Systems of random variables and the free will theorem

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

[1]*Purdue University, West Lafayette, Indiana 47905, USA*
[2]*University of Turku, Turku, FI-20014, Finland*

The title refers to the free will theorem by Conway and Kochen whose flashy formulation is as follows: If experimenters possess free will, then so do particles. In more modest terms, the theorem says that individual pairs of spacelike separated particles cannot be described by deterministic systems provided their mixture is the same for all choices of measurement settings. We reformulate and generalize the free will theorem in terms of systems of random variables and show that the proof is based on two observations: (1) Some compound systems are contextual (nonlocal), and (2) any deterministic system with spacelike separated components is nonsignaling. The contradiction between the two is obtained by showing that a mixture of nonsignaling deterministic systems, if they exist, is always noncontextual. The "experimenters' free will" (independence) assumption is not needed for the proof: It is made redundant by the assumption (1) above, critical for the proof. We next argue that the reason why an individual pair of particles is not described by a deterministic system is more elementary than in the free will theorem. A system, contextual or not and deterministic or not, includes several choices of settings, each of which can be factually used without changing the system. An individual pair of particles can only afford a single realization of random variables for a single choice of settings. With this conceptualization, the "free will of experimenters" cannot be even meaningfully formulated, and the choice between the determinism and "free will of particles" becomes arbitrary and inconsequential.

I. INTRODUCTION

In this paper the issues related to the free will theorem (FWT) [1,2] are discussed in terms of random variables. Conway and Kochen [2,3] emphasize that their theorem does not use probabilistic notions. This seems to plunge our paper in a controversy from the outset [3–5]: Our analysis of the FWT would be suspect if we used conceptual means that are not acceptable in the original formulation of the theorem. This is not the case, however. We use the language of random variables to describe quantum experiments, involving large numbers of replications with multitudes of particles. On the level of individual particles (more specifically, individual pairs of entangled particles), our focus, the same as in the original FWT, is exclusively on whether they can be described as systems of deterministic outcomes (which are, of course, a special case of random variables). Probabilistic description of quantum experiments is hardly controversial, and we have a good demonstration of the benefits it offers for the FWT in Ref. [6]. Moreover, it is unavoidable. It is known [5,6] that the

FWT can be found or extracted from much earlier work than [1,2], with the Kochen-Specker system used by Conway and Kochen being replaced with other contextual systems (in fact, any contextual system, as we will see later). The contextuality of many of these systems is more saliently probabilisitc than that of the Kochen-Specker system. This prominently includes the EPR-Bohm system, with a variant of the FWT being already seen in Bell's pioneering work [7].

Our reformulations lead us to critically reexamine and modify the FWT, although not invalidate it. In particular, we show that the assumption of the free will of experimenters (or independence assumption, as many authors prefer to call it [6,8–12]) is not needed in the FWT. This assumption is only needed to ensure that experimental observations correctly identify the system experimented on as contextual. It is therefore made unnecessary by another assumption, critical for the FWT and underivable from the independence assumption—that a contextual system with certain properties exists. Furthermore, we argue that while the question of whether individual particles can be described by deterministic systems is indeed to be answered negatively, and while the FWT is indeed one way of demonstrating this, there is a more elementary reason for this negative answer: The notion of a system, deterministic or not, is not applicable to an individual pair of particles to begin with. The latter is a realization of random variables for a single choice of settings, whereas the notion of a system involves several mutually exclusive settings, each of which can be factually and repeatedly used.

Our analysis is based on the contextuality-by-default (CbD) theory (e.g., Refs. [13–15]), but its utilization in this

*To whom correspondence should be addressed: ehtibar@purdue.edu
†jvk@iki.fi

paper is confined to only two basic principles. The first one is that each random variable is identified not only by the property it measures but also by the context (settings) in which it measures this property. The second principle is that no two random variables recorded in different, mutually exclusive contexts possess a joint distribution. Moreover, these principles are only applied to a special subset of systems of random variables, the compound, "Alice-Bob"-type systems with spacelike separation. These systems are nonsignaling, and this makes it unnecessary for us to invoke most of the content of CbD. Because of this the reader need not be familiar with CbD to understand this paper.

However, a brief comment may be needed on the two principles just mentioned. The double indexation of random variables means that if Alice chooses a setting x and Bob chooses a setting y, their measurement outcomes (random variables) are represented as, respectively, $A^{x,y}$ and $B^{x,y}$. And if Bob changes his setting to y' while Alice maintains her setting, her measurement is represented by another random variable, $A^{x,y'}$. A reader might erroneously interpret this as indicating that Bob somehow influences Alice's measurements despite their spacelike separation (a "spooky action at a distance"). This is not the case. The distribution of $A^{x,y'}$ is the same as that of $A^{x,y}$, so Bob transfers no information to Alice. There is no "action." The difference between $A^{x,y}$ and $A^{x,y'}$ simply reflects the relational nature of random variables in classical probability theory. A random variable is a measurable function on a probability space, and any variable defined on the same space is jointly distributed with it: Their observed realizations are paired. Therefore, if $A^{x,y}$ and $A^{x,y'}$ were the same random variable, they would be jointly distributed. But this would mean that realizations of $A^{x,y}$ and realizations of $A^{x,y'}$ co-occur (and are equal), while in reality they occur in mutually exclusive contexts.[1]

The scheme of the paper is as follows. In the next section we introduce formal notions and definitions related to systems of random variables. In Secs. III and IV we present the FWT in the language of such systems. In Sec. V we show that a systematic use of the language of random variables makes the FWT unnecessary (though not wrong): The experimenter' free will (independence) becomes unformulable, and the choice between the determinism and free will of particles becomes arbitrary and inconsequential. The concluding section provides a brief summary.

II. PRELIMINARIES

A *compound system* of random variables is an indexed set of random variables,

$$\mathcal{R} = \{(A^{x,y}, B^{x,y}) : (x, y) \in C\}, \qquad (1)$$

where x is a *property* measured by Alice, y is a *property* measured by Bob, (x, y) is the *context* in which the measurements are made, and C is a set of all possible contexts. Every random variable therefore is identified by the property it measures and the context in which it measures it. To simplify discussion, we will assume that all random variables have a finite number of values. The properties x, y are also referred to as *settings*, although there is the obvious semantic difference: A setting x designates the decision and arrangements made to measure property x.

Alice and Bob are always assumed to be *spacelike separated*. Because of this, by special relativity, the system is *nonsignaling*: The distributions of the variables are context-independent,

$$A^{x,y} \overset{\text{dist}}{=} A^{x,y'}, \qquad (2)$$

for any x, y, y' such that $(x, y), (x, y') \in C$. The symbol $\overset{\text{dist}}{=}$ indicates equality of distributions. Analogously,

$$B^{x,y} \overset{\text{dist}}{=} B^{x',y}, \qquad (3)$$

for any y, x, x' such that $(x, y), (x', y) \in C$.

One prominent example of a compound system is the EPR-Bohm system [7,17],

$$
\begin{array}{|c|c|c|c|}
\hline
\multicolumn{4}{c}{\mathcal{R}_{\text{EPRB}}} \\
\hline
A^{1,1} & & B^{1,1} & \\
\hline
A^{1,2} & & & B^{1,2} \\
\hline
& A^{2,1} & B^{2,1} & \\
\hline
& A^{2,2} & & B^{2,2} \\
\hline
\end{array}, \qquad (4)
$$

with $C = \{x = 1, x = 2\} \times \{y = 1, y = 2\}$. Another prominent example is the compound version of the Kochen-Specker-Peres system [18,19],

$$
\begin{array}{|c|c|c|c|c|c|c|c|}
\hline
\multicolumn{8}{c}{\mathcal{R}_{\text{KSP}}} \\
\hline
A^{1,1} & & \cdots & & B^{1,1} & & \cdots & \\
\hline
A^{1,2} & & \cdots & & & B^{1,2} & \cdots & \\
\hline
\vdots & & \vdots\vdots\vdots & & & \vdots & \vdots\vdots\vdots & \\
\hline
A^{1,33} & & & & & & \cdots & B^{1,33} \\
\hline
\vdots & \vdots & \vdots\vdots\vdots & \vdots & \vdots & \vdots & \vdots\vdots\vdots & \vdots \\
\hline
& & \cdots & A^{40,1} & B^{40,1} & & \cdots & \\
\hline
& & \cdots & A^{40,2} & & B^{40,2} & \cdots & \\
\hline
& & \vdots\vdots\vdots & \vdots & & \vdots & \vdots\vdots\vdots & \\
\hline
& & & A^{40,33} & & & \cdots & B^{40,33} \\
\hline
\end{array},
$$
$$(5)$$

with $C = \{x = 1, \ldots, x = 40\} \times \{y = 1, \ldots, y = 33\}$. In $\mathcal{R}_{\text{EPRB}}$, the x values and y values enumerate choices of axes by Alice and Bob, and the random variables are 0/1 (say, spin values in spin-$\frac{1}{2}$ particles). In $\mathcal{R}_{\text{KSP}}$, the y values represent 33 special axes in Peres's proof of the Kochen-Specker theorem [19], and the x values encode 40 Peres's triples formed using these 33 axes; the A variables have values 011,101,110, and the B variables are 0/1.

Any two random variables recorded in the same context, and referred to as an *AB pair*, are *jointly distributed*: This means that $\Pr[A^{x,y} = a, B^{x,y} = b]$ is well defined for $(x, y) \in C$. However, two random variables from different contexts are *stochastically unrelated*, i.e., have no joint distribution; i.e., if $(x, y) \neq (x', y')$, then the event conjunctions

---

[1]Without elaborating (see Refs. [15,16] for detailed argument), another way of understanding the contextual labeling is to observe that if one dropped the second superscript in $A^{x,y}$ and the first superscript in $B^{x,y}$, the system would have to be noncontextual (due to the fact that joint distributions of X, Y and of Y, Z imply the joint distribution of X, Y, Z). The existence of contextual systems therefore is a *reductio ad absurdum* proof that contextual labeling is necessary.

$[A^{x,y} = a, A^{x',y'} = a']$, $[A^{x,y} = a, B^{x',y'} = b]$, etc., are not well-defined events, and no probabilities can be assigned to them. This formal distinction reflects the obvious fact that random variables from mutually exclusive contexts can never be observed together, in any empirical meaning of "together." In particular, $A^{x,y}$ and $A^{x,y'}$ in (2) are not equal, because they are not jointly distributed. All this means that the system $\mathcal{R}$ in (1) is a collection of stochastically unrelated AB pairs, combined within a single system only because every AB pair shares at least one property it measures with at least one other AB pair.

Random variables attaining a given value with probability 1 are *deterministic variables*. A *deterministic system* is a system containing only deterministic variables. Thus, the two systems below are deterministic versions of the EPR-Bohm system, nonsignaling ($\mathcal{D}_{\mathrm{EPRB}}$) and signaling ($\mathcal{D}'_{\mathrm{EPRB}}$):

$\mathcal{D}_{\mathrm{EPRB}}$					$\mathcal{D}'_{\mathrm{EPRB}}$			
1		0			1		0	
1			1		0			1
	0	0				0	1	
	0		1			0		1
$x{=}1$	$x{=}2$	$y{=}1$	$y{=}2$		$x{=}1$	$x{=}2$	$y{=}1$	$y{=}2$

$$\tag{6}$$

In presenting these deterministic systems we conveniently identify the random variables with their supports, say, writing 0 instead of $A^{2,1} \equiv 0$ ($\equiv$ meaning "equal with probability 1"). Because one loses the indexation as a result, one has to indicate for each number what properties it measures (at the bottom of the tables).

Note that we use capital Roman letters to designate random variables, and the script letters $\mathcal{R}, \mathcal{D}$ to refer to systems—because a system is not a random variable, it is a set of stochastically unrelated random variables (the AB pairs).

A *coupling* $\bar{R}$ for a system $\mathcal{R}$ is an identically double-labeled set of *jointly distributed* random variables

$$\bar{R} = \{\bar{A}^{x,y}, \bar{B}^{x,y} : (x,y) \in C\}, \tag{7}$$

such that every AB pair of $\bar{R}$ is distributed as the corresponding AB pair of $\mathcal{R}$:

$$(\bar{A}^{x,y}, \bar{B}^{x,y}) \overset{\mathrm{dist}}{=} (A^{x,y}, B^{x,y}), \tag{8}$$

for every $(x,y) \in C$. Note that we can write $\bar{R}$ rather than $\bar{\mathcal{R}}$ in (7) because a coupling is a random variable in its own right. Thus, while a coupling of $\mathcal{R}_{\mathrm{EPRB}}$ can be presented as

$\bar{R}_{\mathrm{EPRB}}$			
$\bar{A}^{1,1}$		$\bar{B}^{1,1}$	
$\bar{A}^{1,2}$			$\bar{B}^{1,2}$
	$\bar{A}^{2,1}$	$\bar{B}^{2,1}$	
	$\bar{A}^{2,2}$		$\bar{B}^{2,2}$

$$\tag{9}$$

it is no longer a set of four stochastically unrelated pairs, but a random variable

$$\bar{R} = \{\bar{A}^{1,1}, \bar{A}^{1,2}, \bar{A}^{2,1}, \bar{A}^{2,2}, \bar{B}^{1,1}, \bar{B}^{2,1}, \bar{B}^{1,2}, \bar{B}^{2,2}\} \tag{10}$$

with 2^8 possible values.[2] Note also, that any deterministic system $\mathcal{D}$ has a unique coupling D, and the two are easy to

confuse if one uses our convenient identification of deterministic random variables with their supports. Thus, the coupling of $\mathcal{D}_{\mathrm{EPRB}}$ in (6) is written precisely as $\mathcal{D}_{\mathrm{EPRB}}$ itself:

$\bar{\mathcal{D}}_{\mathrm{EPRB}}$			
1		0	
1			1
	0	0	
	0		1
$x{=}1$	$x{=}2$	$y{=}1$	$y{=}2$

$$\tag{11}$$

A (nonsignaling) compound system $\mathcal{R}$ is *noncontextual* if it has a coupling $\bar{R}$ such that

$$
\begin{aligned}
\Pr[\bar{A}^{x,y} = \bar{A}^{x,y'}] &= 1, \\
\text{and} & \\
\Pr[\bar{B}^{x,y} = \bar{B}^{x',y}] &= 1,
\end{aligned}
\tag{12}
$$

whenever the indicated contexts are defined (belong to C). If such a coupling does not exist, the system is *contextual*.

Overlooking logical subtleties [16], this definition is equivalent to the traditional definitions of contextuality and locality, in terms of the nonexistence of joint distributions for single-indexed random variables [20] and in terms of hidden variables with noncontextual or local mapping into observables [7,18]. Perhaps this becomes more clear on observing that noncontextuality is equivalent to the existence of a set of jointly distributed *single-indexed* random variables

$$\tilde{R} = \{\tilde{A}^x, \tilde{B}^y : (x,y) \in C\}, \tag{13}$$

such that

$$(\tilde{A}^x, \tilde{B}^y) \overset{\mathrm{dist}}{=} (A^{x,y}, B^{x,y}), \tag{14}$$

for every $(x,y) \in C$.

III. FREE WILL THEOREM

If all AB pairs in a compound system $\mathcal{R}$ are set to specific values, then the resulting system is called a *realization* of $\mathcal{R}$. The reason this has to be presented as a definition is that, unlike a realization of a random variable (e.g., an AB pair), a realization of a system is not an observable outcome of any experiment: It is a pure mathematical abstraction, as the variables do not co-occur across contexts. Recall that each random variable in $\mathcal{R}$ is finite valued, because of which the set of possible realizations of $\mathcal{R}$ is finite. Thus, the system $\mathcal{R}_{\mathrm{EPRB}}$ has 4^4 realizations, whereas $\mathcal{R}_{\mathrm{KSP}}$ has $6^{40\cdot33}$ realizations.

The question posed in the FWT theorem can be formulated thusly: Given that an idealized experiment involving an unlimited number of particle pairs is described by a system $\mathcal{R}$, is it possible that each individual pair of particles is a deterministic system that coincides with one of the realizations of $\mathcal{R}$? The crux of the issue here is in whether the realizations of the systems can describe individual particle pairs. If not for this constraint, it would be innocuous (although still objectionable, as we will see later) to say that the system $\mathcal{R}$ is presentable

---

[2]To prevent misunderstanding, the term "random variable" is understood here in the standard meaning of modern probability theory,

with no restrictions on the codomain set of values: Random vectors and random sets therefore are also random variables.

as a mixture of some of its realizations $\mathcal{D}_1, \ldots, \mathcal{D}_k$ taken as deterministic systems:

$$\mathcal{R} \stackrel{\text{dist}}{=} \begin{cases} \mathcal{D}_1 & \text{with probability} & p_1 \\ \vdots & \vdots & \vdots \\ \mathcal{D}_k & \text{with probability} & p_k \end{cases}, \sum_{i=1}^{k} p_i = 1. \quad (15)$$

However, in the question asked by the FWT, these deterministic systems are assumed to describe real physical entities (particle pairs), because of which they should be physically realizable. In particular, they are subject to special relativity and have to be nonsignaling. For instance, in the decomposition of $\mathcal{R}_{\text{EPRB}}$, the signaling deterministic system $\mathcal{D}'_{\text{EPRB}}$ in (6) is not allowed.

Theorem (reformulated and generalized FWT). A contextual system $\mathcal{R}$ cannot be decomposed as in (15), where $\mathcal{D}_1, \ldots, \mathcal{D}_k$ are nonsignaling deterministic systems each of which coincides with a realization of $\mathcal{R}$.

Proof. If $k = 0$ (no nonsignaling realizations of $\mathcal{R}$ exist), the theorem is proved. Assume that (15) holds with $k > 0$. Introduce a random variable Λ such that $\Pr[\Lambda = i] = p_i$ ($i = 1, \ldots, k$). Each $\mathcal{D}_i$, being deterministic, has a unique coupling D_i, and the mixture

$$\bar{R} = \begin{cases} D_1 & \text{if} & \Lambda = 1 \\ \vdots & \vdots & \vdots \\ D_k & \text{if} & \Lambda = k \end{cases}, \sum_{i=1}^{k} p_i = 1 \quad (16)$$

is a coupling of $\mathcal{R}$. In this coupling,

$$\bar{A}^{x,y} = \begin{cases} a_1^{x,y} & \text{if} & \Lambda = 1 \\ \vdots & \vdots & \vdots \\ a_k^{x,y} & \text{if} & \Lambda = k \end{cases}, \quad (17)$$

where $a_i^{x,y}$ is the realization of $A^{x,y}$ in system $\mathcal{D}_i$. But $a_i^{x,y} = a_i^{x,y'}$ for any $(x,y), (x,y') \in C$ (nonsignaling). We have then

$$\bar{A}^{x,y} = \begin{cases} a_1^{x,y'} & \text{if} & \Lambda = 1 \\ \vdots & \vdots & \vdots \\ a_k^{x,y'} & \text{if} & \Lambda = k \end{cases} = \bar{A}^{x,y'}. \quad (18)$$

Analogously,

$$\bar{B}^{x,y} = \begin{cases} b_1^{x,y} = b_1^{x',y} & \text{if} & \Lambda = 1 \\ \vdots & \vdots & \vdots \\ b_k^{x,y} = b_k^{x',y} & \text{if} & \Lambda = k \end{cases} = \bar{B}^{x',y}, \quad (19)$$

for all $(x,y), (x',y) \in C$. By definition then, $\mathcal{R}$ is noncontextual, contrary to the theorem's premise.

Equivalently, and perhaps more familiar to physicists, the proof could be formulated as a demonstration that $\mathcal{R}$ has a local hidden variable model. Using the same Λ as in the proof, we have a coupling $\bar{R}$ of $\mathcal{R}$ such that

$$(\bar{A}^{x,y}, \bar{B}^{x,y}) = (f(\Lambda, x, y), g(\Lambda, x, y)), \quad (20)$$

where for each value $\Lambda = i$ and each (x,y), the function (f, g) reads the value of $(a_i^{x,y}, b_i^{x,y})$ in D_i. Since each D_i is

nonsignaling, i.e., $f(i, x, y) = f(i, x)$ and $g(i, x, y) = g(i, y)$ for each $\Lambda = i$, we have

$$(\bar{A}^{x,y}, \bar{B}^{x,y}) = (f(\Lambda, x), g(\Lambda, y)), \quad (21)$$

which is a local (i.e., noncontextual) model with Λ as a hidden variable.

Applying this theorem to $\mathcal{R}_{\text{KSP}}$ used in Conway and Kochen's proof, this system has no nonsignaling realizations (by the Kochen-Specker theorem). Rather surprisingly therefore, the proof of the Conway-Kochen version of the FWT is contained in the first sentence of the proof above. For $\mathcal{R}_{\text{EPRB}}$, we have 16 nonsignaling realizations of this system, and the proof says that their mixtures can only be noncontextual.

IV. WHERE IS THE FREE WILL ASSUMPTION IN THE PROOF?

The formulations and proofs given in the previous section do not even mention the hypothetical freedom with which Alice and Bob choose their settings. How is it possible? The answer is that the experimenters' free will assumption is indeed redundant. The proof above is contingent on the assumption that $\mathcal{R}$ is a correct description of an idealized experiment involving an unlimited number of particle pairs, those whereof we ask whether they could be deterministic systems. The experimenters' free will is only needed to dismiss a conspiracy of nature leading to an incorrect identification of the system in such an experiment. Let us explain this using the system $\mathcal{R}_{\text{EPRB}}$.

Consider the possibility that in an EPR-Bohm experiment only four types of entangled particle pairs are possible, described by the deterministic systems

$\mathcal{D}_1$			
1		1	
1			1
	1	1	
	1		1
$x=1$	$x=2$	$y=1$	$y=2$

$\mathcal{D}_2$			
0		0	
0			0
	0	0	
	0		0
$x=1$	$x=2$	$y=1$	$y=2$

$\mathcal{D}_3$			
0		0	
0			1
	1	0	
	1		1
$x=1$	$x=2$	$y=1$	$y=2$

$\mathcal{D}_4$			
1		0	
1			0
	0	0	
	0		0
$x=1$	$x=2$	$y=1$	$y=2$

$$(22)$$

Then the true system $\mathcal{R}_{\text{EPRB}}$ obtained from any mixture of these four systems is noncontextual. Suppose, however, that whenever Alice and Bob choose $(x, y) = (1, 1)$ or $(2, 1)$ (the first and third rows in the matrices), the nature chooses to supply a pair of particles described either by $\mathcal{D}_1$ or by $\mathcal{D}_2$, equiprobably; whereas for $(x, y) = (1, 2)$ and $(2, 2)$ (the second and fourth rows in the matrices) the nature chooses between $\mathcal{D}_3$ and $\mathcal{D}_4$ equiprobably. Neither Alice nor Bob nor anyone analyzing their experimental data has any way of knowing this. Following many replications, the results will be a statistical estimate of a system in which all random variables

are uniformly distributed,

$$\langle A^{x,y} \rangle = \langle B^{x,y} \rangle = \tfrac{1}{2}, \tag{23}$$

and

$$\langle A^{1,1} B^{1,1} \rangle = \langle A^{2,1} B^{2,1} \rangle = \langle A^{2,2} B^{2,2} \rangle = \tfrac{1}{2},$$
$$\langle A^{1,2} B^{1,2} \rangle = 0. \tag{24}$$

This is a nonsignaling contextual system (a PR box [21]), and if used in the proof of the FWT in place of what we assumed to be the true system, it will lead one to the wrong conclusion that no decomposition (15) is possible.

To avoid such conspiratorial scenarios one can postulate that the distribution of the nonsignaling deterministic systems in (15) is the same for all contexts (all choices of settings). This can be interpreted in terms of Alice's and Bob's free will, but does not have to. It would be better therefore to call this assumption *unbiasedness*, but not to multiply terminology we follow the authors who call it (measurement or setting)*independence*. The assumption, of course, is consistent with the choices of settings being perfectly predetermined, but simply uncorrelated with the occurrences of the different types of deterministic systems. Moreover, Alice's and Bob's choices may very well be correlated, it makes no difference.

The example just given, of a noncontextual system being mistaken for a contextual one, suggests the logical possibility of taking the decomposition (15) for granted, and accounting for the (apparent) contextuality of the observed system $\mathcal{R}$ either by relaxing the requirement of nonsignaling of the deterministic systems $\mathcal{D}_i$ or by exploring deviations from the independence assumption. There is an obvious reciprocity between the two, and it has indeed been researched and quantified [8–11]. This line of study is outside the scope of our paper. In the FWT we are only interested in whether individual particle pairs can be described by deterministic nonsignaling systems, and the answer given is negative.

As the above reasoning shows, the independence assumption is not a necessary premise of this theorem, because it is obviated by the critical assumption that certain contextual systems exist. The situation is this:

(i) if we do not assume, e.g., that $\mathcal{R}_{EPRB}$ for certain quadruples of axes is contextual, then the FWT for this system cannot be proved whether or not one adopts the independence assumption; and

(ii) if we do assume the contextuality of $\mathcal{R}_{EPRB}$ (presumably because we believe experiments or quantum-mechanical theory), the proof can be carried out without mentioning the independence assumption.

Analogously, Conway and Kochen have to postulate the existence of a system $\mathcal{R}_{KSP}$ with certain properties (the SPIN assumption), ensuring that no consistent assignment of values to Alice's measurements is possible (the Kochen-Specker theorem); and they also have to postulate that if Bob's axis coincides with one of the three axes chosen by Alice, then the corresponding measurement outcomes always coincide

(the TWIN assumption). With these postulates, however, the independence (part of their MIN assumption) is not needed.[3]

Summarizing, the independence (or "experimenters' free will") is only needed if we consider the epistemological question: How can one be sure that the compound system estimated from an experiment is truly contextual? The FWT is a conditional statement: If we have a contextual compound system with certain properties, it cannot be decomposed as in (15).

V. SYSTEMS VERSUS ISOLATED AB PAIRS

There is, however, a simpler reason not to use deterministic systems when describing individual particle pairs. The reason is that a single pair of particles is a *realization of an isolated AB pair* of a system rather than a realization of an entire system. The difference is that the AB pair is determined by the factual context chosen by Alice and Bob, while any given realization of a system also includes the counterfactual contexts that Alice and Bob "could have chosen." Thus, experimental trials for $\mathcal{R}_{EPRB}$ produce a series of outcomes, such as

$$
\begin{array}{c|c|c|c}
\vdots & \vdots & \vdots & \vdots \\
0 & & & 1 \\
 & 0 & 1 & \\
 & 1 & 1 & \\
0 & & 0 & \\
 & 0 & & 1 \\
\vdots & \vdots & \vdots & \vdots \\
\hline
x=1 & x=2 & y=1 & y=2
\end{array}
\tag{25}
$$

We see no logical reason to think that the observed value $(A^{1,2}, B^{1,2}) = (0, 1)$ is somehow related to any specific values of the AB pairs for contexts other than the factual $(x, y) = (1, 2)$. Such a relation would only be reasonable if there existed a way to observe these alternative AB pairs by factually performing the measurements for $(x, y) = (1, 1), (2, 1)$, and $(2, 2)$ in addition to $(x, y) = (1, 2)$ on the same pair of particles. This is, however, impossible, and not because a projective measurement is known to destroy the state of entanglement. We are not allowed to use quantum mechanical considerations when conceptually testing the standard quantum mechanical view of elementary particles.[4] The reason we

---

[3]As noted by critics of the FWT [4,5], MIN is not a rigorous statement. It speaks of Alice and Bob's choices as being made "freely" and "independently," of which the former we translate into the independence assumption. The "independently" of Conway and Kochen apparently means that Alice's and Bob's choices are combined in all possible ways. This assumption is not needed either. Thus, it is easy to see that in the system $\mathcal{R}_{KSP}$ one can delete all rows in which Bob's axis is not one of the three axes chosen by Alice. The assumption in question then will be violated, but the proof of the (original) FWT will not change at all (MIN also includes a statement that can be interpreted as nonsignaling principle).

[4]In fact, it has recently been established [22,23] that a sequence of POVM-represented measurements (each of which depends on the settings and outcomes of the previous measurements) can be performed on the same pair of entangled particles without affecting the entanglement state.

cannot speak of "the same" pair of entangled particles being measured repeatedly using variable settings is logical rather than physical. It is critical for our analysis that all particle pairs be generated and prepared in precisely the same way, and their spin values, for any given choice of settings, be measured in precisely the same way. A pair of particles after a measurement has been performed on it simply is not the same pair as it was before, and it should be relabeled accordingly. Barring such relabeling, a sequence like (25) should be treated as several (here, four) unrelated to each other sequences, each defined by a specific context.

One may now ask seemingly the same question as in the FWT but applied to pairs of particles treated as realizations of specific AB pairs: Can they be considered deterministic variables? In other words, the question is whether each pair of particles can be described as

$$
\boxed{A^{x,y} \equiv a^{x,y} \;\|\; B^{x,y} \equiv b^{x,y}}, \tag{26}
$$

with the specific settings (x, y) under which the measurements are recorded. This is, however, a very different question, and the answer is yes, if one so wishes, but this makes no difference. Flips of a fair coin can always be considered a mixture of two deterministic variables with respective values *head* and *tail*, each occurring with probability $1/2$. Following the prevailing tradition in statistics, a sequence of realizations of an AB pair can be treated as a realization of a *random sample* (a set of identically distributed independent random variables). However, it can also be treated as a realization of a set of different random variables (e.g., deterministic ones), randomly alternating. Let us again use $\mathcal{R}_{\text{EPRB}}$ for an illustration. Assume that one of its AB pairs is distributed as

value	(1, 1)	(1, 0)	(0, 1)	(0, 0)
probability	p	$1-p$	0	0

$$\tag{27}$$

Clearly, this variable is indistinguishable from any mixture

$$
\begin{cases}
X & \text{with probability} & q \\
Y & \text{with probability} & 1-q
\end{cases} \tag{28}
$$

of the random variables X, Y with the respective distributions

$$
X : \begin{array}{c|c} (1,1) & (1,0) \\ \hline p_1 & 1-p_1 \end{array}, \qquad
Y : \begin{array}{c|c} (1,1) & (1,0) \\ \hline p_2 & 1-p_2 \end{array}, \tag{29}
$$

provided

$$
q p_1 + (1-q) p_2 = p. \tag{30}
$$

In particular (and trivially), X, Y can be viewed as deterministic variables,

$$
X : \begin{array}{c|c} (1,1) & (1,0) \\ \hline 1 & 0 \end{array}, \qquad
Y : \begin{array}{c|c} (1,1) & (1,0) \\ \hline 0 & 1 \end{array}, \tag{31}
$$

mixed as

$$
\begin{cases}
X & \text{with probability} & p \\
Y & \text{with probability} & 1-p
\end{cases}. \tag{32}
$$

We see that the question of whether the individual pairs of particles have "free will" of their own, i.e., whether they are deterministic entities, loses its meaning.

It should also be noted that once this contextwise view of the individual particle pairs is adopted, one need not be concerned with the independence assumption ("free will" of Alice and Bob). In fact this assumption now is unformulable. Each AB pair corresponds to a fixed context, and one cannot informatively say that the AB pair in a given context does not depend on context.

VI. CONCLUSION

To summarize, the view that follows from the conceptual framework of CbD does not invalidate the FWT-type theorems, but makes them unnecessary. These theorems can be viewed as reductio ad absurdum demonstrations that individual pairs of particles cannot be viewed as deterministic systems. CbD allows one to streamline these demonstrations, ridding them of unnecessary assumptions. Moreover, CbD leads one to the view that the individual pairs of particles should not be viewed as systems to begin with, only as realizations of random variables for a given choice of settings. As Peres famously put it, "unperformed experiments have no results" [24]—and these nonexisting results should not be appended to factual results, lest one runs into a contradiction. The reason we can speak of $\mathcal{R}$ in (1) as a system is that as we switch from one context to another, all other macroscopic circumstances of measurements (overall experimental set-up and preparation procedure) remain the same. We can repeatedly experiment with such a system without changing its defining parameters. In particular, a deterministic system allows one to repeatedly experiment with it and obtain the same results for any given context. Individual pairs of particles, unless they change their identity, afford only one pair of measurements, in one particular context.

[1] J. Conway, S. Kochen, The free will theorem, Found. Phys. **36**, 1441 (2006).

[2] J. Conway and S. Kochen, The strong free will theorem, Not. Am. Math. Soc. **56**, 226 (2009).

[3] J. Conway and S. Kochen, Reply to comments of Bassi, Ghirardi, and Tumulka on the free will theorem, Found. Phys. **37**, 11 (2007).

[4] R. Tumulka, Comment on "The Free Will Theorem," Found. Phys. **37**, 186 (2007).

[5] S. Goldstein, D. V. Tausk, R. Tumulka, and N. Zanghí, What does the free will theorem actually prove? Not. Am. Math. Soc. **57**, 1451 (2010).

[6] E. Cator and K. Landsman, Constraints on determinism: Bell versus Conway–Kochen, Found. Phys. **44**, 781 (2014).

[7] J. S. Bell, On the Einstein–Podolsky–Rosen paradox, Physics **1**, 195 (1964).

[8] J. Barrett and N. Gisin, How Much Measurement Independence is Needed to Demonstrate Nonlocality? Phys. Rev. Lett. **106**, 100406 (2011).

[9] M. J. W. Hall, Relaxed Bell inequalities and Kochen-Specker theorems, Phys. Rev. A **84**, 022102 (2011).

[10] G. Pütz, D. Rosset, T. J. Barnea, Y.-C. Liang, and N. Gisin, Arbitrarily Small Amount of Measurement Independence is Sufficient to Manifest Quantum Nonlocality, Phys. Rev. Lett. **113**, 190402 (2014).

[11] A. S. Friedman, A. H. Guth, M. J. W. Hall, D. I. Kaiser, and J. Gallicchio, Relaxed Bell inequalities with arbitrary measurement dependence for each observer, Phys. Rev. A **99**, 012121 (2019).

[12] M. Esfeld, Bell's theorem and the issue of determinism and indeterminism, Found. Phys. **45**, 471 (2015).

[13] J. V. Kujala, E. N. Dzhafarov, and J.-A. Larsson, Necessary and Sufficient Conditions for Extended Noncontextuality in a Broad Class of Quantum Mechanical systems, Phys. Rev. Lett. **115**, 150401 (2015).

[14] E. N. Dzhafarov, V. H. Cervantes, and J. V. Kujala, Contextuality in canonical systems of random variables, Phil. Trans. Roy. Soc. A **375**, 20160389 (2017).

[15] E. N. Dzhafarov and J. V. Kujala, Probabilistic foundations of contextuality, Fortschr. Phys.-Progr. Phys. **65**, 1600040 (2017).

[16] E. N. Dzhafarov, On joint distributions, counterfactual values, and hidden variables in understanding contextuality, Phil. Trans. Roy. Soc. A **377**, 20190144 (2019).

[17] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[18] S. Kochen and E. P. Specker, The problem of hidden variables in quantum mechanics, J. Math. Mech. **17**, 59 (1967).

[19] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic, Amsterdam, 1995).

[20] A. Fine, Hidden Variables, Joint Probability, and the Bell Inequalities, Phys. Rev. Lett. **48**, 291 (1982).

[21] S. Popescu and D. Rohrlich, Nonlocality as an axiom, Found. Phys. **24**, 379 (1994).

[22] A. Tavakoli and A. Cabello, Quantum predictions for an unmeasured system cannot be simulated with a finite-memory classical system, Phys. Rev. A **97**, 032131 (2018).

[23] G. Foletto, L. Calderaro, A. Tavakoli, M. Schiavon, F. Picciariello, A. Cabello, P. Villoresi, and G. Vallone, Experimental Certification of Sustained Entanglement and Nonlocality After Sequential Measurements, Phys. Rev. Appl. **13**, 044008 (2020); **13**, 069902(E) (2020).

[24] A. Peres, Unperformed experiments have no results, Am. J. Phys. **46**, 745 (1978).