Context-independent mapping and free choice are equivalent: A general proof

Ehtibar N. Dzhafarov

Purdue University, USA, ehtibar@purdue.edu

Abstract

Free choice (or statistical independence) assumption in a hidden variable model (HVM) means that the settings chosen by experimenters do not depend on the values of the hidden variable. The assumption of context-independent (CI) mapping in an HVM means that the results of a measurement do not depend on settings for other measurements. If the measurements are spacelike separated, this assumption is known as local causality. Both free choice and CI mapping assumptions are considered necessary for derivation of the Bell-type criteria of contextuality/nonlocality. It is known, however, for a variety of special cases, that the two assumptions are not logically independent. We show here, in complete generality, for any system of random variables with or without disturbance/signaling, that an HVM that postulates CI mapping is equivalent to an HVM that postulates free choice. If one denies the possibility that a given empirical scenario can be described by an HVM in which measurements depend on other measurements' settings, free choice violations should be denied too, and vice versa.

KEYWORDS: Contextuality; context-independent mapping; free choice; local causality; nonlocality.

1 Introduction

The historical context for our analysis is set by a discussion that A. Shimony, M. A. Horne, and J. F. Clauser had with J. S. Bell [1]. The subject of this discussion was the assumptions underlying derivation of Bell's famous inequality [2] (or its offshoots [3]) for Bohm's version of the Einstein-Podolsky-Rosen (EPR) experiment [4]. Bell originally assumed that this derivation follows from the principle of local causality alone. According to this principle, if outcomes of an experiment are modeled by a function that contains as its arguments a "hidden" random variable and some parameters, these parameters cannot contain spacelike remote experimental settings. Shimony, Horne, and Clauser pointed out to Bell that, in addition to local causality, one had to postulate that the hidden random variable in the hypothetical model does not in any way correlate with experimental settings, local or remote. Bell agreed with this criticism (see [5] for a more detailed historical description). One consequence of the present paper is that Bell did not need to agree. He could have explained instead that one cannot accept local causality without free choice, because denying free choice is equivalent to denying local causality.

Let us define the terminology systematically. For our purposes, all empirical scenarios are described by *systems of random variables* representing measurements,

$$\mathcal{R} = \left\{ R_q^c : c \in C, q \in Q, q \prec c \right\}. \tag{1}$$

Here, R_q^c represents the outcome of measuring property q in *context* c (belonging to corresponding sets Q and C), and $q \prec c$ means that q is measured in c. A context c is any set of systematically recorded circumstances under which the properties are measured. As an example, the following is a system of random variables describing an EPR/Bohm experiment:

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

Here, $q = 1$ and 3 are settings used by Alice, and $q = 2$ and 4 are settings used by Bob. Contexts in (2) are defined by the four combinations of Alice's choice and Bob's choice. In the canonical version of the experiment, Alice's measurements are spacelike separated from Bob's choices, and vice versa.

A hidden variable model (HVM) of system $\mathcal{R}$ is a representation of the variables R_q^c in any given context c in the form

$$
\left\{ R_q^c \right\}_{q \prec c} \overset{d}{=} \left\{ \alpha \left(q, c, \Lambda^c \right) \right\}_{q \prec c},
\tag{3}
$$

where $\{expression\}_{q \prec c}$ is a compact way of writing $\{expression : q \in Q, q \prec c\}$, with c fixed. The symbol $\overset{d}{=}$ means "is distributed as," α is some function, and Λ^c are "hidden random variables." The equation says that the joint distribution of all R_q^c in a given context c is the same as the joint distribution of the corresponding $\alpha \left(q, c, \Lambda^c \right)$. In system (2), e.g., we have

$$
\begin{aligned}
\left\{ R_1^1, R_2^1 \right\} &\overset{d}{=} \left\{ \alpha \left(q = 1, c = 1, \Lambda^1 \right), \alpha \left(q = 2, c = 1, \Lambda^1 \right) \right\}, \\
\left\{ R_2^2, R_3^2 \right\} &\overset{d}{=} \left\{ \alpha \left(q = 2, c = 2, \Lambda^2 \right), \alpha \left(q = 3, c = 2, \Lambda^2 \right) \right\}, \\
\left\{ R_3^3, R_4^3 \right\} &\overset{d}{=} \left\{ \alpha \left(q = 3, c = 3, \Lambda^3 \right), \alpha \left(q = 4, c = 3, \Lambda^3 \right) \right\}, \\
\left\{ R_4^4, R_1^4 \right\} &\overset{d}{=} \left\{ \alpha \left(q = 4, c = 4, \Lambda^4 \right), \alpha \left(q = 1, c = 4, \Lambda^4 \right) \right\}.
\end{aligned}
\tag{4}
$$

We can also present this example in the form of a graph, where q and q' are the two settings used in a given context:

$$\tag{5}$$

An arrow here indicates that the distribution of its terminal node may change with the value of its initial node, and the dotted box indicates that the random variables within it are jointly distributed.

The dependence of Λ^c on c in (3) encompasses the possibility of violations of free choice. If the values of Λ^c and c are somehow interdependent, then different choices of c correspond to different distributions of the hidden variable. (In Conclusion, we will discuss why it is better not to treat c as a random variable.) Note that even if the the values of α do change with c (i.e., c is not a

dummy argument), it is still possible that the distribution of R_q^c does not depend on c. In other words, (3) is compatible with the condition of non-disturbance (non-signaling), and this is also true for the two special cases considered next.

We say that the HVM satisfies the assumption of *context-independent (CI) mapping* (or *local causality*, when spacelike separation is involved) if, for every c,

$$\left\{R_q^c\right\}_{q \prec c} \overset{d}{=} \left\{\beta\left(q, \Lambda^c\right)\right\}_{q \prec c}. \tag{6}$$

For system (2), this means

$$\tag{7}$$

In the model (6) we eliminate c as an explicit argument of function β, but generally allow Λ^c to have different distributions in different contexts c. Clearly, this implies that R_q^c may have different distributions for different c, at a fixed q. However, the usual view (dating back to [1]) is that CI-mapping (or local causality) is not violated here because the dependence of R_q^c on c is "indirect": the value of R_q^c at a fixed value of Λ^c depends on q alone. We will see below, however, that the difference between "direct" and "indirect" dependence on c is specious: they are completely interchangeable.

We say that the HVM satisfies the *free choice* assumption if, for every c,

$$\left\{R_q^c\right\}_{q \prec c} \overset{d}{=} \left\{\gamma\left(q, c, \Lambda\right)\right\}_{q \prec c}. \tag{8}$$

In this model Λ is one and the same for all q in all contexts c. In this general form, the HVM with free choice is allowed to violate the local causality assumption. For system (2), the model (8) can be presented as

$$\tag{9}$$

If both the assumptions of CI-mapping and free choice are satisfied, the HVM has the form

$$\left\{R_q^c\right\}_{q \prec c} \overset{d}{=} \left\{\delta\left(q, \Lambda\right)\right\}_{q \prec c}. \tag{10}$$

This is the HVM of a noncontextual (or locally causal) system of random variables, one for which one derives the traditional Bell-type criteria of noncontextuality/locality. The graph representation

of this model specialized to system (2) is

$$
\begin{array}{cc}
c & \Lambda \\
& \\
R_q^c & R_{q'}^c \\
\uparrow & \uparrow \\
q & q'
\end{array}
\tag{11}
$$

Although the only example we have given relates to the EPR/Bohm paradigm, the definitions just given and the results below apply to all situations described by systems of random variables, such as the classical Kochen-Specker scenario [6], the Klyachko-Can-Binicioğlu-Shumovsky paradigm [7], the Leggett-Garg experiments [8, 9], etc., with or without the assumption of no-disturbance (or no-signaling) [11, 10].

2   Theorem

The main point we make in this paper is very simple and has a very simple demonstration: HVMs (3), (6), and (8) are pairwise equivalent. We prove this by first showing the equivalence (6)⇔(8), and then the equivalence (3)⇔(8).

The proof requires one standard probabilistic notion, and one clarifying observation. The notion in question is (probabilistic) *coupling*: given an indexed set of random variables $\mathcal{X} = \{X_i : i \in I\}$, any set of *jointly distributed* random variables $Y = \{Y_i : i \in I\}$ such that $X_i \stackrel{d}{=} Y_i$ for all $i \in I$ is called a coupling of $\mathcal{X}$.

The observation in question is very simple, but is sometimes misunderstood: any indexed set of jointly distributed random variables is a random variable in its own right, and its components can always be presented as measurable functions of one and the same random variable. For instance, in a vector $U = \{U_1, \ldots, U_n\}$ of jointly distributed ± 1-valued variables, each U_i can be presented as a function of U (namely, its ith projection $\mathrm{Proj}_i(U)$). Equivalently, one can form a random variable V with 2^n values, in a bijective correspondence with $\{-1, 1\}^n$, and present each U_i as some function $f_i(V)$. Another example: if $U_1, \ldots, U_n$ are jointly distributed continuous random variables, V such that $U_i = f_i(V)$ can always be chosen to be uniformly distributed between 0 and 1 [12]. The difference between an indexed set of jointly distributed random variables and a "single" random variable is a matter of choosing between two interchangeable representations.

PROOF **(1a)** To show that, for any c,

$$
\left\{R_q^c\right\}_{q \prec c} \stackrel{d}{=} \left\{\beta\left(q, \Lambda^c\right)\right\}_{q \prec c} \Longrightarrow \left\{R_q^c\right\}_{q \prec c} \stackrel{d}{=} \left\{\gamma\left(q, c, \Lambda\right)\right\}_{q \prec c},
\tag{12}
$$

we form an arbitrary coupling Λ of the random variables $\{\Lambda^c : c \in C\}$. We have

$$
\Lambda^c \stackrel{d}{=} \mathrm{Proj}_c(\Lambda) = \phi(c, \Lambda).
\tag{13}
$$

But then

$$
\left\{\beta\left(q, \Lambda^c\right)\right\}_{q \prec c} \stackrel{d}{=} \left\{\beta\left(q, \phi\left(c, \Lambda\right)\right)\right\}_{q \prec c} = \left\{\gamma\left(q, c, \Lambda\right)\right\}_{q \prec c}.
\tag{14}
$$

(1b) To show the reverse implication

$$\left\{R_q^c\right\}_{q\prec c} \stackrel{d}{=} \left\{\gamma\left(q,c,\Lambda\right)\right\}_{q\prec c} \implies \left\{R_q^c\right\}_{q\prec c} \stackrel{d}{=} \left\{\beta\left(q,\Lambda^c\right)\right\}_{q\prec c}, \tag{15}$$

we define, for every c, the random variable

$$\Lambda^c := \left\{\gamma\left(q,c,\Lambda\right)\right\}_{q\prec c} \tag{16}$$

(whose components are jointly distributed because they are functions of one and the same Λ). The components $\gamma\left(q,c,\Lambda\right)$ in Λ^c are indexed by q, and

$$\gamma\left(q,c,\Lambda\right) = \mathrm{Proj}_q\left(\Lambda^c\right) = \beta\left(q,\Lambda^c\right). \tag{17}$$

But then

$$\left\{\gamma\left(q,c,\Lambda\right)\right\}_{q\prec c} \stackrel{d}{=} \left\{\beta\left(q,\Lambda^c\right)\right\}_{q\prec c}. \tag{18}$$

(2a) To show that, for any c,

$$\left\{R_q^c\right\}_{q\prec c} \stackrel{d}{=} \left\{\alpha\left(q,c,\Lambda^c\right)\right\}_{q\prec c} \implies \left\{R_q^c\right\}_{q\prec c} \stackrel{d}{=} \left\{\gamma\left(q,c,\Lambda\right)\right\}_{q\prec c}, \tag{19}$$

we again form an arbitrary coupling Λ of the random variables $\{\Lambda^c : c \in C\}$. We have

$$\Lambda^c \stackrel{d}{=} \mathrm{Proj}_c\left(\Lambda\right) = \phi\left(c,\Lambda\right). \tag{20}$$

But then

$$\left\{\alpha\left(q,c,\Lambda^c\right)\right\}_{q\prec c} \stackrel{d}{=} \left\{\alpha\left(q,c,\phi\left(c,\Lambda\right)\right)\right\}_{q\prec c} = \left\{\gamma\left(q,c,\Lambda\right)\right\}_{q\prec c}. \tag{21}$$

(2b) Finally

$$\left\{R_q^c\right\}_{q\prec c} \stackrel{d}{=} \left\{\gamma\left(q,c,\Lambda\right)\right\}_{q\prec c} \implies \left\{R_q^c\right\}_{q\prec c} \stackrel{d}{=} \left\{\alpha\left(q,c,\Lambda^c\right)\right\}_{q\prec c} \tag{22}$$

holds trivially. This completes the proof.

The parts (2a) and (2b) imply that the dependence of the hidden variable on context c in $\alpha\left(q,c,\Lambda^c\right)$ is superfluous: it can always be eliminated by replacing $\alpha\left(q,c,\Lambda^c\right)$ with $\gamma\left(q,c,\Lambda\right)$. The parts (1a) and (1b) show that the "indirect" dependence of $\beta\left(q,\Lambda^c\right)$ on c is not a special form of dependence: any $\gamma\left(q,c,\Lambda\right)$ or $\alpha\left(q,c,\Lambda^c\right)$ can be presented as $\beta\left(q,\Lambda^c\right)$.

An unexpected if not paradoxical consequence of the theorem is that an HVM of the form (6), in spite of being introduced as one satisfying CI mapping, does not in fact in fact impose any constraints on the HVMs that can describe the same system of random variables. In particular, (6) does not prevent this system of random variables from being modeled by an HVM of the form (8), one in which CI mapping generally does not hold. Conversely, an HVM of the form (8), in spite of being introduced as one satisfying the free choice assumption, does not in fact prevent the same system of random variables from being modeled by an HVM of the form (6), one in which free choice generally does not hold. This observation provides a simple explanation for the long since noticed reciprocity between measures of the degree to which an HVM violates the assumptions of CI mapping and free choice (as discussed in Conclusion).

Another implication of the same observation is that if one rejects the possibility that a system of random variables can be described by an HVM containing context c as one of its non-dummy arguments, then the possible HVMs for this scenario are of the form (10) rather than (6). This seems to have been John Bell's original idea, criticized in [1].

3   Discussion

Here we discuss a few questions that can be raised in response to the foregoing theorem.

(1) Would the analysis change if $\alpha\left(q,c,\Lambda^c\right)$ in (3) were replaced with a seemingly more general $\xi\left(q,c,\Lambda_q^c\right)$? The answer is it would make no difference, because by forming the random variable $\Lambda^c := \left\{\Lambda_q^c\right\}_{q \prec c}$, we get

$$\xi\left(q,c,\Lambda_q^c\right) = \xi\left(q,c,\mathrm{Proj}_q\left(\Lambda^c\right)\right) = \alpha\left(q,c,\Lambda^c\right). \tag{23}$$

The reason the components of $\left\{\Lambda_q^c\right\}_{q \prec c}$ are jointly distributed is that so are R_q^c for any given c.

(2) Could not the free choice assumption be violated in the form

$$\left\{R_q^c\right\}_{q \prec c} \overset{d}{=} \left\{\varrho\left(q,\Lambda^q\right)\right\}_{q \prec c}, \tag{24}$$

clearly compatible with CI-mapping (or local causality)? The answer is it would make no difference, because $\varrho\left(q,\Lambda^q\right)$ can always be written as $\delta\left(q,\Lambda\right)$. Indeed, considering Λ as the coupling of all Λ^q such that $q \prec c$ for some c, we get

$$\varrho\left(q,\Lambda^q\right) \overset{d}{=} \varrho\left(q,\mathrm{Proj}_q\left(\Lambda\right)\right) = \delta\left(q,\Lambda\right). \tag{25}$$

(3) Does not the difference between $\beta\left(q,\Lambda^c\right)$ and $\gamma\left(q,c,\Lambda\right)$ lie in the physical meaning of the dependence of these functions on c, in spite of their mathematical equivalence? The answer is negative once again. Not having any way to observe Λ, we can impose no physical constraints on what it is and how it can cause changes in the measurement outcomes R_q^c. Thus, for contexts with spacelike separated components, whatever Λ is, if one adopts the local causality assumption, Λ cannot transfer information from spacelike remote components of c to R_q^c. If the dependence of Λ on the remote components of c is explained by a common cause in their inverted light cones, then the same explanation applies to R_q^c and c directly, leading to the representation (8).

(4) Should not the free choice assumption be formulated in terms of the (non-)independence of the hidden variable Λ and context c treated as another random variable? The answer is that, first, this makes no difference, and second, treating c as a random variable is conceptually dubious. The reason this makes no difference is that, even if c is a random variable such that c and Λ are jointly distributed, conditioning Λ on different values of c creates the variables Λ^c of the analysis above. The reason why treating c as a random variable is dubious is that one can easily realize experimental procedures in which c is not chosen randomly: e.g., one can run four side-by-side EPR/Bohm experiments, each with a fixed value of c for years; or one can change the values of c in accordance with a deterministic algorithm (say, $1, 2, 3, 4, 1, 2, 3, 4, \ldots$). The notions of randomness and of a random variable are not identical, so the procedures just mentioned may still allow one to treat c as a random variable — but this cannot be done uniquely, in a standard way based on frequencies of occurrences. Conditioning Λ on c, on the other hand, is innocuous: realizations of Λ cannot be controlled in any way.

4   Conclusion

The analysis presented in this article has consequences beyond just the issue of how one derives the Bell-type criteria. There is a sizable body of insightful literature on the reciprocity between the

degree free choice is violated and the degree of dependence of measurements on contexts [13, 14, 15, 16]. It is clear now that the logical basis of this literature, usually focusing on specific cases, such as the EPR/Bohm type experiment, is the equivalence of (8) and (6). We know now that this equivalence holds in complete generality, for all possible systems of random variables (1), with and without disturbance alike. This equivalence implies, in particular, that any measure of deviation of an HVM from the model (10) should equally be interpretable as the degree of violation of CI mapping (local causality) and the degree to which the experimenters lack free choice. This is most clearly indicated in the recent paper by Blasiak et al. [17]. The abstract of their paper states that "causal explanations resorting to either locality or free choice violations are fully interchangeable." This coincides with the results of the present work, except that here they are established by a very different argument and in greater if not maximal generality.

Choosing between equivalent representations may seem a worthless exercise, but one may have grounds for preference in this case. Local causality is a principle derived from special relativity, and its applicability in physics is universal. In Kochen-Specker-type contextuality the CI mapping is derived from the definition of observables in quantum mechanics. Its applicability, too, therefore is much wider than establishing and measuring contextuality/nonlocality. Violation of free choice, on the other hand, is merely an abstract logical possibility narrowly aimed at explaining how a noncontextual world could fool us into observing contextual systems [18].

Conflict of Interests Author declares no conflicts of interests

Data Availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

[1] J. S. Bell, A. Shimony, M. A. Horne, J. F. Clauser, An exchange on local beables. Dial. 39, 85-110 (1985).

[2] J. S. Bell, On the problem of hidden variables in quantum mechanics, Rev. Mod. Phys. 38, 447 (1966).

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

[4] D. Bohm, Y. Aharonov, Discussion of experimental proof for the paradox of Einstein, Rosen and Podolski. Phys. Rev. 108, 1070-1076 (1957).

[5] T. Norsen, J.S. Bell's Concept of Local Causality. Am. J. Phys. 79, 1261 (2011)

[6] S. Kochen, E.P. Specker, The problem of hidden variables in quantum mechanics. J. Math. Mech., 17:59–87 (1967).

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

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

[9] J. Kofler and C. Brukner, Condition for macroscopic realism beyond the Leggett-Garg inequalities, Phys. Rev. A 87, 052115 (2013).

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

[11] E. N. Dzhafarov, J. V. Kujala, V.H. Cervantes, Contextuality and noncontextuality measures and generalized Bell inequalities for cyclic systems. Phys. Rev. A 101, 042119 (2020). [Erratum notes in Phys. Rev. A 101, 069902 (2020) and Phys. Rev. A 103:059901 (2021).]

[12] A. S. Kechris. Classical descriptive set theory. New York: Springer (1995).

[13] J. Barrett, N. Gisin, How much measurement independence is needed to demonstrate nonlocality? Phys. Rev. Lett. 106, 100406 (2011).

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

[15] G. Pütz, D. Rosset, T. J. Barnea, Y.-C. Liang, N. Gisin, Arbitrarily small amount of measurement independence is sufficient to manifest quantum nonlocality. Phys. Rev. Lett. 113, 190402 (2014).

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

[17] P. Blasiak, E. M. Pothos, J. M. Yearsley, C. Gallus, E. Borsuk, Violations of locality and free choice are equivalent resources in Bell experiments. PNAS 118 e2020569118 (2021).

[18] E. N. Dzhafarov, J. V. Kujala, Systems of random variables and the Free Will Theorem. Phys. Rev. Res. 2:043288 (2020).