Minimal Assumption Derivation of a Bell-type Inequality

Article *in* The British Journal for the Philosophy of Science · December 2003

DOI: 10.1093/bjps/axi140 · Source: OAI

3 authors, including:

Gerd Graßhoff
Humboldt-Universität zu Berlin

Samuel Portmann
University of Geneva

Minimal assumption derivation of a Bell-type inequality

Gerd Graßhoff,[1, *] Samuel Portmann,[1, 2, †] and Adrian Wüthrich[1, 2, ‡]

[1] *History and Philosophy of Science, Exact Sciences, Sidlerstrasse, University of Bern, CH-3012 Bern, Switzerland*[§]

[2] *Institute of Theoretical Physics, University of Bern, Switzerland*[¶]

(Dated: February 1, 2008)

John Bell showed that a big class of local hidden-variable models stands in conflict with quantum mechanics and experiment. Recently, there were suggestions that empirical adequate hidden-variable models might exist, which presuppose a weaker notion of local causality. We will show that a Bell-type inequality can be derived also from these weaker assumptions.

PACS numbers: 03.65.Ud

I. INTRODUCTION

The violation of Bell's inequality by the outcome of an EPR-type spin experiment [1, 2] seems to exclude a local theory with hidden variables. The underlying *reductio ad absurdum* proof infers on the grounds of the empirical falsification of the derived inequality that at least one of the required assumptions must be false. The force of the argument requires that the derivation be deductive and that all assumptions be explicit. We aim to extract a minimal set of assumptions needed for a deductive derivation of Bell's inequalities given perfect correlation of outcomes of an EPR-type spin experiment with parallel settings.

One of the assumptions in Bell's original derivation [3] was determinism. Later, he succeeded in deriving a similar inequality without determinism [4], placing in its stead an assumption later dubbed *local causality* [5]. As Bell stressed, the notion of local causality he and others used might be challenged. In [6] it was pointed out, that Reichenbach's Common Cause Principle [7] indeed suggests a weaker form of local causality. We will prove here, however, that even from this weaker notion Bell's inequality can still be derived.[1]

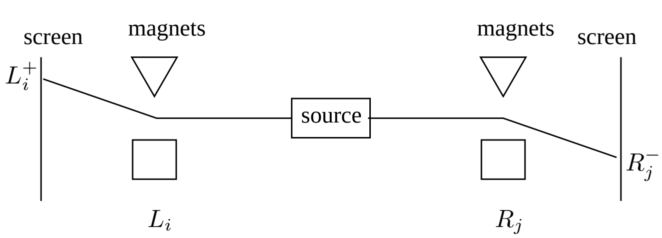

FIG. 1: Setup of the EPR-Bohm experiment. Cf. [9, p. 140].

II. THE EPR-BOHM EXPERIMENT

Consider the so-called *EPR-Bohm (EPRB) experiment* [1, 2]. Two spin-$\frac{1}{2}$ particles in the *singlet state*

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) \tag{1}$$

are separated in such a way that one particle moves to a measurement apparatus in the left wing of the experimental setting and the other particle to a measurement apparatus in the right wing (see FIG. 1). The experimenter can choose arbitrarily one of three directions in which the spin is measured with a Stern-Gerlach magnet.

The following terminology follows the reconstruction by Wigner [10] and which van Fraassen [11] has subsequently expanded on. The event type[2] that the left (right) measurement apparatus is set to measure the spin in direction $i \in \{1, 2, 3\}$ is symbolized by L_i (R_i). L_i^a (R_i^a) symbolizes the event type that the measurement outcome in the left (right) wing of a spin measurement in direction i is a. There are two possible measurement outcomes *spin up* ($a{=}+$) and *spin down* ($a{=}-$) for each particle in each direction. The letter $j \in \{1, 2, 3\}$ will be used like i to symbolize directions and $b \in \{+, -\}$ like a to symbolize measurement outcomes. Formulas in which the variables i, j, a, and b appear are meant to hold—if not otherwise stated—for all possible values of the variables. $p(X)$ denotes the probability of an event type X, which is empirically measurable as the relative frequency of all runs of an EPRB experiment in which the event type X is instantiated, with respect to all runs. $p(X \wedge Y)$ is the probability of the event type 'X and Y', measurable as the relative frequency of all runs in which both X and Y are instantiated. $p(X|Y) = p(X \wedge Y)/p(Y)$ is the conditional probability of the event type X *given* the event type Y, measurable as the relative frequency of instantiations of X with respect to the subensemble of

---

[*] Electronic address: gerd.grasshoff@philo.unibe.ch
[†] Electronic address: portmann@itp.unibe.ch
[‡] Electronic address: awuethr@itp.unibe.ch
[§] URL: `http://www.philoscience.unibe.ch`
[¶] URL: `http://www-itp.unibe.ch`

[1] Several of the issues we present in this paper are discussed in more detail in [8].

[2] We will speak of *event types* to distinguish them from the *token events* which instantiate corresponding event types.

all runs in which Y is instantiated. E.g.

$$p(L_i^a \wedge R_j^b | L_i \wedge R_j) \tag{2}$$

denotes the probability that the measurement outcome is a on the left and b on the right, when measuring in direction i on the left and in direction j on the right. These probabilities are predicted by quantum mechanics as

$$p(L_i^+ \wedge R_j^+ | L_i \wedge R_j) = \frac{1}{2}\sin^2\frac{\varphi_{ij}}{2}, \tag{3}$$

$$p(L_i^- \wedge R_j^- | L_i \wedge R_j) = \frac{1}{2}\sin^2\frac{\varphi_{ij}}{2}, \tag{4}$$

$$p(L_i^+ \wedge R_j^- | L_i \wedge R_j) = \frac{1}{2}\cos^2\frac{\varphi_{ij}}{2}, \tag{5}$$

$$p(L_i^- \wedge R_j^+ | L_i \wedge R_j) = \frac{1}{2}\cos^2\frac{\varphi_{ij}}{2}, \tag{6}$$

where φ_{ij} denotes the angle between the two measurement directions i and j. Also, the outcomes on each side are predicted separately to be completely random:

$$p(L_i^a | L_i \wedge R_j) = \frac{1}{2}, \tag{7}$$

$$p(R_j^b | L_i \wedge R_j) = \frac{1}{2}. \tag{8}$$

III. LOCAL CAUSALITY

The derivations of Bell-type inequalities known to us which *do not presuppose determinism* assume instead what John Bell calls *local causality* [5, 12]. That is, the assumption that there is common cause variable[3] V which takes on values $q \in I = \{q_1, q_2, q_3, \ldots, q_k\}$ such that for event types 'the variable V has the value q' (Vq) we have $\sum_q p(Vq) = 1$ and

$$p(L_i^a \wedge R_j^b | Vq \wedge L_i \wedge R_j) = p(L_i^a | Vq \wedge L_i) \times p(R_j^b | Vq \wedge R_j). \tag{9}$$

Other frequently used names for this condition are *factorizability* [13] and *strong locality* [14, 15]. It is usually justified by pointing out that it follows from the conjunction of the following three conditions, which are called *completeness* (equation (10)) and *locality* (equations (11) and (12)) [14, 15], *outcome independence* and *parameter independence* [16], or *causality* and *hidden locality* [11]:

$$p(L_i^a \wedge R_j^b | Vq \wedge L_i \wedge R_j) = p(L_i^a | Vq \wedge L_i \wedge R_j) \times p(R_j^b | Vq \wedge L_i \wedge R_j), \tag{10}$$

$$p(L_i^a | L_i \wedge R_j \wedge Vq) = p(L_i^a | L_i \wedge Vq), \tag{11}$$

$$p(R_j^a | L_i \wedge R_j \wedge Vq) = p(R_j^a | R_j \wedge Vq). \tag{12}$$

[3] For the sake of simplicity, we assume that this partition is discrete and finite. As will become clear in the following, the derivation of Bell's inequality can also be done without this restriction.

Equation (10) says that event types Vq or the variable V *"screens off"* L_i^a and R_j^b from each other [11, 13]. Van Fraassen [11] pointed out, that equation (10) can be motivated through *Reichenbach's Principle of Common Cause* (PCC) [7]. The principle states, that whenever two different event types A and B are statistically correlated

$$p(A \wedge B) \neq p(A)p(B) \tag{13}$$

and neither A is causally relevant for B nor B for A, there exists a common cause variable V with values $q \in I = \{q_1, q_2, q_3, \ldots, q_k\}$ ($\sum_q p(Vq) = 1$) such that A and B given Vq are uncorrelated:

$$p(A \wedge B | Vq) = p(A|Vq)p(B|Vq). \tag{14}$$

In its original formulation the principle is stated only for a common cause event type C, which is included in our formulation as the special case where Vq can take only two values: $Vq_1 = C$, $Vq_2 = \neg C$ ('not C'). The principle was formulated for general common cause variables by Hofer-Szabó et al. [17] and Placek [18]. Besides the screening-off condition Reichenbach [7] and Hofer-Szabó et al. [17] stipulate further restrictions on the common cause variable, which are, however, irrelevant for our purposes.

Now, as can be seen from equations (3)-(6), the event type L_i^a is in general correlated with event type R_j^b. It is

$$p(L_i^a | L_i \wedge R_j) = p(R_j^b | L_i \wedge R_j) = \frac{1}{2}, \tag{15}$$

and therefore

$$p(L_i^a \wedge R_j^b | L_i \wedge R_j) \neq p(L_i^a | L_i \wedge R_j) p(R_j^b | L_i \wedge R_j) \quad \text{except for} \quad \varphi_{ij} = \frac{\pi}{2} \mod \pi. \tag{16}$$

Supposing that L_i^a is not causally relevant for R_j^b and vice versa (which is reinforced by the fact that the setup of the experiment can be chosen so that the instantiations of L_i^a and R_j^b in each run of the experiment are space-like separated), PCC requires a common cause variable which fulfills equation (10). There are several different correlations; e.g. L_1^+ is correlated with R_2^+, and L_2^+ is correlated with R_3^+. For each of these correlations PCC enforces the consequence that a common cause variable exists. As stressed in [6] nothing in PCC dictates that the common cause variables of the different correlations have to be the same. However, in all the derivations of Bell's inequality known to us this identification is made nevertheless. It is further shown in [6] and [17], that for any set of correlations it is mathematically possible to construct common cause variables. The authors concluded in [6] that the apparent contradiction between this possibility and the claim that the EPRB correlations do not allow for a common cause variable [11, 13], is resolved by pointing out that in the derivation of Bell's inequality a *common* common cause variable for all measurements is assumed:

> "The crucial assumption in the [...] derivation of the [Clauser-Horne] inequality is that [the two-valued common cause variable] is a [two-valued common cause variable] *for all four* correlated pairs, i. e. that [Vq] is a *common* common cause [variable], shared by different correlations. Without this assumption Bell's inequality *cannot* be derived. But there does not seem to be any obvious reason why common causes should also be common common causes, whether of quantum or of any other sort of correlations." (Italics in the original.)

Showing the mathematical possibility of constructing common cause variables for any set of correlations and in particular for the correlations found in the EPRB experiment is not sufficient for proving the existence of a physically "natural" hidden-variable model for that experiment, however. Besides being common cause variables (thus fulfilling equation (10)), parameter independence should hold, too (equations (11) and (12)). Also, they should not be correlated with the measurement choices. As shown by Szabó [19], it is possible to construct a model which fulfills these requirements for each of the common cause variables separately. However, the *conjunctions* and other logical combinations of the event types that the common cause variables have certain values correlate in that model with the measurement operations. Whether a model can be constructed without these correlations was posed as an open question by Szabó. This question is answered negatively by the derivation of Bell's inequality that we present in the remainder of this article.

IV. BELL'S INEQUALITY FROM SEPARATE COMMON CAUSES

A. A weak screening-off principle

Consider an EPRB experiment where the same direction i ($i \in \{1, 2, 3\}$) is chosen in both wings. That is, in each run the event type $L_i \wedge R_i$ is instantiated. With this special setting quantum mechanics predicts (see equations (3)-(8), with $\varphi_{ij} = 0$) that the measurement outcomes in each wing are random but that the outcomes in one wing are perfectly correlated with the outcomes in the other wing: if and only if the spin of the left particle is up, then the spin of the right particle is down, and vice versa. We refer to this assumption as *perfect correlation*, or PCORR for short.

Assumption 1 (PCORR)

$$p_{ii}(R_i^-|L_i^+) = 1 \textit{ and } p_{ii}(L_i^+|R_i^-) = 1. \qquad (17)$$

We use here the definition

$$p_{ij}(\dots) \doteq p(\dots|L_i \wedge R_j). \qquad (18)$$

Large spatial separation of coinciding events of type L_i^a and R_j^b suggests that the respective instances are indeed distinct events. This excludes an explanation of the correlations by *event identity*, as is the case, for example, with a tossed coin for the perfect correlation of the event types 'heads up' and 'tails down'. Such a perfect correlation is explained in that every instance of 'heads up' is also an instance of 'tails down', and vice versa. Since the separation is even space-like, no L_i^a or R_j^b should be causally relevant for the other. We refer to these two assumptions as *separability*, SEP for short, and *locality 1* (LOC1).

Assumption 2 (SEP) *The coinciding instances of L_i^a and R_j^b are distinct events.*

Assumption 3 (LOC1) *No L_i^a or R_j^b is causally relevant for the other.*

Rather, there should be a common cause variable; that is, we assume PCC.

Assumption 4 (PCC) *If two event types A and B are correlated and the correlation cannot be explained by direct causation nor event identity, then there exists a common cause variable V, with values $q \in I = \{q_1, q_2, q_3, \dots, q_k\}$ such that $\sum_q p(Vq) = 1$ and*

$$p(A \wedge B|Vq) = p(A|Vq)p(B|Vq), \qquad \forall q.$$

As already mentioned, we omit the other Reichenbachian conditions [7, 17] since they are not necessary for our derivation.

This principle together with the assumptions PCORR, SEP and LOC1 implies that there is for each of the EPRB correlations a (separate) common cause variable V_{ij}^{+-} with $q \in I_{ij}^{+-}$.

Result 1

$$p_{ii}(L_i^+ \wedge R_i^-|V_{ii}^{+-}q) = p_{ii}(L_i^+|V_{ii}^{+-}q) \times p_{ii}(R_i^-|V_{ii}^{+-}q). \qquad (19)$$

Note that common cause variables can be different for different correlations.

B. Perfect correlation and "determinism"

We now show that from the fact that a *perfect* correlation is screened off by some variable it follows that without loss of generality the common cause variable can be assumed to be two-valued and that the having of one of the two values of the variables is necessary and sufficient for the instantiation of the two perfectly correlated event types, cf. [20].

Let A and B be perfectly correlated,

$$p(A|B) = p(B|A) = 1,$$

and screened-off from each other by a common cause variable,

$$p(A \wedge B|Vq) = p(A|Vq)p(B|Vq).$$

We can split the set I of all values V completely into two disjunct subsets, namely in the subset I^+ of those values of V for which $p(A \wedge Vq)$ is not zero and in the subset I^- of those for which it is zero:

$$\begin{aligned} I^+ &= \{q \in I : \; p(A \wedge Vq) \neq 0\}, \\ I^- &= \{q \in I : \; p(A \wedge Vq) = 0\}, \\ I &= I^- \cup I^+, \; I^- \cap I^+ = \emptyset. \end{aligned}$$

¿From this definition of I^- it follows already that

$$p(A|Vq) = 0, \quad \forall q \in I^-, \tag{20}$$

i. e. that Vq with $q \in I^+$ is necessary for A. Moreover, for all $q \in I^+$ we have by screening off and perfect correlation

$$p(A|Vq) = p(A|B \wedge Vq) = 1. \tag{21}$$

That the variable has a value in I^+ is a necessary and sufficient condition for A. The following calculation shows that Vq with $q \in I^+$ is also necessary and sufficient for B.

¿From perfect correlation it follows that

$$p(B|A \wedge Vq) = 1, \quad \forall q \in I^+.$$

That Vq screens off B from A yields

$$p(B|A \wedge Vq) = p(B|Vq).$$

Together with the previous equation this implies that Vq is sufficient for B for all $q \in I^+$:

$$p(B|Vq) = 1 \quad \forall q \in I^+. \tag{22}$$

If $q \in I^-$ we have by definition $p(A \wedge Vq) = 0$, which implies

$$p(A \wedge B \wedge Vq) = 0.$$

By perfect correlation we have therefore also $p(B \wedge Vq) = 0$, which in turn implies that

$$p(B|Vq) = 0, \quad \forall q \in I^-, \tag{23}$$

which means that Vq with $q \in I^+$ is also necessary for B.

This calculation shows that in the case of a perfect correlation the set of values of the common cause variable decomposes into two relevant sets. This means that whenever there is an (arbitrarily-valued) common cause variable for a perfect correlation, there is also a two-valued common cause variable, namely the disjunction of all event types Vq for which $q \in I^+$ or $q \in I^-$, respectively.

$$\begin{aligned} C &= \vee_{q \in I^+} Vq, \\ \neg C &= \vee_{q \in I^-} Vq. \end{aligned}$$

We refer to C as a *common cause event type*. In the case of a perfect correlation no generality is achieved by allowing for a more than two-valued common cause variable; if there is a common cause variable for a perfect correlation, there is also a common cause event type. Moreover, the common cause event type is a necessary and sufficient condition for the event types that are screened off by it (equations (20), (21), (22) and (23)).

Result 1 thus implies that there is a common cause event type C_{ii}^{+-} such that

$$p_{ii}(L_i^+|C_{ii}^{+-}) = p_{ii}(R_i^-|C_{ii}^{+-}) = 1, \tag{24}$$
$$p_{ii}(L_i^+|\neg C_{ii}^{+-}) = p_{ii}(R_i^-|\neg C_{ii}^{+-}) = 0. \tag{25}$$

The sub- and superscripts of C_{ii}^{+-} refer to C_{ii}^{+-} being the common cause event type of L_i^+ and R_i^-.

The outcome of a spin measurement is always either $+$ or $-$ and nothing else. We call this assumption *exactly one of exactly two possible outcomes* (EX).

Assumption 5 (EX)

$$p_{ii}(L_i^+) + p_{ii}(L_i^-) = 1, \qquad p_{ii}(L_i^+ \wedge L_i^-) = 0, \tag{26}$$
$$p_{ii}(R_i^+) + p_{ii}(R_i^-) = 1, \qquad p_{ii}(R_i^+ \wedge R_i^-) = 0. \tag{27}$$

As stressed by Fine [21], among the actual measurements there are always runs in which no outcome is registered, which is normally attributed to the limited efficiency of the detectors and not taken to the statistics. If one assumes instead, that part of these no-outcome runs are caused by the hidden variable, then it is possible to construct empirically adequate models for the EPRB experiments [22, 23]. With assumption 5, we explicitly exclude such models.

With assumption 5, while C_{ii}^{+-} is necessary and sufficient for L_i^+ and R_i^-, its complement, i. e. $\neg C_{ii}^{+-}$ is necessary and sufficient for the opposite outcomes, i. e. L_i^- and R_i^+:

$$p_{ii}(L_i^-|C_{ii}^{+-}) = p_{ii}(R_i^+|C_{ii}^{+-}) = 0, \tag{28}$$
$$p_{ii}(L_i^-|\neg C_{ii}^{+-}) = p_{ii}(R_i^+|\neg C_{ii}^{+-}) = 1. \tag{29}$$

C. A *minimal theory* for spins

In section IV B it was found that C_{ii}^{+-} is sufficient for L_i^+ *given* parallel settings $(L_i \wedge R_i)$, see equation (24). I. e. the conjunction $C_{ii}^{+-} \wedge L_i \wedge R_i$ is sufficient for L_i^+. But because of space-like separation of events of type L_i^+ and R_i that are instantiated in the same run, the latter types should not be causally relevant for the former. The measurement choice in one wing should be causally irrelevant for the outcomes (and the choices) in the other wing. Therefore we should discard R_i from the sufficient conjunction. The part $C_{ii}^{+-} \wedge L_i$ *alone* is sufficient for L_i^+. A similar reasoning can be applied to R_j^+, R_j and $\neg C_{jj}^{+-}$, cf. equation (29). This is our assumption *locality 2* (LOC2).

Assumption 6 (LOC2) *If $L_i \wedge R_i \wedge X$ is sufficient for L_i^+, then $L_i \wedge X$ alone is sufficient for L_i^+; and similarly for R_j^+, i. e. if $L_j \wedge R_j \wedge Y$ is sufficient for R_j^+, then $R_j \wedge Y$ alone is sufficient for R_j^+.*

Moreover, the remaining part $C_{ii}^{+-} \wedge L_i$ is *minimally* sufficient, in the sense that none of its parts is sufficient on its own.[4] If, for example, C_{11}^{+-} is instantiated, but we do not choose to measure L_1, then L_1^+ will not be instantiated. That is to say, we cannot discard yet another conjunct of $L_i \wedge C_{ii}^{+-}$ as we discarded R_i from $C_{ii}^{+-} \wedge L_i \wedge R_i$.

Let us turn to *necessary* conditions for L_i^+. To begin with, L_i is necessary: If there is no Stern-Gerlach magnet properly set up (L_i) the particle is not deflected either up- or downwards; similarly for L_i^-, R_j^+ and R_j^-. Roughly speaking, *no outcome without measurement* (NOWM).

Assumption 7 (NOWM)

$$p(L_i^+ \wedge \neg L_i) = 0, \qquad p(L_i^- \wedge \neg L_i) = 0, \tag{30}$$
$$p(R_j^+ \wedge \neg R_j) = 0, \qquad p(R_j^- \wedge \neg R_j) = 0. \tag{31}$$

Second, we saw in section IV B that if parallel settings are chosen and $\neg C_{ii}^{+-}$ is instantiated an event of type L_i^+ does never occur. In other words, $\neg C_{ii}^{+-} \wedge L_i \wedge R_i$ implies $\neg L_i^+$:

$$\neg C_{ii}^{+-} \wedge L_i \wedge R_i \rightarrow \neg L_i^+. \tag{32}$$

Again we propose a locality condition based on the idea that the measurement choice in one wing should be causally irrelevant for the outcomes (and the choices) in the other wing:[5] If $\neg C_{ii}^{+-} \wedge L_i \wedge R_i$ are sufficient for $\neg L_i^+$, then $\neg C_{ii}^{+-} \wedge L_i$ *alone* should be sufficient for $\neg L_i^+$. A similar reasoning can be applied to R_j^+, R_j and C_{jj}^{+-}, cf. equation (28).

Assumption 8 (LOC3) *If $L_i \wedge R_i \wedge X$ is sufficient for $\neg L_i^+$, then $L_i \wedge X$ alone is sufficient for $\neg L_i^+$; and similarly for $\neg R_j^+$, i. e. if $L_j \wedge R_j \wedge Y$ is sufficient for $\neg R_j^+$, then $R_j \wedge Y$ alone is sufficient for $\neg R_j^+$.*

By LOC3 it follows from equation (32) that

$$\neg C_{ii}^{+-} \wedge L_i \rightarrow \neg L_i^+. \tag{33}$$

This is equivalent to

$$L_i^+ \wedge L_i \rightarrow C_{ii}^{+-}, \tag{34}$$

and also to

$$L_i^+ \wedge L_i \rightarrow C_{ii}^{+-} \wedge L_i. \tag{35}$$

[4] Minimal sufficient conditions as definied by [24] and [25].

[5] The following version of LOC3 is slightly different from an earlier version of the article. We thank Gabor Hofer-Szabó, Miklos Rédei and Iñaki San Pedro for their comments.

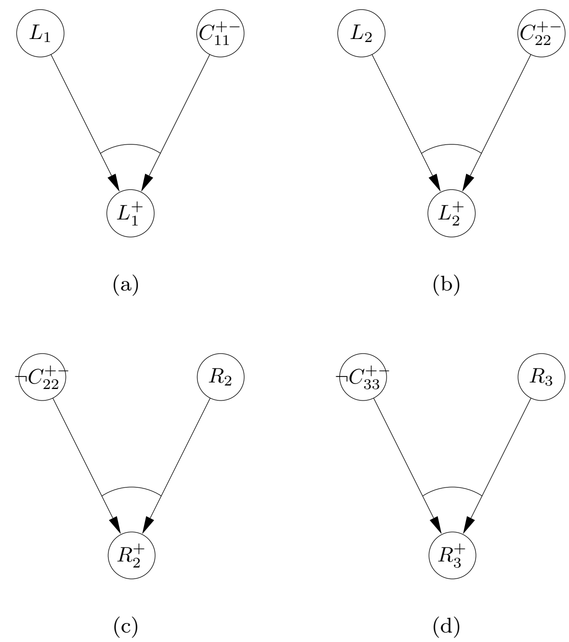

FIG. 2: Minimal theories for outcomes of spin measurements.

According to equation (30), L_i is necessary for L_i^+. That means $L_i^+ \rightarrow L_i$, but also $L_i^+ \rightarrow L_i^+ \wedge L_i$. Above, we have found (eq. (35)) that $L_i^+ \wedge L_i \rightarrow C_{ii}^{+-} \wedge L_i$. Altogether, this entails $L_i^+ \rightarrow L_i \wedge C_{ii}^{+-}$, i. e. that $L_i \wedge C_{ii}^{+-}$ is necessary for L_i^+. Moreover, it is a *minimally* necessary condition in the sense of [24] since it does not contain any disjuncts. All in all: $C_{ii}^{+-} \wedge L_i$ is a minimally necessary and minimally sufficient condition for L_i^+. In a similar vein we find that $R_j \wedge \neg C_{jj}^{+-}$ is minimally necessary and minimally sufficient for R_j^+. We have thus derived in particular the four *minimal theories* in the sense of [24] as illustrated in FIG. 2.

In a formal notation the four minimal theories read as the following four equations, where $\leftrightarrow$ is the usual *biconditional*, which means that the left-hand side implies the right-hand side and vice versa.[6] This intermediate result is referred to as *minimal theories* (MTH).

Result 2 (MTH)

$$(L_1 \wedge C_{11}^{+-}) \leftrightarrow L_1^+, \tag{2(a)}$$
$$(L_2 \wedge C_{22}^{+-}) \leftrightarrow L_2^+, \tag{2(b)}$$
$$(R_2 \wedge \neg C_{22}^{+-}) \leftrightarrow R_2^+, \tag{2(c)}$$
$$(R_3 \wedge \neg C_{33}^{+-}) \leftrightarrow R_3^+. \tag{2(d)}$$

[6] For details see [24] and [25]. Note in particular that a correct formal notation of a minimal theory uses what Graßhoff et. al. [24, 25] call a *double conditional*.

From the logical relations (2(a)), (2(b)), (2(c)) and (2(d)) the following probabilities can be derived:

$$p(L_1^+ \wedge R_2^+) = p(L_1 \wedge C_{11}^{+-} \wedge R_2 \wedge \neg C_{22}^{+-}),$$
$$p(L_2^+ \wedge R_3^+) = p(L_2 \wedge C_{22}^{+-} \wedge R_3 \wedge \neg C_{33}^{+-}),$$
$$p(L_1^+ \wedge R_3^+) = p(L_1 \wedge C_{11}^{+-} \wedge R_3 \wedge \neg C_{33}^{+-}).$$

By NOWM (equations (30) and (31)) $p(L_1^+ \wedge R_2^+)$ is the same as $p(L_1^+ \wedge R_2^+ \wedge L_1 \wedge R_2)$ etc. and the above equations read:

$$p(L_1^+ \wedge R_2^+ \wedge L_1 \wedge R_2) = p(L_1 \wedge C_{11}^{+-} \wedge R_2 \wedge \neg C_{22}^{+-}), \quad (36)$$
$$p(L_2^+ \wedge R_3^+ \wedge L_2 \wedge R_3) = p(L_2 \wedge C_{22}^{+-} \wedge R_3 \wedge \neg C_{33}^{+-}), \quad (37)$$
$$p(L_1^+ \wedge R_3^+ \wedge L_1 \wedge R_3) = p(L_1 \wedge C_{11}^{+-} \wedge R_3 \wedge \neg C_{33}^{+-}). \quad (38)$$

D. No conspiracy

The events of type C_{ii}^{+-} are not supposed to be influenced by the measuring operations L_i and R_j. One reason for this assumption is that the measurement operations can be chosen arbitrarily before the particles enter the magnetic field of the Stern-Gerlach magnets and that an event of type C_{ii}^{+-} is assumed to happen *before* the particles arrive at the magnets. Therefore a causal influence of the measurement operations on events of type C_{ii}^{+-} would be tantamount to *backward causation.* Also an inverse statement is supposed to hold: The event types C_{ii}^{+-} are assumed not to be causally relevant for the measurement operations. This is meant to rule out some kind of "cosmic conspiracy" that whenever an event of type C_{ii}^{+-} is instantiated, the experimenter would be "forced" to use certain measurement operations. This *causal* independence between C_{ii}^{+-} and the measurement operations is assumed to imply the corresponding *statistical* independence. The same is assumed to hold also for *conjunctions* of common cause event types. We refer to this condition as *no conspiracy* (NO-CONS).

Assumption 9 (NO-CONS)

$$p(C_{ii}^{+-} \wedge \neg C_{jj}^{+-} | L_i \wedge R_j) = p(C_{ii}^{+-} \wedge \neg C_{jj}^{+-}). \quad (39)$$

By this condition of statistical independence the three probabilities considered above can be transformed. That is, we have, for instance

$$\begin{aligned} p(L_1^+ \wedge R_2^+ | L_1 \wedge R_2) &\doteq \frac{p(L_1^+ \wedge R_2^+ \wedge L_1 \wedge R_2)}{p(L_1 \wedge R_2)} \\ &\overset{\text{(i)}}{=} \frac{p(L_1 \wedge C_{11}^{+-} \wedge R_2 \wedge \neg C_{22}^{+-})}{p(L_1 \wedge R_2)} \\ &\doteq p(C_{11}^{+-} \wedge \neg C_{22}^{+-} | L_1 \wedge R_2) \\ &\overset{\text{(ii)}}{=} p(C_{11}^{+-} \wedge \neg C_{22}^{+-}) \\ &\overset{\text{(iii)}}{=} p(C_{11}^{+-} \wedge \neg C_{22}^{+-} \wedge C_{33}^{+-}) \\ &\quad + p(C_{11}^{+-} \wedge \neg C_{22}^{+-} \wedge \neg C_{33}^{+-}). \end{aligned}$$

The dotted equations are true by definition of conditional probability. In step (i) equation (36) was used. Step (ii) is valid by "no conspiracy" (equation (39)), and (iii) by a theorem of probability calculus, according to which $p(A) = p(A \wedge B) + p(A \wedge \neg B)$ for any A and B. Transforming the other two expressions in a similar way, we arrive at

$$p(L_1^+ \wedge R_2^+ | L_1 \wedge R_2) = p(C_{11}^{+-} \wedge \neg C_{22}^{+-} \wedge C_{33}^{+-}) + p(C_{11}^{+-} \wedge \neg C_{22}^{+-} \wedge \neg C_{33}^{+-}), \quad (40)$$
$$p(L_2^+ \wedge R_3^+ | L_2 \wedge R_3) = p(C_{11}^{+-} \wedge C_{22}^{+-} \wedge \neg C_{33}^{+-}) + p(\neg C_{11}^{+-} \wedge C_{22}^{+-} \wedge \neg C_{33}^{+-}), \quad (41)$$

$$p(L_1^+ \wedge R_3^+ | L_1 \wedge R_3) = p(C_{11}^{+-} \wedge C_{22}^{+-} \wedge \neg C_{33}^{+-}) + p(C_{11}^{+-} \wedge \neg C_{22}^{+-} \wedge \neg C_{33}^{+-}). \quad (42)$$

Since both terms of the right-hand side of the last equation appear in the sum of the right-hand sides of the first two equations, the following version of the Bell inequality (BELL) follows[7].

Result 3 (BELL)

$$p(L_1^+ \wedge R_3^+ | L_1 \wedge R_3) \leq p(L_1^+ \wedge R_2^+ | L_1 \wedge R_2) + p(L_2^+ \wedge R_3^+ | L_2 \wedge R_3). \quad (43)$$

This inequality has been empirically falsified, see e. g. [26].
The inequality was derived from the following assumptions.

- Perfect correlation (PCORR),
- separability (SEP),

[7] It was first derived in this form by Wigner [10].

- locality 1 (LOC1),
- principle of common cause (PCC),
- exactly one of exactly two possible outcomes (EX),
- locality 2 (LOC2),
- no outcome without measurement (NOWM),
- locality 3 (LOC3),
- no conspiracy (NO-CONS).

This is a version of Bell's theorem. It says: If these assumptions are true, the Bell inequality is true. The derivation of the Bell inequality presented here is an improvement on the usual Bell-type arguments, such as [5] and [11], in two respects: First, it does not assume a *common* common cause variable for different correlations. Second, contrary to the usual locality conditions, the ones assumed here do not presuppose a solution to the problems posed by the relation between causal and statistical (in)dependence (see e. g. [27]).

V. DISCUSSION

Our claim to have presented a minimal assumption derivation of a Bell-type inequality is relative: our set of assumptions is weaker than any set known to us from which a Bell-type inequality can be derived and that contains the assumption of *perfect* correlation (PCORR). It was one of the achievements of Clauser and Horne [12] to show that a Bell-type inequality can be derived also if the correlations of outcomes of parallel spin measurements are not assumed to be perfect. Our assumption of correlation is stronger than the one used by Clauser and Horne. However, they assume a *common* common cause variable for all correlations, which is a stronger assumption than our assumption of possibly different common cause variables for each correlation (PCC). We have not been able to derive a Bell-type inequality without assuming perfect correlation and allowing different common cause variables. If PCORR is indeed a necessary assumption for our derivation of the Bell inequality, it should be possible to construct a model in which PCORR does not hold (being violated by an arbitrary small deviation, say). Since the actually measured correlations are never perfect—a fact that is usually attributed to experimental imperfections—it is not obvious how such a model could be refuted.

Our notion of local causality might be challenged as follows. Even though nothing in PCC dictates that in general the common cause variables of different correlations have to be the same, there might be strong grounds for why they are the same in the context of the EPRB experiment. Indeed, Bell argued for his choice of local

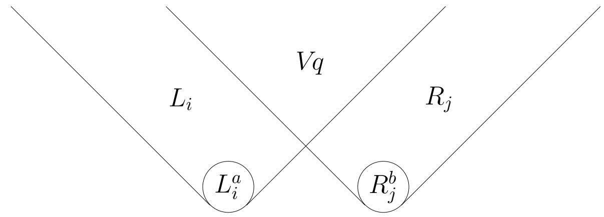

FIG. 3: The two backward light cones of two measurement outcomes. The total state of the overlap is taken to define a common common cause variable V, which can take on certain values q. Cf. [9, p. 55]

.

causality along the following lines.[8] Assume that L_i^a and R_j^b are positively correlated. Then

$$p(L_i^a|R_j^b \wedge L_i \wedge R_j) > p(L_i^a|L_i \wedge R_j). \qquad (44)$$

Since coinciding instances of L_i^a and R_j^b are space-like separated, neither is causally relevant for the other. Rather, the correlation should be explained by exhibiting some common causes in the overlap of the backward light cones of the coinciding instances. An instance of, say, L_i^a raises the probability of an instantiation of one of the common causally relevant factors, and this raises the probability of an instantiation of R_j^b. But *given the total state* of the overlap of the backward light cones of two coinciding instances, the probability of, say, R_j^b is assumed to be the same whether L_i^a is instantiated or not. If the total state of the overlap of the backward light cones is already given, nothing more that could be causally relevant for R_j^b can be inferred from an instance of L_i^a.

Along this line of reasoning the total state V of the overlap of the backward lightcones[9] of L_i^a and R_j^b is a common cause variable which screens off the correlation:

$$p(L_i^a \wedge R_j^b|L_i \wedge R_j \wedge Vq) = p(L_i^a|L_i \wedge R_j \wedge Vq) \times p(R_j^b|L_i \wedge R_j \wedge Vq). \qquad (45)$$

The common past Vq cannot be altered by choosing one or the other direction for the spin measurement—*"facta infecta fieri non possunt"* [18, p. 185]. Therefore the total state Vq of the common past is indeed a *common* common cause variable for all correlated outcomes, see FIG. 3.

This reasoning can be questioned along the following lines. It is reasonable that not all event types that are instantiated in the overlap of the backward light cones of two coinciding instances of the correlated event types are causally relevant for these latter event types. Therefore

[8] For a very good and more detailed discussion of this, see [13].

[9] One might argue that the total state of the *union* of the backward lightcones is a better candidate for a common cause variable [13]. The following discussion carries over also to this case.

conditionalizing on the total state is conditionalizing not only on the relevant factors but also on the irrelevant. Moreover, it is conceivable that which event types of the common past are relevant and which are not differs for different measurements. Claiming that the total state of the common past is a common common cause variable, one is thus committed to assume that

> "conditionalizing on all other events [...] in addition to those affecting [the correlated event types], does not disrupt the stochastic independence induced by conditionalizing on the affecting events." [13]

In particular in the light of *Simpson's paradox* [28] this assumption has been challenged [29]. Here, we will not assess arguments in favour of or against the possibility that conditionalizing on irrelevancies yields unexpected statistical dependencies. Our point is that by weakening the assumption in the way we did, our derivation is conclusive whatever may be the answer to this question.

Acknowledgments

We would like to thank Guido Bacciagaluppi, Miklós Rédei, Gabor Hofer-Szabó and Iñaki San Pedro for fruitful discussions.

[1] A. Einstein, B. Podolsky, and N. Rosen, Physical Review **47**, 777 (1935).
[2] D. Bohm, *Quantum Theory* (Prentice Hall, New York, 1951).
[3] J. S. Bell, Physics **1**, 195 (1964), reprinted in [9, p. 14].
[4] J. S. Bell, in *Foundations of Quantum Mechanics*, Proceedings of the International School of Physics 'Enrico Fermi' (New York, Academic, 1971), p. 171, reprinted in [9, p. 29].
[5] J. S. Bell, *The theory of local beables*, TH-2053-CERN (1975), presented at the Sixth GIFT Seminar, Jaca, 2–7 June 1975, reproduced in *Epistemological Letters*, March 1976, and reprinted in [9, p. 52].
[6] G. Hofer-Szabó, M. Rédei, and L. E. Szabó, British Journal for the Philosophy of Science **50**, 377 (1999), quant-ph/9805066.
[7] H. Reichenbach, *The Direction of Time* (University of California Press, Los Angeles, 1956).
[8] A. Wüthrich, *Quantum Correlations and Common Causes* (Bern Studies in the History and Philosophy of Science,, 2004).
[9] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987).
[10] E. Wigner, American Journal of Physics **38**, 1005 (1970).
[11] B. C. van Fraassen, in [30].
[12] J. F. Clauser and M. A. Horne, Physical Review D **10**, 526 (1974).
[13] J. Butterfield, in [30].
[14] J. P. Jarrett, Noûs **18**, 569 (1984).
[15] J. P. Jarrett, in [30], pp. 60–79.
[16] A. Shimony, *Search for a naturalistic world view*, vol. 2 (Cambridge University Press, 1993).
[17] G. Hofer-Szabó, M. Rédei, and L. E. Szabó, *Reichenbachian Common Cause Systems*, Forthcoming in *International Journal of Theoretical Physics*, URL `http://philsci-archive.pitt.edu/archive/00001246/`.
[18] T. Placek, *Is Nature determinstic?* (Jagiellonian University Press, Kraków, 2000), 1st ed.
[19] L. E. Szabó (1998), quant-ph/9806074.
[20] P. Suppes and M. Zanotti, in *Logic and Probability in Quantum Mechanics*, edited by P. Suppes (Dordrecht: Reidel, 1976), pp. 445–455.
[21] A. Fine, Synthese **50**, 279 (1982).
[22] L. Szabó (2000), quant-ph/0002030.
[23] L. Szabó and A. Fine (2002), URL `http://philsci-archive.pitt.edu/archive/00000642/`.
[24] G. Graßhoff and M. May, in *Current Issues in Causation*, edited by W. Spohn, M. Ledwig, and M. Esfeld (Paderborn: Mentis, 2001), pp. 85–114.
[25] G. Graßhoff and M. Baumgartner, *Kausalität und kausales Schliessen* (Bern Studies in the History and Philosophy of Science, forthcoming).
[26] A. Aspect, J. Dalibard, and G. Roger, Physical Review Letters **49**, 1804 (1982).
[27] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search* (Springer Verlag, 1993).
[28] E. H. Simpson, Journal of the Royal Statistical Society, Ser. B. **13**, 238 (1951).
[29] N. Cartwright, Noûs **13**, 419 (1979).
[30] J. T. Cushing and E. McMullin, eds., *Philosophical Consequences of Quantum Theory. Reflections on Bell's Theorem* (University of Notre Dame Press, Notre Dame, 1989).