0.9485, -0.9772,$  and  $-0.0025$  respectively.

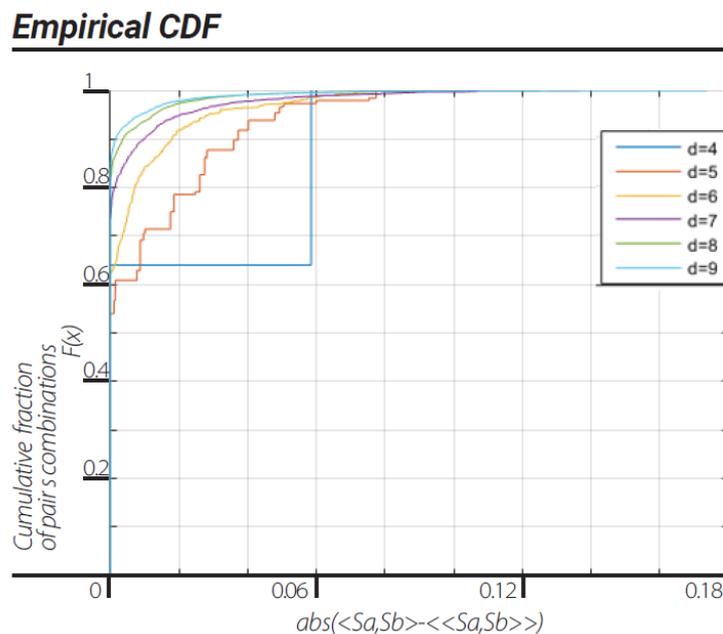


**Figure 2.** CHSH values computed according to the first method in Section 9 step 3. (A) Cdf’s of CHSH values with Euclidean view under “ward” linkage clustering algorithm. Each cdf corresponds to block size  $d = 4, 5, \dots, 9$ . (B) Cdf’s of CHSH values with Minkowski view under “ward” linkage clustering algorithm. Each cdf corresponds to block size  $d = 4, 5, \dots, 9$ . (C) Cdf’s of CHSH values with Lorentzian view under “ward” linkage clustering algorithm. Each cdf corresponds to block size  $d = 4, 5, \dots, 9$ . (D) fraction of CHSH with values above 2 in each view, Euclidean, Minkowski and Lorentzian, with block sizes  $d = 4, 5, \dots, 9$ .



**Figure 3.** Ergodicity score values were computed according to the scheme described in Section 10.1. (A) CDFs of ergodicity score values with Euclidean view under “ward” linkage clustering algorithm. Each CDF corresponds to a block size of  $d = 4, 5, \dots, 9$ . (B) CDFs of ergodicity score values with Minkowski view under “ward” linkage clustering algorithm. Each CDF corresponds to a block size of  $d = 4, 5, \dots, 9$ . (C) CDFs of ergodicity score values with Lorentzian view under “ward” linkage clustering algorithm. Each CDF corresponds to a block size of  $d = 4, 5, \dots, 9$ . (D) Mean of ergodicity score in each view, Euclidean, Minkowski, and Lorentzian, with block sizes of  $d = 4, 5, \dots, 9$ .

We note that for all views, the dendrogram data series with block sizes of  $d = 4$  show that the data is ergodic, thus the violation of CHSH inequality for a block size of  $d = 4$  can be a result of correlation nonergodicity. We carried on to investigate ergodicity in terms of two stationary ergodic series correlations in the Euclidean view. This ergodic test for correlations is described in Section 10.2. The results of this investigation clearly show all  $d = 4, 5, \dots, 9$  block size dendrograms are nonergodic, in terms of correlations, in the Euclidean view (Figure 4).



**Figure 4.** Ergodicity score values computed according to the scheme described in Section 10.2. CDFs of ergodicity score values with Euclidean view under “ward” linkage clustering algorithm. Each CDF corresponds to a block size of  $d = 4, 5, \dots, 9$ .

### 11.3. CHSH Second Method

In line with the second method of Step 3 in Section 9 (numerical example in Figure 1 step E), we calculated the CHSH values for each combination of features

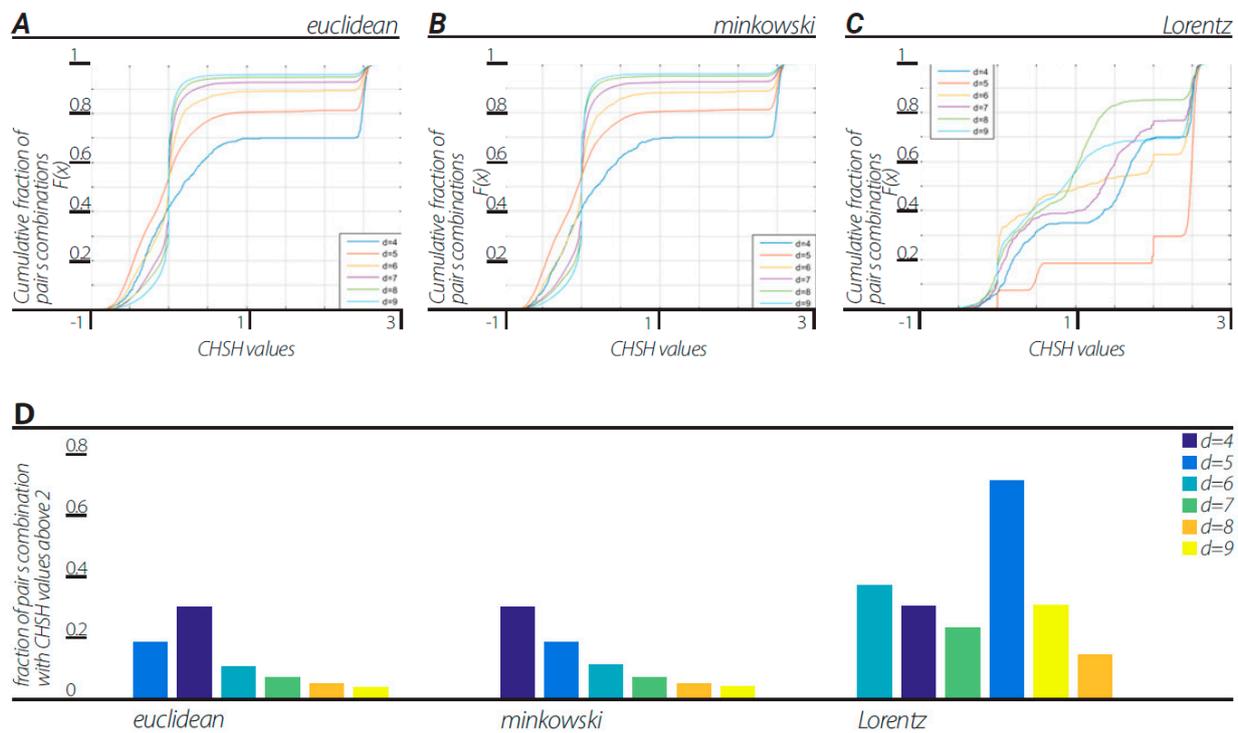
$$ab = [[a \ b] \ [a \ b'] \ [a' \ b] \ [a' \ b']]$$

where  $a, a', b,$  and  $b'$  were chosen out of the number of features in the whole two-dendrogram series. Cumulative distribution functions (CDFs) of the CHSH values of all pair combinations are shown in Figure 4. Dendrograms were constructed with block sizes of  $d = 4, 5, \dots, 9$  in three different ways corresponding to the three views described above. Again, as with the first method, all views in all block sizes of  $d = 4, 5, \dots, 9$  show a fraction of the pairs producing CHSH values above 2, resulting in a violation of CHSH inequality (Figure 5). Interestingly, the fraction of CHSH values above 2 in the Euclidean and Minkowski views show descending fraction values in respect to increasing  $d$  values of  $d = 4, 5, \dots, 9$  in contrast to the first method, which showed an ascending relation with increasing  $d$  values.

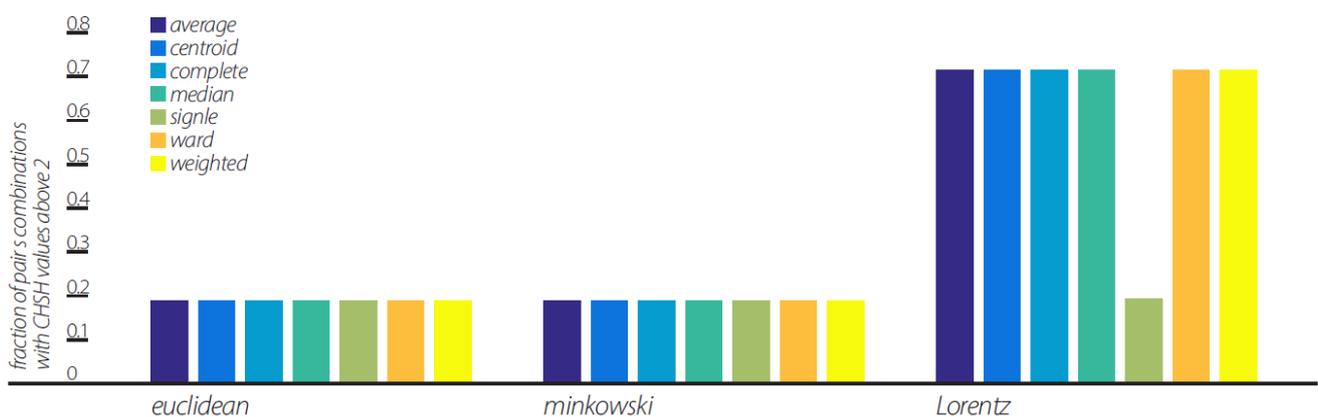
Please note that the Euclidean CDFs rise in correlation to the second ergodic test described in Figure 4.

We further studied the dependence of the three views’ violations of CHSH inequality on the linkage clustering algorithm (Figure 6). For that purpose, we used seven different algorithms, all with  $d = 5$ , which showed no changes in the Euclidean and Minkowski views as to the fraction of pairs violating CHSH inequality. The Lorentz view showed a

significant change in value in the “single” algorithm, resulting in fractions of pairs’ values comparable to the Euclidean and Minkowski views.



**Figure 5.** CHSH values computed according to the second method of Step 3 in Section 9. (A) CDFs of CHSH values with Euclidean view under “ward” linkage clustering algorithm. Each CDF corresponds to a block size of  $d = 4, 5, \dots, 9$ . (B) CDFs of CHSH values with Minkowski view under “ward” linkage clustering algorithm. Each CDF corresponds to a block size of  $d = 4, 5, \dots, 9$ . (C) CDFs of CHSH values with Lorentzian view under “ward” linkage clustering algorithm. Each CDF corresponds to a block size of  $d = 4, 5, \dots, 9$ . (D) Fraction of CHSH with values above a value of 2 in each view: Euclidean, Minkowski and Lorentzian, with block sizes  $d = 4, 5, \dots, 9$ .



**Figure 6.** The influence of seven linkage clustering algorithms on CHSH inequality violation. Fractions of CHSH with values above 2 in each view, Euclidean, Minkowski, and Lorentzian, with block sizes of  $d = 5$  and in each of the seven linkage clustering algorithms.

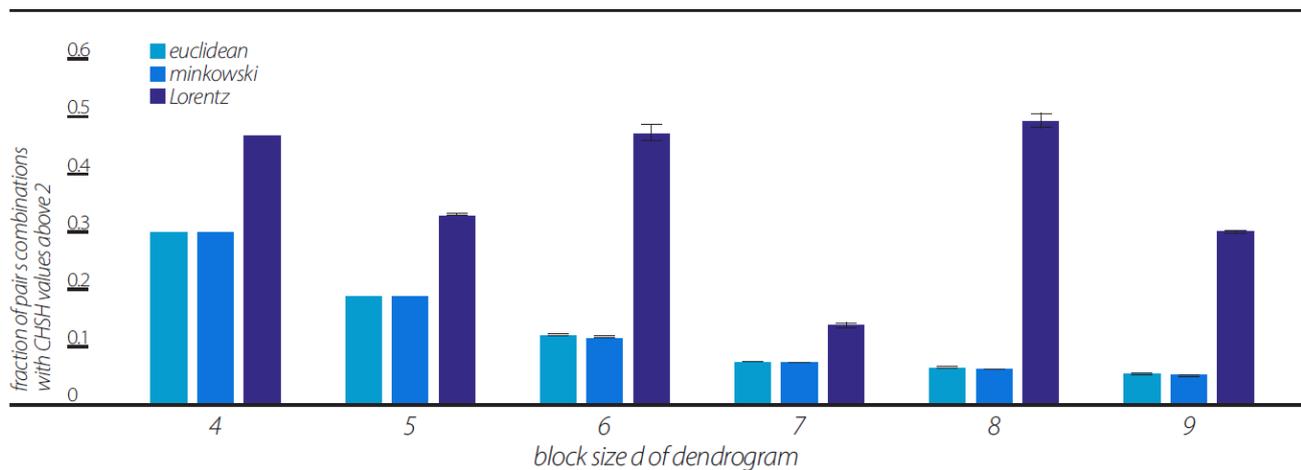
#### CHSH Second Method Random Data

We carried on to produce 10 random number sequences (comparable in size with the number of frames in the 2-slit diffraction experiment) in order to compare the analysis with the 2-slit diffraction experiment correlation analysis.

For this purpose, for each random number in one of the ten sequences, we used the following scheme:

First, we randomly chose a number between 1 and 7 which indicated the number of random numbers generated from the interval [1, 512]; this was in correspondence with the 2-slit diffraction experiment with 512 detectors.

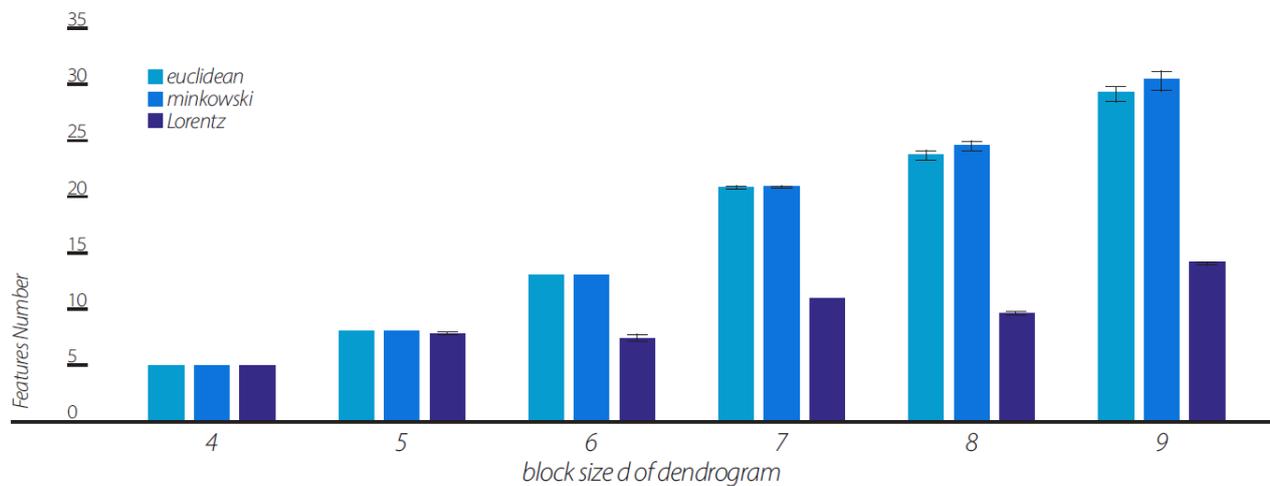
We multiplied these numbers and calculated the  $\log_{10}$  value of the outcome, which resulted in a unique number representing each “randomly generated frame detection pattern”. Each such frame was accompanied by a timestamp, which indicated the time of generation of the “randomly generated frame detection pattern”. We then analyzed, in each of the “views” as defined in Section 10, the fraction of  $ab$  pairs that violated CHSH inequality. The mean fraction values of  $ab$  pairs of ten random sequences that violate the CHSH inequality are shown in Figure 7.



**Figure 7.** Mean CHSH values computed from 10 sequences of random data according to the second method of Step 3 in Section 9. Fractions of CHSH with values above 2 in each view, Euclidean, Minkowski, and Lorentzian, with block sizes of  $d = 4, 5, \dots, 9$ , and in each with a “ward” linkage clustering algorithm.

#### 11.4. Lorentz-Transformed Data

As was clearly indicated by the above analyses, the Lorentz-transformed data resulted in higher fractions of pairs violating the CHSH inequality under most of the clustering algorithms (although, the “single” algorithm showed comparable CHSH fraction values, see Figure 6). One might wonder why. The explanation for this phenomenon is that most clustering algorithms, except for the “single” algorithm, produce dendrogram series with much a smaller phase space of the possible topology of dendrograms. This is clearly seen in Figure 8, which shows the number of features contained in the dendrogram series of the 10 random data sequences under each of the three views. The number of features in the Lorentzian view remains small for all  $d = 4, 5, \dots, 9$ .



**Figure 8.** Mean feature numbers computed from 10 sequences of dendrograms composed out of random data. Mean feature number of 10 sequences in each view, the Euclidean, Minkowski, and Lorentzian and in block sizes of  $d = 4, 5, \dots, 9$ .

## 12. Discussion on the Results of Numerical Data Analysis

To emphasize our basic idea, quantum-likeness is not the property of systems but of a time series of observations. In the same way as randomness, quantum-likeness can be established on the basis of tests for data. For this study, we considered only one (but very important) test, the CHSH test. The following are our main results:

- Violation of the CHSH inequality can be demonstrated by a time series generated by measurements on classical systems.
- The key point is a hierarchic dendrographic representation of data. It can be described by a local hidden variable model.
- CHSH violation is closely correlated with a violation of ergodicity.

Generally, nonergodicity is the fundamental property of a hierarchic dendrographic representation of statistical data.

This representation is based on the application of clustering algorithms. Although different algorithms can generate different dendrographic series from the same experimental data, general properties of representations do not differ; all algorithms show CHSH violation and nonergodicity of data or of correlations. Of course, the degrees of violation and nonergodicity vary with algorithms (see Figures 2–5 and Figure 7).

By taking into account the temporal component of the data series, we generated two-dimensional geometry endowed with the Minkowski pseudo-metric. We applied DH theory to this geometry. It was found that the basic properties of the Minkowski DH theory do not differ from its Euclidean version (Figures 2–5 and Figure 7). For the last step of our playing with algorithms, we performed a Lorentz transformation of data and then proceeded with dendrographic modeling. Again, the Lorentz-transformed data showed CHSH violations and nonergodicity of data (Figures 2–6), although not with comparable values to the Euclidean and Minkowski views. We show that these CHSH violations can become comparable with the violations under the Euclidean and Minkowski views under certain clustering algorithms (Figure 6, “single” clustering algorithm). We also show that the high fraction of pairs showing CHSH violations in the Lorentzian view is a consequence of the smaller phase space of the possible topology of dendrograms under most clustering algorithms, which result in a smaller number of features (Figure 8).

The two methods in Step 3 of Section 9 use the hierarchical information of the data differently. As a result, these two method adaptations to the formalization of the CHSH experiment, in terms of the hierarchical information, are different. This different adaptation leads to somewhat different results, where CHSH violations seem to be with the opposite

phase with a  $d$  block size. Although such correlations disappeared under Lorentz transformation, we note that, as shown in Figure 2, we used a particular linkage algorithm that behaves differently for Lorentz-transformed data, as can be seen in all figures. Interestingly, Figure 6 shows that the “single” linkage method results in the same performance for the Euclidean, Minkowski, and Lorentz “views”. Thus, we suspect that these correlations will be recovered under the “single” linkage algorithm. Although this still needs to be investigated, we note that both methods show CHSH violations.

Some more simulations should be considered with a  $d$  block size larger than 9 in order to again verify the transition from quantum to classical correlations already discussed in the previous study [1].

### 13. Concluding Remarks

In this study, DH theory [1] was structured in the ontic–epistemic (implicate–explicate order) framework. The experimental data was described by dendrograms generated by clustering algorithms—the epistemic counterpart of our theory. In the limit, dendrograms generated the infinite  $p$ -adic tree—the mathematical model of the ontic universe.

DH theory is realistic, not only because the epistemic theory of observations has an ontic background but also because, even at the epistemic level, it is possible to introduce hidden variables. Ontic realism is very exotic; *the universe is zero-dimensional, totally disconnected, and disordered*. This universe is bounded, but with respect to the very special distance, the  $p$ -adic ultrametric. However, maybe it is even more surprising that the local realistic description can be used eventually at the epistemic level, with hierarchic hidden variables and observables as their functions.

DH theory provides a novel viewpoint on classical–quantum interrelations. First, we stress that the ontic  $p$ -adic world is classical. In the epistemic dendrographic world, simpler dendrograms have higher degrees of quantum-likeness. For the moment, the latter is characterized with just one test of quantumness—the CHSH test. Its violation is a consequence of the nonergodicity of empirical data. However, it is crucial that this data is represented dendrographically. This representation gives us a local hidden variable model that generates correlations that violate the CHSH inequation (cf. [67–69,74,75]). The violation of ergodicity is closely coupled with the interrelation of the Kolmogorov measure-theoretic and the von Mises frequency approaches to probability (see [68] and references therein).

Finally, we noticed that the size of the baselevel of a dendrogram can be considered as a system’s dimension. It is interesting that we are able to violate the CHSH inequality only for  $d > 3$ .

**Author Contributions:** Conceptualization, methodology, software, validation, formal analysis investigation, resources, writing—original draft preparation, writing—review and editing, project administration, funding acquisition, O.S., F.B. and A.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** No experimental data was collected.

**Acknowledgments:** Igor Volovich, Vladimir Anashin, Evgenii Zelenov, Varda and Boaz Dotan, Masha Shor.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A. Ultrametric Spaces, Trees, $p$ -Adic Numbers I

In this appendix, we follow the non-expert friendly presentation from paper [76]. In some applications, the point structure of a set  $X$  and the properties of a metric  $\rho$  may

essentially differ from the Euclidean case. We are interested in metric spaces  $X$  where, instead of the standard triangle inequality, the strong triangle inequality,

$$\rho(x, y) \leq \max[\rho(x, z), \rho(z, y)] \quad (\text{A1})$$

is valid. Such a metric is called an ultrametric, and such metric spaces are called ultrametric spaces. The strong triangle inequality can be stated geometrically: all triangles are isosceles.

Let us discuss the main properties of ultrametric space  $X$ . We set

$$B_r(a) = \{x \in X: \rho(x - a) < r\}, B_r \sim (a) = \{x \in X: \rho(x - a) \leq r\}, r > 0, a \in X \quad (\text{A2})$$

These are balls of the radius  $r$  with the center at the point  $a$ . Our standard intuition tells us that  $B_r(a)$  is a closed ball, but not open, and  $B_r \sim (a)$  is an open ball, but not closed. However, it is not valid for ultrametric spaces.

In an ultrametric space, each ball in  $X$  is open and closed at the same time. Each point of a ball may serve as a center. A ball may have infinite radii.

Let  $U$  and  $V$  be two balls in ultrametric space  $X$ . Thus, there are only two possibilities: (1) the balls are ordered by inclusion (i.e.,  $U \subset V$  or  $V \subset U$ ); (2) the balls are disjointed.

Thus, if two balls have a common point, then one has to be a part of another.

The symbol  $S_r(a)$  denotes the sphere  $\{x \in X: \rho(x, a) = r\}$  of the radius  $r > 0$  with the center at  $a$ . There is also a large deviation from the Euclidean case; the sphere  $S_r(a)$  is not a boundary of  $B_r(a)$  or  $B_r \sim (a)$ .

Consider the following class of ultrametric spaces  $(X, \rho)$ . Every point  $x$  has an infinite number of coordinates  $x = (a_0, a_1, \dots, a_n)$ . Each coordinate yields a finite number of values  $a \in \{0, \dots, m - 1\}$ , where  $p > 1$  is a natural number. We denote the space of sequences (1) by the symbol  $X = \mathbf{Zp}$ . The standard ultrametric is introduced in this set in the following way.

Let  $x = (a_0, a_1, a_2, \dots, a_n)$ ,  $y = (b_0, b_1, b_2, \dots, b_n) \in \mathbf{Zp}$ . We set

$$r_p(x, y) = 1/m^k \quad \text{if } a_j = b_j, j = 0, 1, \dots, k - 1, \text{ and } a_k \neq b_k \quad (\text{A3})$$

This is a metric and even an ultrametric. To find the distance  $r_p(x, y)$  between two strings of digits,  $x$  and  $y$ , we have to find the first position  $k$  at which the strings have different digits. The space  $X = \mathbf{Zm}$  coincides with the unit ball centered at zero,  $X = B_1(0)$ ; this space is compact. Geometrically, it can be represented by the tree (see Figure 8 for the 2-adic tree representing  $\mathbf{Z2}$ ). Here, one vertex, the root labeled as  $R$ , is incidental for two edges and other vertices are incidental for three edges. We noticed that it is convenient to consider this tree as the directed graph; for each vertex  $I$  different from  $R$ , one edge comes from the branch starting at  $R$ , the "input edge", and two edges go out from  $I$ , the "output edges". These two edges (or vertices at their ends) are labeled by  $a = 0, 1$ . In Figure A1, the order of labeling of the output edges is based on the embedding of the tree in the plane, the upper output edges are labeled by 0 and the lower by 1. This leads to the concrete numerical representation of this tree. However, the rule used for the labeling of edges is not obligatory; for each vertex  $I$ , we can assign 0/1 to each of the output edges in an arbitrary way and obtain another numerical representation of this tree.

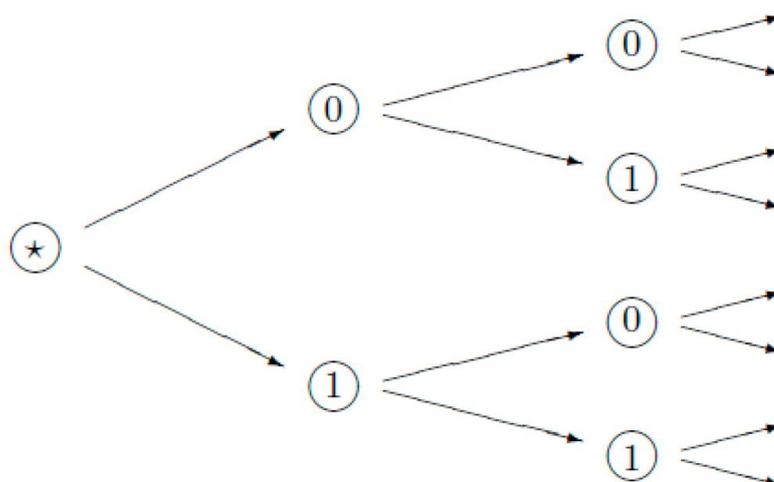


Figure A1. 2-adic hierarchic tree.

## References

- Shor, O.; Benninger, F.; Khrennikov, A. Representation of the Universe as a Dendrographic Hologram Endowed with Relational Interpretation. *Entropy* **2021**, *23*, 584. [[CrossRef](#)] [[PubMed](#)]
- Atmanspacher, H. Determinism is ontic, determinability is epistemic. In *Between Chance and Choice: Interdisciplinary Perspectives on Determinism*; Atmanspacher, H., Bishop, R.C., Eds.; Imprint Academic: Thorverton, UK, 2002; pp. 49–74.
- Atmanspacher, H.; Primas, H. Epistemic and ontic quantum realities. In *Time, Quantum and Information*; Castell, L., Ischebeck, O., Eds.; Springer: Berlin/Heidelberg, Germany, 2003.
- Hertz, H. *The Principles of Mechanics: Presented in a New Form*; Macmillan: London, UK, 1899.
- Boltzmann, L. Über die Frage nach der objektiven Existenz der Vorgänge in der unbelebten Natur. In *Populre Schriften*; Barth, J.A., Ed.; Springer: Leipzig, Germany, 1905.
- Khrennikov, A. Quantum epistemology from subquantum ontology: Quantum mechanics from theory of classical random fields. *Ann. Phys.* **2017**, *377*, 147–163. [[CrossRef](#)]
- Khrennikov, A. Hertz's Viewpoint on Quantum Theory. *Act. Nerv. Super.* **2019**, *61*, 24–30. [[CrossRef](#)]
- Brenner, J.E.; Igamberdiev, A. *Philosophy in Reality: A New Book of Changes*; Springer: Berlin/Heidelberg, Germany, 2021.
- Bohm, D. *Wholeness and the Implicate Order*; Routledge: London, UK, 1980.
- Khrennikov, A. *Beyond Quantum*; Pan Stanford Publishing: Singapore, 2014.
- Jaeger, G. *Quantum Objects: Non-Local Correlation, Causality and Objective Indefiniteness in the Quantum World*; Springer: Berlin/Heidelberg, Germany, 2013.
- Boughn, S. Making Sense of Bell's Theorem and Quantum Nonlocality. *Found. Phys.* **2017**, *47*, 640–657. [[CrossRef](#)]
- Fuchs, C.A. On participatory realism. In *Information and Interaction*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 113–134.
- Smolin, L. *Einstein's Unfinished Revolution: The Search for What Lies beyond the Quantum*; Penguin Press: London, UK, 2019.
- Cohen, E.; Cortés, M.; Elitzur, A.C.; Smolin, L. Realism and Causality II: Retrocausality in Energetic Causal Sets. *Phys. Rev. D* **2020**, *102*, 124028. [[CrossRef](#)]
- Cohen, E.; Cortés, M.; Elitzur, A.; Smolin, L.; Gan, R. Realism and causality I: Pilot wave and retrocausal models as possible facilitators. *Phys. Rev. D* **2020**, *102*. [[CrossRef](#)]
- Kupczynski, M. Closing the Door on Quantum Nonlocality. *Entropy* **2018**, *20*, 877. [[CrossRef](#)]
- Sciarretta, A. A local-realistic quantum mechanical model of spin and spin entanglement. *Int. J. Quant. Inf.* **2021**, *19*, 2150006. [[CrossRef](#)]
- Volovich, I. p-adic string. *Class. Quantum Gravity* **1987**, *4*, 83–87. [[CrossRef](#)]
- Vladimirov, V.S.; Volovich, I.V.; Zelenov, E.I. *p-Adic Analysis and Mathematical Physics*; World Scientific: Singapore, 1994.
- Aref'eva, I.Y.; Dragovich, B.; Frampton, P.H.; Volovich, I.V. The wave function of the Universe and p-adic gravity. *Int. J. Mod. Phys. A* **1991**, *6*, 4341–4358. [[CrossRef](#)]
- Freund, P.G.O.; Witten, E. Adelic string amplitudes. *Phys. Lett. B* **1987**, *199*, 191–194. [[CrossRef](#)]
- Khrennikov, A. p-adic quantum mechanics with p-adic valued functions. *J. Math. Phys.* **1991**, *32*, 932–937. [[CrossRef](#)]
- Khrennikov, A. *p-Adic Valued Distributions in Mathematical Physics*; Springer: Berlin/Heidelberg, Germany, 1994.
- Parisi, G.; Sourlas, N. P-adic numbers and replica symmetry breaking. *Eur. Phys. J. B Condens. Matter Complex. Syst.* **2000**, *14*, 535–542. [[CrossRef](#)]
- Volovich, I.V. Number theory as the ultimate physical theory. *P-Adic Numbers Ultrametric Anal. Appl.* **2010**, *2*, 77–87. [[CrossRef](#)]

27. Dragovich, B.; Khrennikov, A.Y.; Kozyrev, S.V.; Volovich, I.V.; Zelenov, E.I. P-adic mathematical physics: The first 30 years. *p-Adic Numbers Ultrametric Anal. Appl.* **2017**, *9*, 87–121. [CrossRef]
28. Anashin, V. Discreteness causes waves. *Facta Univ. Ser. Phys. Chem. Technol.* **2016**, *14*, 143–196. [CrossRef]
29. Bohr, N. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **1935**, *48*, 696–702. [CrossRef]
30. Bohr, N. The quantum of action and the description of nature. In *Foundations of Quantum Physics I (1926–1932)*, Niels Bohr Collected Works; Kalckar, J., Ed.; Elsevier B.V.: Amsterdam, The Netherlands, 1985; Volume 6, pp. 201–217.
31. Plotnitsky, A. *Epistemology and Probability: Bohr, Heisenberg, Schrödinger and the Nature of Quantum-Theoretical Thinking*; Springer: Berlin/Heidelberg, Germany, 2009.
32. Plotnitsky, A. *Niels Bohr and Complementarity: An Introduction*; Springer: Berlin/Heidelberg, Germany, 2012.
33. Plotnitsky, A. “The Unavoidable Interaction Between the Object and the Measuring Instruments:” Reality, Probability, and Nonlocality in Quantum Physics. *Found. Phys.* **2020**, *50*, 1824–1858. [CrossRef]
34. Buonomano, V. *Quantum Uncertainties, Recent and Future Experiments and Interpretations*; Honig, W.M., Kraft, D.W., Panarella, E., Eds.; Plenum Press: New York, NY, USA, 1986.
35. Gallavotti, G. Ergodicity, ensembles, irreversibility in Boltzmann and beyond. *J. Stat. Phys.* **1995**, *78*, 1571–1589. [CrossRef]
36. Khrennikov, Y. p-adic probability predictions of correlations between particles in the two slit and neutron interferometry experiments. *II Nuovo Cimento B* **1998**, *113*, 751–760.
37. Khrennikov, A. Buonomano against Bell: Nonergodicity or nonlocality? *Int. J. Quant. Inf.* **2017**, *15*, 1740010. [CrossRef]
38. Hnilo, A. Beyond loophole-free experiments: A search for nonergodicity. In *Quantum Worlds: Perspectives on the Ontology of Quantum Mechanics*; Lombardi, O., Fortin, S., López, C., Holik, F., Eds.; Cambridge University Press: Cambridge, UK, 2019; pp. 245–266.
39. Barbour, J.; Smolin, L. Extremal variety as the foundation of a cosmological quantum theory. *arXiv* **1992**, arXiv:hep-th/9203041.
40. Barbour, J. Shape dynamics. An introduction. In *Quantum Field Theory and Gravity*; Springer: Basel, Switzerland, 2012.
41. Rovelli, C. Relational quantum mechanics. *Int. J. Theor. Phys.* **1996**, *35*, 1637–1678. [CrossRef]
42. Smolin, L. Quantum Mechanics and the Principle of Maximal Variety. *Found. Phys.* **2016**, *46*, 736–758. [CrossRef]
43. Cortês, M.; Smolin, L. The universe as a process of unique events. *Phys. Rev. D* **2014**, *90*, 1–30. [CrossRef]
44. Plotnitsky, A. Spooky predictions at a distance: Reality, complementarity and contextuality in quantum theory. *Philos. Trans. R. Soc. A* **2019**, *377*, 20190089. [CrossRef]
45. Fuchs, C.-A.; Mermin, N.-D.; Schack, R. An Introduction to QBism with an Application to the Locality of Quantum Mechanics. *Am. J. Phys.* **2014**, *82*, 749–761. [CrossRef]
46. Murtagh, F.; Contreras, P. Algorithms for hierarchical clustering: An Overview. *Data Min. Knowl. Discov.* **2011**, *2*, 86–97. [CrossRef]
47. Murtagh, F. Ultrametric and Generalized Ultrametric in Computational Logic and in Data Analysis. *Horiz. Comput. Sci. Res.* **2011**, *2*.
48. Murtagh, F.; Contreras, P.; Downs, G. Hierarchical clustering of massive, high dimensional data sets by exploiting ultrametric embedding. *SIAM J. Sci. Comput.* **2008**, *30*, 707–730. [CrossRef]
49. Available online: [https://www.mathworks.com/help/stats/linkage.html#mw\\_08b425f7-fc8c-480a-b618-f768817e8e11](https://www.mathworks.com/help/stats/linkage.html#mw_08b425f7-fc8c-480a-b618-f768817e8e11) (accessed on 25 July 2021).
50. Bell, J.S. On the Einstein-Podolsky-Rosen paradox. *Physics* **1964**, *1*, 195–200. [CrossRef]
51. Bell, J.S. *Speakable and Unsayable in Quantum Mechanics*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2004.
52. Bell, J.S. On the problem of hidden variables in quantum theory. *Rev. Mod. Phys.* **1966**, *38*, 450–481. [CrossRef]
53. 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. [CrossRef]
54. Aspect, A. Experimental tests of Bell’s inequalities in atomic physics. In *Atomic Physics 8, Proceedings of the Eighth International Conference on Atomic Physics, Gothenburg, Sweden, 2–6 August 1982*; Lindgren, I., Rosen, A., Svanberg, S., Eds.; Plenum Press: New York, NY, USA, 1983; pp. 103–128.
55. Aspect, A. Bell’s Theorem: The naive view of an experimentalist. In *Quantum [Un] Speakables*; Springer: Berlin/Heidelberg, Germany, 2002; pp. 119–153.
56. Khrennikov, A. *Quantum Theory: Reconsideration of Foundations*; Växjö Univ. Press: Växjö, Sweden, 2002.
57. Khrennikov, A. *Foundations of Probability and Physics-2*; Växjö Univ. Press: Växjö, Sweden, 2003.
58. Adenier, G.; Khrennikov, A.; Nieuwenhuizen, T.M. *Quantum Theory: Reconsideration of Foundations-3*; AIP: Melville, NY, USA, 2006; p. 810.
59. Adenier, G.; Fuchs, C.; Khrennikov, A. *Foundations of Probability and Physics-3*; AIP: Melville, NY, USA, 2007; p. 889.
60. Adenier, G.; Khrennikov, A.; Lahti, P.Y.; Manko, V.I.; Nieuwenhuizen, T.M. *Quantum Theory: Reconsideration of Foundations—4*; AIP: Melville, NY, USA, 2008; p. 962.
61. Aspect, A.; Dalibard, J.; Roger, G. Experimental test of Bell’s Inequalities using time-varying analyzers. *Phys. Rev. Lett.* **1982**, *49*, 1804–1807. [CrossRef]
62. Weihs, G.; Jennewein, T.; Simon, C.; Weinfurter, H.; Zeilinger, A. Violation of Bell’s Inequality under Strict Einstein Locality Conditions. *Phys. Rev. Lett.* **1998**, *81*, 5039–5043. [CrossRef]

63. Hensen, B.; Bernien, H.; Dreau, A.E.; Reiserer, A.; Kalb, N.; Blok, M.S.; Ruitenberg, J.; Vermeulen, R.F.; Schouten, R.N.; Abellan, C.; et al. Experimental loophole-free violation of a Bell inequality using entangled Electron spins separated by 1.3 km. *Nature* **2015**, *526*, 682. [[CrossRef](#)] [[PubMed](#)]
64. Giustina, M.; Versteegh, M.A.; Wengerowsky, S.H.; Steiner, J.; Hochrainer, A.; Phelan, K.; Steinlechner, F.; Kofler, J.; Larsson, J.Å.; Abellan, C.; et al. A significant-loophole-free test of Bell's theorem with entangled photons. *Phys. Rev. Lett.* **2015**, *115*, 250401. [[CrossRef](#)]
65. Shalm, L.K.; Meyer-Scott, E.; Christensen, B.G.; Bierhorst, P.; Wayne, M.A.; Stevens, M.J.; Gerrits, T.; Glancy, S.; Hamel, D.R.; Allman, M.S.; et al. A strong loophole-free test of local realism. *Phys. Rev. Lett.* **2015**, *115*, 2504. [[CrossRef](#)] [[PubMed](#)]
66. De Broglie, L. *The Current Interpretation of Wave Mechanics: A Critical Study*; Elsevier: Amsterdam, The Netherlands, 1964.
67. Khrennikov, A.Y. *Interpretations of Probability*, 2nd ed.; Walter De Gruyter: Berlin, Germany, 2009.
68. Khrennikov, A. Einstein and Bell, von Mises and Kolmogorov: Reality and locality, frequency and probability. *Found. Phys.* **2002**, *32*, 1159–1174. [[CrossRef](#)]
69. Khrennikov, A. *Contextual Approach to Quantum Formalism*; Springer: Berlin/Heidelberg, Germany, 2009.
70. von Mises, R. *The Mathematical Theory of Probability and Statistics*; Academic Press: London, UK, 1964.
71. Kolmogoroff, A.-N. *Grundbegriffe der Wahrscheinlichkeitsrechnung*; Springer: Berlin/Heidelberg, Germany, 1933.
72. Khrennikov, A. Get rid of nonlocality from quantum physics. *Entropy* **2019**, *21*, 806. [[CrossRef](#)] [[PubMed](#)]
73. Aspden, R.S.; Padgett, M.J.; Spalding, G. Video recording true single-photon double-slit interference Video recording true single-photon double-slit interference. *Am. J. Phys.* **2018**, *671*. [[CrossRef](#)]
74. Hess, K.; Philipp, W. Bell's theorem: Critique of proofs with and without inequalities. *AIP Conf. Proc.* **2005**, *750*, 150.
75. Hess, K. *Einstein Was Right!* Jenny Stanford Publishing: Singapore, 2014.
76. Khrennikov, A.; Oleschko, K.; López, D.J.C. Modeling fluid's dynamics with master equations in ultrametric spaces representing the treelike structure of capillary networks. *Entropy* **2016**, *18*, 249. [[CrossRef](#)]