Stochastic Outcomes in Branching Space-Time: Analysis of Bell's Theorem*

Tomasz Placek

Abstract

The paper extends the framework of outcomes in branching space-time (Kowalski, Placek 1999) by assigning probabilities to outcomes of events, where these probabilities are interpreted as weighted possibilities. In resulting models I define the notion of common cause of correlated outcomes of a single event, and investigate which setups allow for the introduction of common causes. It turns out that a deterministic common can always be introduced, but (surprisingly) only special setups permit the introduction of truly stochastic common causes. I analyze next the Bell-Aspect experiment and derive the Bell-CH inequalities. I observe that we postulate there not a common cause for outcomes of a single event but rather a *common* common cause that accounts for outcomes of many events, where 'events' mean 'measurements with (different) directions of polarization'. Since the inequalities are violated, I claim that no causal story can be told about the Bell correlations, where causality is subluminal and restricted by screening-off and locality conditions. I also prove an intimate link between the so-called quantum holism and the satisfaction of all the conditions used in the definition of common cause.

1 Introduction

Belnap's (1992) concept of branching space-time has recently been used in (Kowalski, Placek 1999) to obtain mathematically rigorous models of outcomes in branching

*The paper is a part of a larger project carried out with Thomas Müller. I would like to thank him for numerous helpful suggestions and comments. I also benefited from comments of audiences in Cracow and of Oxford philosophy of physics seminar. The project is partially supported by KBN grant 191/H01/97/12.

space-time (OBST), the principal algebraic result being that the family of outcomes of any fixed event constitutes a Boolean algebra. The models allow for a unified treatment of modality, space-time, and causal signalling, and thus, for an analysis of Bell's theorems. However, since they do not incorporate probabilities, we could only define the notion of non-stochastic common cause and then analyze the non-stochastic Bell-type arguments, i.e., GHZ theorems (Greenberger *et. al.* 1989 and 1990). We proved that a structure of spatiotemporal events can, mathematically speaking, always be enlarged by adding to it an event (common cause) that would screen off correlated outcomes of a *single* event. We also showed that in the setups of GHZ experiments, it is impossible to enlarge the corresponding structures of events so that an added event (*common* common cause) screened off correlated outcomes of more than one events.

My goal in this paper is to further extend the framework of OBST so that it will allow for the analysis of stochastic causation and, in particular, (stochastic) Bell-type theorems. Technically, this amounts to introducing probabilities to models of OBST and then defining the common cause, so that, given some additional assumptions, the Bell-type inequalities will be provable in these models. In OBST it is assumed that, for some events, if an event occurred, there are many histories containing this event. Now we want to assign weights to such histories, which amounts to assigning probabilities to some structures in models of OBST. Note that once we agreed on the openness of future, which means that the course of events may go one way or another, there should not be much resistance to the introduction of objective probabilities. Accordingly, our main task is a technical problem of appropriately defining probabilities on some structures of branching space-time. A simple and hence tempting idea is to identify the probability of an event with the ratio of the number of histories that contain the event in question to the number of all histories of some set. This, however, would tacitly assume that histories are equiprobable, which is rather counterintuitive. Thus, we had rather take advantage of the principal algebraic result of (Kowalski, Placek 1999), i.e., that the family of outcomes of any event constitutes a Boolean algebra, and assign probabilities to outcomes of events.[1] To introduce these probabilities we first need to recall same basic facts about outcomes in branching space-time.

2 Rudiments of outcomes in branching space-time

The point of departure are structures $\mathbf{W} = \langle W; \leq \rangle$, in which objects similar to possible worlds of the familiar Kripke semantics will be identified. W is a non-empty set, partially ordered by $\leq$, that consists of spatiotemporal points understood as concrete particulars. The relation $x \leq y$ is interpreted as meaning that x is in the backward light cone of y, which entails (barring some pathologies) that $\leq$, interpreted as above, is indeed a partial ordering. In a familiar way, we have: $x < y$ iff $x \leq y$ and $x \neq y$. This relation is used to introduce some ordering of subsets of W:

DEFINITION 1 (OF PRECEDENCE)
For $E, F \subseteq W$, $x \in W$ *(1)* $E \prec x$ *iff* $\forall_{e \in E}\ e < x$*; (2)* $E \prec F$ *iff* $\forall_{x \in F}\ E \prec x$.

To define histories, I need first the notion of *compatible* points:

[1]Note that this approach allows for assigning probabilities to histories as well, by first considering events that the history contains, and then multiplying probabilities of the atomic outcomes that the history passes through. It is rather doubtful, however, whether the notion of probability of a history can have any meaning but purely mathematical.

DEFINITION 2 (OF COMPATIBILITY)
$x, y \in W$ are upward compatible iff there is a $z \in W$ with $z \geq x$ and $z \geq y$. In standard mathematical usage sets composed entirely of upward compatible elements are called upward directed.

The expression $x \perp y$ means that x and y are *not* upward compatible; such elements are referred to as *orthogonal*.

Next, I introduce some special subsets of W, which will be called 'histories'. Intuitively, a history is to represent a possible course of events.

DEFINITION 3 (OF A HISTORY)
A subset h of W is a history iff h is an upward directed subset of W and for all upward directed $h' \subseteq W$ we have: $h' \supseteq h$ implies $h' = h$. In other words: histories are maximal upward directed subsets of W.

Since a course of events may go one way or another, histories split. Without yet knowing whether any exist, let us define the sets of *splitting points* for points in W.

DEFINITION 4 (OF SPLITTING POINTS FOR POINTS)
For any two orthogonal points $x, y \in W$, we define $C(x, y) \subseteq W$ by putting $z \in C(x, y)$ iff z is a maximal element in $\{z \in W : z \leq x \ \& \ z \leq y\}$. If x and y are not orthogonal, we put $C(x, y) = \emptyset$.

To ensure that for any pair of orthogonal points x and y, $C(x, y)$ is non-empty and that sets of splitting points behave "nicely", I assume these two conditions:

(C1) For any $x, y, z \in W$, if $x \perp y$ and $z \leq x$, $z \leq y$, then there is some $t \in C(x, y)$ with $t \geq z$.

(C2) For any $x, y, z, t \in W$, if $x \geq z$ and $y \geq t$, then $C(x, y) \supseteq C(z, t)$.

Now I define some notions that are required to introduce outcomes of events in W:

DEFINITION 5 (OF RELATIVE ORTHOGONALITY)
Elements x, y of W are orthogonal relative to E, written $x \perp_E y$, iff $E \prec x$, $E \prec y$ and $C(x, y) \cap E \neq \emptyset$.

DEFINITION 6 (OF ORTHOGONAL COMPLEMENT)
For $F \subseteq W$, the orthogonal complement of F relative to E is the set $F^{\perp_E}$ such that $x \in F^{\perp_E}$ iff $\forall_{y \in F}\ x \perp_E y$.

DEFINITION 7 (OF OUTCOME)
A subset F of W is an outcome of $E \subset W$ iff $F = F^{\perp_E \perp_E}$

This definition ensures that the outcome of E is preceded by E and is located as close as possible to E. Also, what the outcomes of E look like crucially depends on whether and, if so, how many, histories split in E.

Now, here is the main algebraic result, proved in (Kowalski, Placek 1999):

THEOREM 1
The family F_E of outcomes of $E \subset W$ forms a complete and atomic Boolean algebra

$$\mathcal{F}_E = \langle F_E, \cap, \cup, \perp_E, \Omega_E, \emptyset \rangle,$$

where $\cap$ and $\cup$ are the familiar set-theoretical operations, $\Omega_E = \{x \in W : E \prec x\}$, and $\emptyset$ is the empty set.

Since only specific subsets of W have non-trivial outcomes, these subsets will be dubbed 'events' and defined as follows:

DEFINITION 8 (OF EVENTS)
$E \subset W$ is an event if $E \neq \emptyset$ and $\exists_{x \in W}\ E \prec x$.

In short, the event is a subset of a history that is bounded from above. It can be proved that

LEMMA 1
$E \subset W$ is an event iff E has a non-empty outcome.

In what follows, I will frequently refer to *atomic* outcomes of events; hence the definition:

DEFINITION 9 (OF ATOMIC OUTCOMES)
e is an atomic outcome of $E \subset W$ iff (1) e is a non-empty outcome of E and (2) there is no non-empty outcome u of E such that $u \subset e$.

I end this section by quoting the definition of space-like events that will be needed to introduce (stochastic) common causes:

DEFINITION 10 (OF SPACE-LIKE EVENTS)
The set $\{E_1, E_2, E_3, \ldots, E_n\}$ of events is space-like only if

- *$\bigcup_{i=1}^{n} E_i$ is an event;*
- *E_i does not overlap with any outcome of E_j, i.e. for all $i, j \leq n$,*

$$E_i \cap \Omega_{E_j} = \emptyset.$$

3 Probabilities

Since the family F_E of outcomes of event E constitutes a complete and atomic Boolean algebra:

$$\mathcal{F}_E = \langle F_E, \cap, \cup, \perp_E, \Omega_E, \emptyset \rangle,$$

to obtain a classical σ-additive probability space $\{\mathcal{F}_E, \mu\}$ it suffices to equip the algebra with a measure $\mu : F_E \to [0, 1]$ that satisfies

$$\mu(\Omega_E) = 1 \qquad \text{where} \qquad \Omega_E = \{x \in W : E \prec x\}$$

$$\text{If } \forall_{A_i, A_j \in F_E, i \neq j}\ A_i \cap A_j = \emptyset, \text{ then } \mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

Within this probability space we have probabilities of outcomes of E and conditional probability that an outcome of E takes place given that another outcome of E takes place. To clarify the notation, I shall write $\mu_E(x)$ for the probability of outcome x of E and $\mu_E(x \mid y)$ for the probability that outcome x of E happened given that outcome y of E happened, where the latter concept is defined in the familiar way as

$$\mu_E(x \mid y) = \begin{cases} \frac{\mu_E(x \cap y)}{\mu_E(y)} & \text{iff } \mu_E(y) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

The introduction of probabilities to the model of OBST has an effect of making the resulting model more fine-grained. Mathematically speaking, the model of stochastic OBST is a quadruple $\langle W, \leq, \mathcal{E}, \Upsilon \rangle$, where W is non-empty set, $\leq$—partial ordering, $\mathcal{E}$—the set of events in W, and Υ—the collection of probability measures, where these probability measures are defined on algebras of outcomes of events from $\mathcal{E}$.

Note that the approach allows only for assigning probabilities to outcomes of events. In particular, probability is not defined on combinations of outcomes of *different* events. Any model has a collection of probability spaces and at this stage there are no constraints whatsoever on how the probability measures are related one to another. Ideally the models should incorporate such constraints, a good idea being to mimic probabilistic conditionalization in this framework. Rather than pursuing here this project, I will limit myself to one simple constraint. In some rather special cases, I will require that probability measures μ_E and μ_F corresponding to events E and F, respectively, be related to probability measure $\mu_{E \cup F}$ corresponding to the sum of events E and F. The requirement, to be specified soon, relies on the following two facts.[2]

FACT 1 *If e and f are outcomes of E and F, respectively, then $e \cap f$ is an outcome of $E \cup F$.*
If a is an outcomes of $E \cup F$, then there are outcomes e of E and f of F such that $a = e \cap f$.

FACT 2 *If e and f are atomic outcomes of E and F, respectively, then $e \cap f$ is either empty or an atomic outcome of $E \cup F$.*

To simplify the calculations, it is henceforth assumed that every family of outcomes is finite. Suppose now E and F are events with outcomes $e_1, \ldots, e_N$ and $f_1, \ldots, f_M$, respectively. Let us also assume that "no history escapes", that is, every history containing E, contains F, and the other way round. This obviously guarantees that $E \cup F$ is an event. More importantly, this assumption holds only if $e_i \cap f_j \neq \emptyset$, for any non-empty e_i and f_j. Combining this observation with the facts above yields:

REQUIREMENT
Given that E and F are events with outcomes $e_1, \ldots, e_N$ and $f_1, \ldots, f_M$, respectively, and such that $e_i \cap f_j \neq \emptyset$ for any non-empty e_i and f_j, it is required that:

$$\mu_E(e_i) = \sum_{f_j} \mu_{E \cup F}(e_i \cap f_j) = \mu_{E \cup F}(e_i \cap \Omega_F)$$
$$\mu_F(f_j) = \sum_{e_i} \mu_{E \cup F}(e_i \cap f_j) = \mu_{E \cup F}(\Omega_E \cap f_j),$$

where the summation is over atomic outcomes only. The right hand side equations follow from the fact that Ω_F (Ω_E) is equal to the sum of all atomic outcomes of F (E); since these atoms are disjoint, the rest follows from the definition of measure. Although REQUIREMENT closely resembles what is known in probability theory as marginal

[2]For proofs of Fact 1 and Fact 2, see Appendix.

property, I am going to consider it a semi-empirical claim capable of refutation. Moreover, there is every good reason to accept it as applied to observable probabilities—no experiment has given even the slightest evidence against REQUIREMENT as applied to observable probabilities. However, in what follows REQUIREMENT will frequently be applied to unobservable probabilities, where one might be more sceptical as to its validity. This scepticism will probably be reinforced by noting that this condition entails some premises of Bell's theorem. For the sceptic of this kind, it is however extremely difficult to explain how, in some hidden layers of reality, REQUIREMENT could possibly fail, its intuitive meaning being "no history escapes".

Apart from being frequently applied in the construction of common causes, REQUIREMENT legitimizes this simplification of the notation: I shall frequently skip the reference to outcome $\Omega_{E(F)}$, and write:

$$\mu_{E\cup F}(e) \text{ for } \mu_{E\cup F}(e \cap \Omega_F)$$

and similarly for $\mu_{E\cup F}(e)$.

Before we proceed, let us assume a convention. The result of an experiment, intuitively, is not an outcome (in technical sense) of a corresponding event E, as the outcome is open to the future. However, it is reasonable to assume that atomic outcomes e_i of E correspond one-to-one to results r^{ei}, the latter being a click of a detector, a scintillation or the like. Thus, any two different results are always in two different *atomic* outcomes of E and every non-empty outcome of E contains one or another result of the experiment. Thus, although the experimenter measures probabilities of clicks or scintillations rather than probabilities of outcomes, these probabilities are mathematically the same. Accordingly, probabilities assigned to outcomes are here understood as the probabilities of the corresponding results. Hence, whenever confusion is not likely, I will refer to probabilities of outcomes, meaning also probabilities of corresponding results.

Now, the researcher carrying out a typical correlation experiment aims to find long run frequencies of results $e_i \cap f_j$ of the "double" measurement $E \cup F$ and long run frequencies of results e_i (f_j) of the single measurement E (F), where E and F can be thought of as the event of measuring a quantity at one apparatus, and the event of measuring a quantity at the other apparatus, respectively. The former frequencies are represented by probabilities of the form $\mu_{E\cup F}(e_i \cap f_j)$. As for the latter frequencies, they result from ignoring the results at the other wing of the experiment. Mathematically this means that the probability of obtaining the result e_i, given measurements E and F, is

$$\mu_{E\cup F}(e_i) = \sum_{f_j} \mu_{E\cup F}(e_i \cap f_j) = \mu_{E\cup F}(e_i \cap \Omega_F),$$

where the summation is over atomic f_j only (an analogous formula holds for $\mu_{E\cup F}(f_j)$ as well).

One may wonder why there is still a lingering reference to F in the subscript on the left hand side, but this is a consequence of working with a measure defined on the family of outcomes of $E \cup F$. Besides, in a correlation experiment, while considering results of one apparatus, you may still have the other apparatus turned on, its indications being neglected. Of course, you may as well switch off the other apparatus, or even remove it, and then measure $\mu_E(e_i)$. Nevertheless, if the experimenter wants to use this probability to model the correlation experiment, he must check that $\mu_E(e_i) =$

$\mu_{E\cup F}(e_i)$, its meaning being: the other apparatus does not interfere with the nearby results. Moreover, since this is a correlation experiment, the experimenter works to ensure that in every run of the experiment, a result is produced in one region and a result is produced in the other region. Accordingly, every non-empty outcome of E has a non-empty intersection with any non-empty outcome of F. Thus, by REQUIREMENT, we may identify:

$$\mu_{E\cup F}(e_i) = \mu_E(e_i) \qquad \mu_{E\cup F}(f_i) = \mu_F(f_i)$$

Moreover, in line with what was said above, these identifications are assumed to be empirical. In other words, by performing two kinds of the experiment, with the distant apparatus switched on and switched off, we may see if in a given setup our REQUIREMENT holds.

Thus, for basic frequencies the experimenter measures, I take frequencies of outcomes of a "double" measurement. Other frequencies are recalculated on the ground of these findings. This means that "canonical" probability spaces are spaces generated by sums of two "measurement events", i.e., the probabilities I will work with are of the form:

$$\mu_{E\cup F}(e_i \cap f_j).$$

4 The idea of conditioning on a common cause

Suppose in a correlation experiment, with E and F standing, respectively, for events of measuring some quantities at two (possibly distant) apparatuses, particular results $e_i \cap f_j$ are counted. The observation is that neither all the runs nor none of the runs result in $e_i \cap f_j$, which means (on the level of probabilities) that $0 \neq \mu_{E\cup F}(e_i \cap f_j) \neq 1$. Importantly, the results are correlated, i.e.,

$$\mu_{E\cup F}(e_i \cap f_j) \neq \mu_{E\cup F}(e_i) \cdot \mu_{E\cup F}(f_j) = \mu_{E\cup F}(e_i \cap \Omega_F) \cdot \mu_{E\cup F}(\Omega_E \cap f_j).$$

There are, moreover, grounds to think that initial segments of these runs,[3] are exact replicas, exactly the same fragments of the world. Now, the idea of hidden variables is a break with this belief. One says that the runs are differentiated by some hidden factors. To be as clear as possible, let us talk only of two differentiating factors, say V and H. The runs (you think now) are divided into two sorts, those with a V factor and those with a H factor. In a similar vein, the cases of $e_i \cap f_j$ are also of two sorts as well as the cases of observing no $e_i \cap f_j$. The momentous decision to be made now is how you want to think of the effects of these two factors, V and H. You may give way to your deterministic inclinations and say that all the runs with, say, the V factor result in $e_i \cap f_j$ whereas all the runs with H factor do not produce $e_i \cap f_j$. In the terminology of hidden-variables theories, these models will be called "determinate-value (deterministic) hidden variable models". The other option is to require that the factors will "clean" the statistics, that is, in each population of runs, V population and H population, e_i and f_j will be stochastically independent. In the terminology just mentioned, these models are dubbed "stochastic hidden-variable models". Since this option is more general, I shall investigate rather this than the deterministic case.

A natural way to translate the concept of differentiating factors into the framework of OBST is to assume that the factors are outcomes of some event C, which will be called 'common cause'. With this idea, some runs are different because they are actualized in different histories, that is, histories passing through different outcomes of C.

[3] 'Initial' means 'before the registration of results'.

Thus, there is a common cause for correlated e_i and f_j if the initial model of OBST can be enlarged by adding a properly located event C and in the resulting model specific probabilistic constraints hold. In this enlarged model we want all results r^{ei} and r^{fj}, which correspond to atomic outcomes e_i and f_j, respectively, to be differentiated into two (or more) varieties (see Figure 1). As an effect of this differentiation, each atomic outcome e_i (or f_j) will be transformed into the non-atomic outcome e_i^* (or f_j^*), containing all the multiplied results r_k^{ei} (or r_k^{fj}). The resulting outcome of E (or of F) is extended in respect to the initial outcome, where this notion is defined as follows:

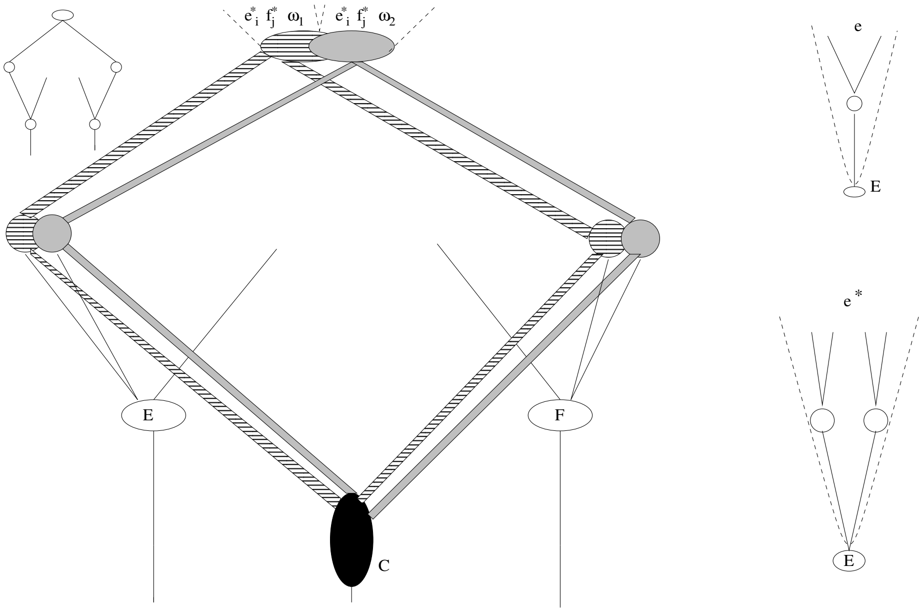

Figure 1: The initial model (in the left upper corner) and the enlarged model. Event C has two outcomes: with the grey pattern (ω_1) and the horizontal pattern (ω_2). Event $E \cup F \cup C$ has also two outcomes: grey $e_i^* \cap f_j^* \cap \omega_1$ and horizontal $e_i^* \cap f_j^* \cap \omega_2$. The unfinished lines point to other possible outcomes. On the right, outcome e of E (up) and the extended outcome e^* of E (down). The lines represent the relation of precedence, $\prec$. Although C does not precede E (F), this does not preclude C exerting causal influence on E (F), as a part of the former can precede the latter.

DEFINITION 11 (OF EXTENDED OUTCOME)
In model $\langle W, \leq, \mathcal{E}, \Upsilon \rangle$ let E be an event and e—its atomic outcome that corresponds to result r^e. To construct a K-extended outcome e^* in an enlarged model, introduce K varieties of the result, $r_1^e, \ldots r_K^e$, such that any two r_m^e, r_n^e ($m \neq n$ and $m, n \leq K$) split in E, i.e.,

$$\forall_{x \in r_m^e} \forall_{y \in r_n^e} C(x, y) \cap E \neq \emptyset.$$

The outcome e^* of E that is extended in respect to e is

$$e^* \stackrel{df}{=} \bigcup_{k \leq K} r_k^{e \perp_E \perp_E}$$

□

Note that each $r_k^{e\perp_E\perp_E}$ is an atomic outcome of E. Thus, the extended outcome e^* results from summing all atomic outcomes corresponding to K-multiplied results r_k^e. Turning next to space-like events E and F, by extending their respective atomic outcomes e and f, a similar extension is induced in respect to outcomes of $E \cup F$. To guarantee that in this enlarged model the resulting outcomes of E and F mesh properly, I require that for non-empty $e \cap f$

$$r_m^{e\,\perp_E\perp_E} \cap r_n^{f\perp_F\perp_F} \neq \emptyset \iff n = m.$$

A consequence of this requirement is that each initial atomic outcome $e \cap f$ of $E \cup F$ is replaced by a collection of atomic outcomes, all of them being of the form $r_k^{e\perp_E\perp_E} \cap r_k^{f\perp_F\perp_F}$, where r_k^e and r_k^f are K-multiplied results r^e and r^f, respectively. This collection can be summed to obtain an extended outcome of $E \cup F$, extended in respect to $e \cap f$. It remains only to relate the differentiated outcomes of E, F, and $E \cup F$ to atomic outcomes of common cause C. This can be done by first requiring that the multiplication number K is the same as the number of atomic outcomes ω_k of C, and then demanding that, for any $k \leq K$

$$r_n^e \subset \omega_k \Leftrightarrow n = k \text{ and } r_n^f \subset \omega_k \Leftrightarrow n = k.$$

This requirement entails that C precedes all results $r_1^e \ldots r_K^e$ and $r_1^f \ldots r_K^f$—after all, C is a common cause. It also guarantees that a history that contains C passes through $e^* \cap f^*$, which in turn entails that outcome $e^* \cap f^* \cap \omega_k$ of $E \cup F \cup C$ is non-empty, for some outcome ω_k of C, $k \leq K$. It is provisionally assumed that E and F are not differentiated and C does not precede measurement events E and F; these limitations will be soon discussed, however. An enlarged model, for $K = 2$ and with C taking care of correlated e and f is displayed in Figure 1. The two outcomes of C are indicated there by different patterns.

5 Common cause

As was argued above, to find a common cause of correlated outcomes in a model of stochastic OBST is to enlarge this initial model by introducing a properly located event C and extending the initial outcomes, the probabilities in the enlarged model satisfying some constraints. In this section I will specify these constraints and define the notion of common cause for (possibly many) correlated outcomes of an event.

DEFINITION 12 (OF COMMON CAUSE)
Let $\{E, F\}$ be a space-like set in model $\langle W, \leq, \mathcal{E}, \Upsilon\rangle$ and let $e_1, \ldots, e_N$ be atomic outcomes of E and $f_1, \ldots f_M$ are atomic outcomes of F, which correspond to results $r^{e1}, \ldots, r^{eN}$ and $r^{f1}, \ldots r^{fM}$, respectively. Let also some outcomes of $E \cup F$ be correlated, i.e., for some (possibly many or even all) $i \leq N$ and $j \leq M$:

$$\mu_{E\cup F}(e_i \cap f_j) \neq \mu_{E\cup F}(e_i) \cdot \mu_{E\cup F}(f_j).$$

A common cause C that accounts for these correlations is an event with K atomic outcomes ω_k in a model $\langle W', \leq', \mathcal{E}', \Upsilon'\rangle$ that is enlarged in respect to $\langle W, \leq, \mathcal{E}, \Upsilon\rangle$ by:

1. *K-multiplying results r^{ei} and r^{fj} to obtain r_k^{ei} and r_k^{fj}, respectively, for all $k \leq K$, $i \leq N$ and $j \leq M$;*

2. *K-extending outcomes e_i and f_j into e_i^* and f_j^*, respectively, for all $i \leq N, j \leq M$;*

3. *requiring that, for each atomic outcome $e_i \cap f_j$ of $E \cup F$*

$$(r_n^{ei})^{\perp_E \perp_E} \cap (r_m^{fj})^{\perp_F \perp_F} \neq \emptyset \;\Leftrightarrow n = m;$$

4. *adding an event C with K atomic outcomes.*

The common cause C should satisfy:

$$\forall_{i \leq N,\, j \leq M,\, k \leq K}\; C \prec r_k^{ei} \wedge C \prec r_k^{fj}; \qquad \text{(Causal Precedence)}$$

Every history that contains C, contains E and F; *(Conservation)*

$$\forall_{i \leq N, j \leq M, k \leq K}\; \overset{*}{\mu}_{E \cup F \cup C}(e_i^* \cap f_j^* \mid \omega_k) = \overset{*}{\mu}_{E \cup F \cup C}(e_i^* \mid \omega_k) \cdot \overset{*}{\mu}_{E \cup F \cup C}(f_j^* \mid \omega_k),$$

where $\overset{*}{\mu}_{E \cup F \cup C}(\omega_k) = \overset{*}{\mu}_{E \cup F \cup C}(\Omega_E^* \cap \Omega_F^* \cap \omega_k)$; *(Screening-off)*

$$\forall_{i \leq N, j \leq M}\; \mu_{E \cup F}(e_i \cap f_j) = \overset{*}{\mu}_{E \cup F}(e_i^* \cap f_j^*). \qquad \text{(Adequacy)}$$

To comment on the clauses above, Screening-off is a part of Reichenbach's (1956) principle of the common cause. Causal Precedence guarantees that if only $e_i \cap f_j$ is non-empty, then for some atomic ω_k, $e_i^* \cap f_j^* \cap \omega_k$ is also non-empty. Conservation allows for strengthening this relation to: if only $e_i \cap f_j$ is non-empty, then for *any* atomic ω_k, $e_i^* \cap f_j^* \cap \omega_k$ is also non-empty. Thus the two conditions make the model satisfy the premises of REQUIREMENT[4], so for any non-empty $e_i \cap f_j$ we have:

$$\overset{*}{\mu}_{E \cup F}(e_i^* \cap f_j^*) = \overset{*}{\mu}_{E \cup F \cup C}(e_i^* \cap f_j^* \cap \Omega_C).$$

Then Adequacy says that no matter how the initial model is enlarged, it should yield the same observable probabilities. Therefore, it identifies $\overset{*}{\mu}_{E \cup F}(e_i^* \cap f_j^*)$ (in the enlarged model) and $\mu_{E \cup F}(e_i \cap f_j)$ (in the initial model).

Note some peculiarities of the definition. It recommends the extention of even the outcome of, say, E that has empty intersections with any outcome of F. It also requires placing C below even such multiplied results r_k^{ei} that result r^{ei} (in the initial model) does not share a history with any result r^{fj}. Finally, Screening-off concerns correlated outcomes as well as non-correlated outcomes. Nevertheless, these peculiarities do not seem to do any mathematical harm, and will not arise in many models, in particular, in models of the Bell-Aspect experiment.

[4]But note that in this application REQUIREMENT is not testable.

6 Do common causes exist?

Naturally we want to know whether common causes exist in setups fulfilling the assumptions of the definition of common cause. To investigate this, let $\{E, F\}$ be space-like and E and F have, respectively, outcomes $e_1 \dots e_M$ and $f_1 \dots f_N$, where these outcomes correspond to results $r^{e1} \dots r^{eM}$ and $r^{f1} \dots r^{fN}$, respectively. Finally, some (possibly all) outcomes of E and F are correlated, i.e. for some $i \leq M, j \leq N$,

$$\mu_{E \cup F}(e_i \cap f_j) \neq \mu_{E \cup F}(e_i)\mu_{E \cup F}(f_j).$$

Now, while enlarging this initial model, results r^{ei}, r^{fj} should be K-multiplied to obtain r^{ei}_k, r^{fj}_k (for all $i \leq M, j \leq N, k \leq K$) and outcomes e_i, f_j correspondingly extended into e^*_i, f^*_j. Next, it should be ensured that $(r^{ei}_k)^{\perp_E \perp_E} \cap (r^{fj}_l)^{\perp_F \perp_F} \neq \emptyset$ iff $l = k$. Then C should be made to precede r^{ei}_k and r^{fj}_k. Finally, to satisfy Conservation no other lines representing precedence should be drawn. Figure (2) displays this procedure for $M = N = K = 2$. The figure intends to represent EPR-like correlations, with E and F standing for events of measuring (by two distant apparatuses) spin projections of pairs of particles, where directions of polarization set in two apparatuses are fixed and not necessarily parallel.

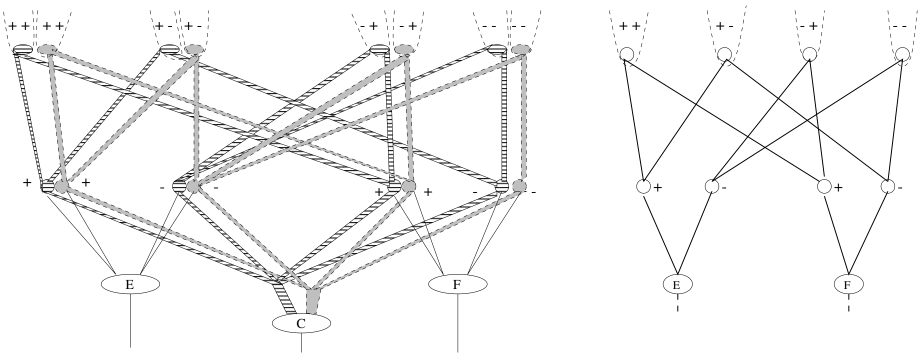

Figure 2: The enlarged model (left) and the initial model (right). Each event E and F has initially two atomic outcomes that correspond either to result $+$ or to result $-$, each of these outcomes being extended in the enlarged model. C has two atomic outcomes, with the grey pattern (ω_1) and the horizontal pattern (ω_2). These outcomes differentiate the initial outcomes of $E \cup F$ into grey $e^*_i \cap f^*_j \cap \omega_1$ and horizontal $e^*_i \cap f^*_j \cap \omega_2$.

In the next step, we must ensure that probabilistic constraints of Adequacy and of Screening-off are satisfied. To simplify the notation, put

$$\mu_{E \cup F}(e_i \cap f_j) = q_{ij} \qquad \overset{*}{\mu}_{E \cup F \cup C}(e^*_i \cap f^*_j \cap \omega_k) = p_{ijk}$$

Adequacy (via Causal Precedence, Conservation and REQUIREMENT) yields $N \cdot M$ formulas:

$$q_{ij} = \sum_{k}^{K} p_{ijk} \tag{1}$$

and Screening-off unwinds into:

$$p_{ijk} \sum_{m,n}^{M,N} p_{mnk} = (\sum_{m}^{M} p_{imk})(\sum_{m}^{N} p_{mjk}), \tag{2}$$

which can be more conveniently re-written as

$$\sum_{m\neq i, n\neq j}^{M,N} (p_{ijk}p_{mnk} - p_{ink}p_{mjk}) = 0 \tag{3}$$

Putting now $K = M \cdot N$, there are always these solutions of Eqs. (1) and (3):

$$p_{ijk} = q_{ij}\, \delta^{k}_{i+(j-1)M} \tag{4}$$

This implies that for any $i \leq M, j \leq N, k \leq K$,

$$\overset{*}{\mu}_{E\cup F\cup C}(e_i^* \cap f_j^* \mid \omega_k) = 0 \text{ or } 1.$$

Thus, the solutions entail a kind of determinism: if we are in outcome ω_k of C, then there is exactly one outcome of $E \cup F$ that is bound to happen, the chances of any other outcome of $E \cup F$ to happen being null.[5] The existence of these solutions means that determinism is always a mathematically viable option for outcomes of a single event of the form $E \cup F$. Nevertheless, this option is neither philosophically nor physically interesting. On the statistical level, it says that one can always partition the set of runs of the correlation experiment into those in which $e_i \cap f_j$ has been observed and those in which $e_i \cap f_j$ has not been observed. Thus, a more interesting question is whether Eqs. (1) and (3) allow for non-deterministic solutions. Are there solutions of Eqs. (1) and (3) for $K < M \cdot N$? Or, are there truly stochastic solutions p_{ijk}? The last problem resembles the question of whether the existence of a determinate-value (deterministic) hidden variable model is equivalent to the existence of a stochastic factorizable hidden variable model, and answered positively by Fine (1982). However, in the discussion of this equivalence, it is simply assumed that a model with trivial probabilities only (i.e., a deterministic model) counts as a stochastic model, which makes the implication from deterministic model to stochastic model trivial. A philosophically interesting question remains, namely whether probabilistic causation subject to screening-off substantially relaxes the constraint of determinism. What was called above "truly stochastic model" is a model in which some conditional probabilities are non-trivial. In the framework of stochastic OBST this amounts to saying that for any given k at least two probabilities p_{ijk} are non-zero.

Let us thus see if there are solutions p_{ijk} of Eqs. (1) and (3) such that for any given k at least two probabilities p_{ijk} are non-zero. Note that in general the equations may yield $M{\cdot}N + M{\cdot}N{\cdot}K$ constraints, the number of variables p_{ijk} being $M{\cdot}N{\cdot}K$. However, in the following cases of $\langle M, N\rangle$: $\langle 1, N\rangle$, $\langle M, 1\rangle$, $\langle 2, 3\rangle$, and $\langle 3, 2\rangle$, the number of non-trivial constraints imposed by Eqs. (3) is substantially smaller than $M{\cdot}N\cdot K$. This fact allows for truly stochastic solutions in these cases.[6] In all the other cases, in general there may be no truly stochastic solutions p_{ijk}. The simplest case of this kind is $M = N = 3$ (see Appendix).

[5] In the hidden-variable terminology, this is a determinate-value hidden variable model.

[6] The fact that there are solutions of Eqs. (1) and (3) for $N = M = 2$ squares nicely with the existence of a hidden variable model for EPR-like correlations, where the directions of polarization are fixed—see (Bell 1964).

Thus, surprisingly, a deterministic common cause can always be posited for correlated outcomes of any event of the form $E \cup F$, but a truly stochastic common cause is viable only for rather special correlated outcomes of an event of this kind. Contrary to the received wisdom, it is easier to accommodate determinism than truly probabilistic causation with screening-off!

7 A model for the Bell-Aspect experiment

My aim now is to attempt a construction of a common cause for the correlations observed in the Bell-Aspect experiment.[7] Let us assume two possible polarization directions α, σ in the left (L) apparatus and two possible polarization directions β, γ in the right (R) apparatus. By $L\alpha$ ($L\sigma$) and $R\beta$ ($R\gamma$). I will mean events of measuring spin projections on the corresponding directions on the left and on the right, respectively. These events, in turn, allow one to build events consisting of measuring a spin projection on the left and measuring a spin projection on the right, like for instance $L\alpha \cup R\beta$. In sum, there will be four such events: $L\alpha \cup R\beta$, $L\alpha \cup R\gamma$, $L\sigma \cup R\beta$, and $L\sigma \cup R\gamma$. For each of these events, correlations between outcomes are found. In order to explain these, the present enterprise goes much further than the familiar quest for a common cause of correlated outcomes of a *single* event. In essence, we are asking for one common cause that should take care of outcomes of the four events above. Such a common cause will be dubbed here, after (Hofer-Szabó, Rédei, and Szabó 1999), *common* common cause. My strategy is to show that if there is a common common cause in an enlarged model, then Bell-Clauser-Horne (Bell-CH) inequality[8] is provable in this model. But since quantum mechanics predicts this inequality to be violated, and almost certainly the inequality is also experimentally violated, this is a *reductio ad absurdum* of the existence of the common common cause in the set-up considered. To continue with the notation, $+$ and $-$ stand for results, while $L+$ ($L-$), $R+$ ($R-$) stand for outcomes of measuring a spin projection on the left and on the right, respectively, where the subscript under the probability μ will indicate which polarizations are considered. Finally, A_L and A_R stand for events of selecting the direction settings in the left and right apparatus, resp. Outcomes of the events involved are correlated, that is, we have many (possibly as many as 16) formulas like this:

$$\mu_{L\alpha \cup R\beta}(L+ \cap R+) \neq \mu_{L\alpha \cup R\beta}(L+)\mu_{L\alpha \cup R\beta}(R+).$$

Let us tell a story about what is happening in one wing of the set-up. There is an apparatus there, in which one of the two polarization settings can be chosen. Clearly, an apparatus before the selection can be viewed as an event. Since the selection is free, there is a split in this event, so that there are two possible measurements (i.e., on different directions of polarization), both preceded by the "selection" event. Next, no matter which selection has been made, the measurement yields one of two possible results, $+$ or $-$. Finally, a result from one apparatus is observed together with a result from the other apparatus, and observation of many runs of the experiment shows a correlation.

To construct a stochastic OBST model for the experiment considered, one should read off the precedence relations from the corresponding space-time diagram. Consider thus the left hand side of Figure 3. Clearly, the "double" result, here $++$, is preceded

[7] See (Aspect *et al* 1982) and (Bell 1964).

[8] The inequality was first derived in (Clauser and Horne 1974).

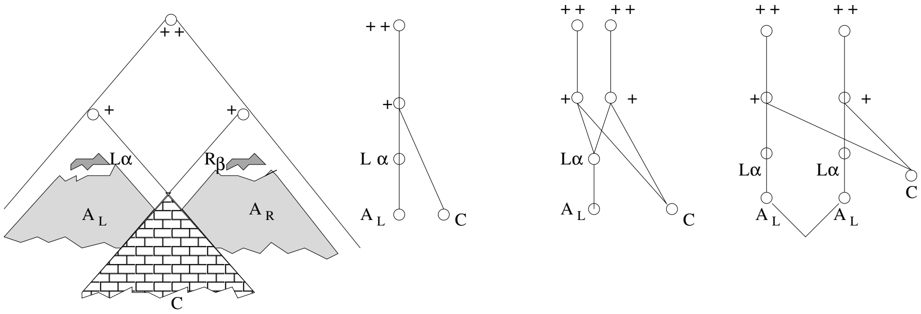

Figure 3: On the left, a space-time diagram for the Bell-Aspect setup. To the right, the precedence relations read from the diagram. On the right, two possible locations of splittings. It is the model to the left (no hidden states of the apparatus) that is described in this paper.

by $+$ on the left and $+$ on the right. Next, to focus our attention on the left wing of the setup, $+$ on the left is preceded by the measurement event with setting α, which in turn is preceded by the event A_L of selecting the settings. Turning to common cause C, since it must precede results at the two wings, the largest region one can take is the common past of $+$ on the left and $+$ on the right. One may want to cut it from below, as on the diagram. Clearly, C precedes pluses in the two wings and the "double" result $++$. However, it is neither the case that $C \prec A_L$ nor that $C \prec L\alpha$. This obviously does not mean that C cannot causally influence A_L or $L\alpha$, as a proper part of C precedes these events. This however is irrelevant for what the outcomes of C are. The precedence relations are displayed immediately right to the space-time diagram of Figure 3. To account for the "white" regions below the level of $L\alpha$ one should introduce new events that precede $L+$, but have no precedence relations with $L\alpha$ and A_L.

The next decision is where the splittings should be located. The general idea is that each: $+$ on the left, $+$ on the right, and $++$ are differentiated by outcomes of C in K varieties, each variety belonging to different atomic outcome of C. Thus, we must have splittings in C that yield K atomic outcomes of this event. However, the model requires more splittings, and there does not appear to be any clear ground for locating them in one place rather than some other. The conservative option taken in this paper says that although results in one wing, say $+$ in the left, are differentiated into K varieties, neither the measurement event, say $L\alpha$, nor the event A_L of the selection of settings is differentiated. This means that in all the relevant runs of the experiment we consider all instances of $L\alpha$ to be identical and all instances of A_L also to be identical. Of course, this concerns as well all the other measurement events and the events of the selection of settings. The location of splittings in this option is depicted in the middle of Figure 3. Another option is to assume that because of some hidden states of the apparatus, already the "selection" events A_L and A_R are differentiated, and consequently, the differentiation carries over to all the four measurement events $L\alpha$, $L\sigma$, $R\beta$, and $R\gamma$, this idea being illustrated in the far right part of Figure 3. The first option appears to be mathematically simpler, which is the reason it is taken in this paper. But the derivation of the Bell-CH inequality can be carried out in models with other locations of splittings, though to this end the definitions of extended outcomes and common cause should be appropriately modified, one modification being discussed in Appendix.

To see how the setup is represented in the present framework, consider Figure 4. The figure also shows the location of a stipulated common common cause C. The source of the particles is not indicated, since it will be treated as a part of C.

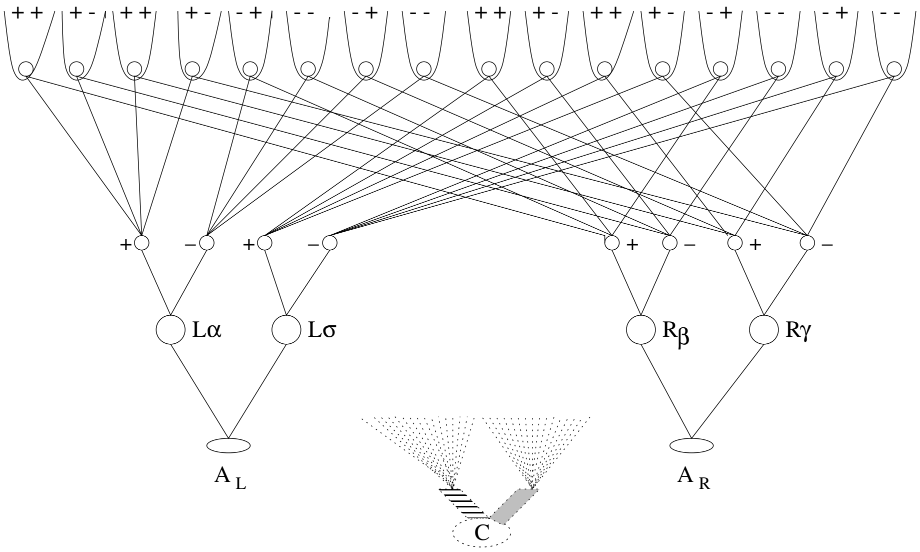

Figure 4: The model for the setup of the Bell-Aspect experiment, with the location of C indicated.

Contemplation of this figure brings us to the probably most intricate problem of the experiment, modality. The way the experiment is usually described suggests that there are two kinds of modality. On the one hand, we have selections of settings on both the apparatuses, the assumption being that they are done by a *free* agent (the experimenter) or produced by a "*random*" device.[9] Most significantly, while postulating a cause for the correlations one wants to retain this freedom: a postulated cause should not have any influence on which settings are selected. In the framework of stochastic OBST this means that selection of settings is independent of which outcome of C is actualized. As a justification of this belief, I am inclined to argue that it is not, and even cannot, be experimental, since no truly random *physical* device can be build. No matter how hard we try, there will be doubts as to whether the underlying physical processes are truly random. The status of the requirement is rather that of a condition on the possibility of experimental science. Were it not true, nature could be systematically hiding its secrets, by directing the experimenter what to measure on any given occasion. Turning now to the other kind of modality, it is concerned with results, and amounts to saying that given the setup, both results $+$ and $-$ are possible. This is the modality that we want to either eliminate by postulating deterministic causes or substantially tame by introducing probabilistic causes. Observe thus that our approach to the two kinds of modalities involved is somewhat contrary: asking for explanation of the experiment, we want a causal explanation of correlated results, but refuse to admit that selections of settings have causes other than free agents or random devices.

[9]In the actual experiment a semi-random device was used (Aspect *et al* 1982).

Before we begin the proof, let us clarify a possible confusion that the condition of Conservation may produce. In the definition of common cause for outcomes of a *single* event $E \cup F$, this condition requires that every history containing C contains both measurement events E, F as well. Clearly, in the present setup we do not want every history containing C to contain all $L\alpha, L\sigma, R\beta$, and $R\gamma$. Even a single history that had this feature would be bad, since this would suggest that, for instance, there is a history in which both settings on the left are selected simultaneously. Thus, in the definition of *common* common cause, Conservation must be appropriately modified. Since this modification is already given in the definition of non-stochastic common common cause of (Kowalski, Placek 1999),[10] I place the formal definition of common common cause in the Appendix, here modifying this condition only for the set-up considered. The modification is this:
Every history containing C contains exactly one of the following events:
$L\alpha \cup R\beta, L\alpha \cup R\gamma, L\sigma \cup R\beta$ or $L\sigma \cup R\gamma$.

8 Bell-CH inequality

To begin with the proof of the Bell-CH inequality, take note of an arithmetical fact that for any real numbers u, u', v, and v' from the $[0, 1]$ interval the following holds:

$$-1 \leq uv + uv' + u'v' - u'v - u - v' \leq 0 \tag{5}$$

In the proof below we allow for any finite number K of atomic outcomes of C.

The first assumption, Locality, says that the probability of obtaining the result at one wing should be independent from the setting chosen at a remote apparatus:

$$\begin{aligned} &\forall_{\alpha,\beta,\gamma}\ \mu_{L\alpha\cup R\beta}(L\pm) = \mu_{L\alpha\cup R\gamma}(L\pm) \\ &\forall_{\alpha,\beta,\gamma}\ \mu_{L\alpha\cup R\beta}(R\pm) = \mu_{L\gamma\cup R\beta}(R\pm) \end{aligned} \tag{LOCALITY}$$

It is a significant fact that Locality is a consequence of our REQUIREMENT.

Next, we have similar conditions, but imposed on the "bigger" probability spaces; these concern extended outcomes $L+^*, L-^*, R+^*$ and $R-^*$:

$$\begin{aligned} &\forall_{\alpha,\beta,\varphi}\ \overset{*}{\mu}_{L\alpha\cup R\beta\cup C}(L\pm^*) = \overset{*}{\mu}_{L\alpha\cup R\varphi\cup C}(L\pm^*) \\ &\forall_{\alpha,\beta,\varphi}\ \overset{*}{\mu}_{L\alpha\cup R\beta\cup C}(R\pm^*) = \overset{*}{\mu}_{L\varphi\cup R\beta\cup C}(R\pm^*) \end{aligned} \tag{LOCALITY'}$$

The motivation behind these conditions is that once Locality *simpliciter* assures that probability of the outcome of one event is independent of the remote event, another probability that is defined on an enlarged algebra should behave likewise. These conditions can also be viewed as a constraint on how the enlarged model should be constructed: the new model should retain the property of the initial model. Mathematically speaking, this can always be guaranteed, so it is rather Locality than Locality' that is at stake in the Bell-CH inequalities. It can also be seen that given Causal Precedence and Conservation, Locality' is a consequence of our REQUIREMENT. (In this area, REQUIREMENT is not testable, though.)

[10] It is Definition 14.

The last assumption is known as 'No Conspiracy' requirement, which roughly says that polarization directions and hidden states should be independent. This translates into a claim that polarization settings and outcomes ω_i of the common common cause C are independent, which technically means that the probability of ω_i should be the same, no matter by which probability measure: $\overset{*}{\mu}_{L\alpha\cup R\beta\cup C}, \overset{*}{\mu}_{L\alpha\cup R\gamma\cup C}, \overset{*}{\mu}_{L\sigma\cup R\beta\cup C}$ or $\overset{*}{\mu}_{L\sigma\cup R\gamma\cup C}$, it is calculated.

$$\forall_{\alpha,\beta,\gamma,\varphi,i}\ \overset{*}{\mu}_{L\alpha\cup R\beta\cup C}\ (\omega_i) = \overset{*}{\mu}_{L\gamma\cup R\varphi\cup C}\ (\omega_i) \overset{df}{=} \overset{*}{\mu}_C(\omega_i) \qquad \text{(NO CONSPIRACY)}$$

A bit surprisingly, No Conspiracy also follows from REQUIREMENT, given that Causal Precedence and Conservation hold.

The derivation begins with these substitutions:

$$\begin{aligned} u &= \overset{*}{\mu}_{L\alpha\cup R\beta\cup C}(L+^* \mid \omega_i) = \overset{*}{\mu}_{L\alpha\cup R\gamma\cup C}(L+^* \mid \omega_i) \\ v &= \overset{*}{\mu}_{L\alpha\cup R\beta\cup C}(R+^* \mid \omega_i) = \overset{*}{\mu}_{L\sigma\cup R\beta\cup C}(R+^* \mid \omega_i) \\ u' &= \overset{*}{\mu}_{L\sigma\cup R\beta\cup C}(L+^* \mid \omega_i) = \overset{*}{\mu}_{L\sigma\cup R\gamma\cup C}(L+^* \mid \omega_i) \\ v' &= \overset{*}{\mu}_{L\sigma\cup R\gamma\cup C}(R+^* \mid \omega_i) = \overset{*}{\mu}_{L\alpha\cup R\gamma\cup C}(R+^* \mid \omega_i), \end{aligned}$$

where the right-hand side equations rest on the Locality' condition.

Next, multiply the sides of Inequality 5 by $\overset{*}{\mu}_C(\omega_i)$, sum over i and, taking advantage of Screening-off and No Conspiracy, obtain this inequality:

$$\begin{aligned} -1 \leq &\overset{*}{\mu}_{L\alpha\cup R\beta\cup C}(L+^*\cap R+^*) + \overset{*}{\mu}_{L\alpha\cup R\gamma\cup C}(L+^*\cap R+^*) + \overset{*}{\mu}_{L\sigma\cup R\gamma\cup C}(L+^*\cap R+^*) \\ &- \overset{*}{\mu}_{L\sigma\cup R\beta\cup C}(L+^*\cap R+^*) - \overset{*}{\mu}_{L\alpha\cup R\beta\cup C}(L+^*) - \overset{*}{\mu}_{L\alpha\cup R\gamma\cup C}(R+^*) \leq 0. \end{aligned} \tag{6}$$

Since Adequacy makes identifications like

$$\overset{*}{\mu}_{L\sigma\cup R\beta}(L+^*\cap R+^*) = \mu_{L\sigma\cup R\beta}(L+\cap R+),$$

Inequality (6) indeed yields (via Causal Precedence, Conservation, and REQUIREMENT) a constraint on observable probabilities known as the Bell-Clauser-Horne inequality:

$$\begin{aligned} -1 \leq &\mu_{L\alpha\cup R\beta}(L+\cap R+) + \mu_{L\alpha\cup R\gamma}(L+\cap R+) + \mu_{L\sigma\cup R\gamma}(L+\cap R+) \\ &- \mu_{L\sigma\cup R\beta\cup C}(L+\cap R+) - \mu_{L\alpha\cup R\beta}(L+) - \mu_{L\alpha\cup R\gamma}(R+) \leq 0. \end{aligned} \tag{7}$$

Turning now to the question whether the search for a common common cause is reasonable in the setup considered, observe first that this search goes far and above Reichenbach's idea of finding a common cause for outcomes of a single event. Reichenbach's approach, if applied literally to outcomes of many events, will presumably yield many separate common causes, each taking care of outcomes of a single event. Thus, postulating a common common cause for the setup considered amounts to identifying *prima facie* separate common causes, which correspond to the four events: $L\alpha \cup R\beta, L\alpha \cup R\gamma, L\sigma \cup R\beta$, and $L\sigma \cup R\gamma$. But is there any reason for this identification? To see that indeed the identification has strong support, note that in any history containing one or another pair of settings, there is an agent that (who) selects the settings, the source of the particles in the same quantum state, and much more. In the present framework it natural to read this as the claim that all the histories share some event, i.e., a common common cause. Still remaining in the framework of stochastic OBST, you may object to this, claiming that there are four different events C_i, each taking care of exactly one of $L\alpha \cup R\beta$, $L\alpha \cup R\gamma$, $L\sigma \cup R\beta$, and $L\sigma \cup R\gamma$. Most importantly, however, each of these four measurement events occur *after* the corresponding C_i. Accordingly, events C_i should not be made to happen by something that occurs later. In short, if for instance C_2 happened, then the later occurrence of $L\alpha \cup R\beta$ has no power to change it into C_1. *Facta infecta fieri non possunt*. It is this powerful claim that supports the identification of possibly many separate common causes into a single common common cause. The opponent of common common causes may nevertheless go to the extreme, saying that there are as many types of runs as runs themselves, since the pasts of all these runs have nothing in common, no invariant. Clearly, this move blocks the quest for a common common cause, and if done sufficiently generally, the application of branching models to empirical reality. Note, however, that this stance undermines the significance of experimental sciences as well: how then could experiments be used to test theories, especially probabilistic theories? Thus, to conclude, the fact that Bell-CH inequality is provable in every OBST model with a common common cause, constitutes a powerful argument that no causal story can be told about the Bell correlations, where causality is meant as restricted by prohibition on superluminal signalling and Screening-off and Locality conditions.

The significance of this result depends on whether there are strong grounds for believing that Screening-off should hold in the setup of the experiment. In this respect our results push in the opposite directions. On the one hand, in many cases there are no non-trivial (i.e., non-deterministic) common causes for correlated outcomes of single events of the form $E \cup F$. Thus, one may become skeptical about the existence of common common causes. On the other hand, in (Kowalski, Placek 1999) we proved GHZ theorems, where we used a non-stochastic *ersatz* of screening-off that appears to be rather trivial. (GHZ theorems can be viewed as non-stochastic Bell-type theorems: they follow from certain "natural" assumptions but are predicted to be violated.) Following this line of thinking, one may decide to reject Locality or some other premise of the Bell theorem. However, in the present framework, all these premises are consequences of the innocuous-looking REQUIREMENT! And it is hard to see how our REQUIREMENT can possibly fail.[11]

Before putting forward my final conclusions, a remark about the so-called quantum holism seems in place. After all, many researchers see it as responsible for the violation of the Bell inequalities.

9 What's quantum holism?

Holism and atomism are opposing philosophical doctrines concerning the relation between properties of the whole and properties of the parts. Atomism claims that properties of the whole are entirely determined by the properties of its parts, whereas holism

[11] Other recent results related to validity of screening-off in the EPR-like correlations are even more confusing: on the one hand, the condition is provable in two different, and mathematically rigorous models, see (Spohn 1990 and 1994) and (Hofer-Szabó, Rédei and Szabó 1999). On the other hand, (Cartwright 1989) and (Cartwright and Chang 1993) have a proof, in the framework of stochastic causal linear modeling, that in EPR-like setups some generalized condition other than screening-off holds.

argues that there is more to the properties of the whole than the properties of its parts. To finesse these rather obscure claims you may first translate the notions of whole and part into those of compound-system and component-system, and then decide to refer rather to physical states than properties of physical systems. Since the physical state of the system can be viewed as assigning probabilities to all experimental propositions pertaining to the system, the holism vs. atomism controversy has a probabilistic dimension as well. Atomism thus says that probabilities assigned to propositions pertaining to components determine probabilities assigned to propositions pertaining to the compound system whereas holism denies this. In this sense quantum mechanics has a holistic feature: for a certain class of compound systems, the probability assignments for the components do not determine probability assignments for the compound system. Mathematically speaking, this is a feature of probability, and this sort of holism can emerge on the classical level as well.[12]

Now the probabilistic feature of holism can easily be modeled in the framework of stochastic OBST. To test for holism, one may think of a correlation experiment, and ask whether probabilities of results measured on the compound system (e.g., two particles) supervene on probabilities of corresponding results measured on component systems (e.g., single particles). Moreover, it may be further required that both the measuring apparatuses be turned on, so every history containing one measuring event contains the other as well, and the other way round. This will make the premises of our REQUIREMENT satisfied. To carry on this discussion in a more rigorous way, suppose E and F are events forming a space-like set, and e_i and f_j ($i \leq I, j \leq J$) are their atomic outcomes, respectively. We are interested in relations between probability assignments to outcomes e_i and f_j, on the one hand, and the probability assignment to outcome $e_i \cap f_j$ of $E \cup F$, on the other. The questions to be posed are:

1. Do probability assignments to all outcomes $e_i \cap f_j$ of $E \cup F$ ($i \leq I, j \leq J$) fully determine probability assignments to any given outcome e_i of E and any given outcome f_j of F?
2. Do probability assignments to all outcomes e_i of E and all outcomes f_j of F ($i \leq I, j \leq J$) fully determine the probability assignment to any given outcome $e_i \cap f_j$ of $E \cup F$?

Clearly, the answer to the first problem is in the positive, since

$$\mu_{E\cup F}(e_i \cap \Omega_F) = \mu_{E\cup F}(e_i) = \mu_E(e_i) \qquad \mu_{E\cup F}(\Omega_E \cup f_j) = \mu_{E\cup F}(f_j) = \mu_F(f_j).$$

Does however the converse relation obtain?

$$\begin{array}{l} \mu_E(e_i) = \mu_{E\cup F}(e_i \cap \Omega_F) \\ \mu_F(f_j) = \mu_{E\cup F}(\Omega_E \cap f_j) \end{array} \stackrel{?}{\Longrightarrow} \quad \mu_{E\cup F}(e_i \cap f_j)$$

In a rather special case, if e_i and f_j are stochastically independent, the relation above holds, since then:

$$\mu_{E\cup F}(e_i \cap f_j) = \mu_{E\cup F}(e_i \cap \Omega_F) \cdot \mu_{E\cup F}(\Omega_E \cap f_j) = \mu_E(e_i) \cdot \mu_F(f_j)$$

In general, however, this atomistic feature does not obtain, which means that the models of stochastic OBST are impartial with respect to the holism vs. atomism controversy.

[12] See (van Fraassen 1991, pp. 73–76).

But, again in rather special circumstances, if the model is committed to the quest for a common cause, the formula above holds, and consequently, the model is atomistic. To see this, suppose e_i and f_j are correlated in a model, and, in accord with the idea of common cause, an enlarged model is posited. In this enlarged model, by the probability calculus, we have:

$$\overset{*}{\mu}_{E\cup F\cup C}(e_i \cap \Omega_F) = \sum_{k}^{K} \overset{*}{\mu}_{E\cup F\cup C}(e_i \cap \Omega_F \mid \omega_k) \cdot \overset{*}{\mu}_{E\cup F\cup C}(\Omega_E \cap \Omega_F \cap \omega_k)$$

and analogously for outcome f_j. But then any two $\overset{*}{\mu}_{E\cup F\cup C}(e_i \cap \Omega_F \mid \omega_k)$ and $\overset{*}{\mu}_{E\cup F\cup C}(\Omega_E \cap f_j \mid \omega_k)$ determine, by Screening-Off, $\overset{*}{\mu}_{E\cup F\cup C}(e_i \cap f_j \mid \omega_k)$. These last factors yield $\overset{*}{\mu}_{E\cup F\cup C}(e_i \cap f_j \cap \Omega_C)$, which then Adequacy identifies with $\mu_{E\cup F}(e_i \cap f_j)$ (by REQUIREMENT).

The immediate moral from this fact is a new face of the conditions employed in the definition of common cause: they introduce an atomistic feature to models of correlated outcomes of space-like events. Another moral concerns consequences of violation of the Bell-CH inequalities. Since as we saw, probabilistic atomism is not much different than the satisfaction of all the conditions of common cause, to say that the violation testifies to quantum holism does not say anything more than "the conjunction of the conditions of common cause cannot be true". Thus, since the issue of holism tends to muddle the matter, I would recommend not refering to it in the discussions of Bell's theorem.

10 Open questions and conclusions

The big question, which is related to a theorem due to Fine (Fine 1982), is whether violation of Bell's inequalities can indicate anything about issues like causality, locality or other features of the physical world. The theorem says that satisfaction of the Bell-CH inequalities is necessary and sufficient for the existence of a joint distribution of four statistical observables, where the joint distribution fulfills certain additional requirements. This joint distribution, however, returns joint distributions for pairs of non-commuting observables, which by standard quantum mechanical wisdom cannot exist. Hence, Fine's theorem can be read as saying that violation of Bell-CH inequalities is rather a consequence of a well-known fact about quantum mechanical probabilities than a lesson about some features of physical world. Now, one may use the framework of stochastic OBST to clarify this issue, and in particular, to show that this interpretation is incorrect. To do so one needs to explicitly construct a joint distribution that Fine'e theorem refers to. This joint distribution will return, as Fine's does, the joint distributions for incompatible events; however, the existence of these probabilities will turn out non-controversial. An additional bonus of carrying out this construction is to see what mathematical meaning conditions of Adequacy, Locality, Locality', Screening-off, and No Conspiracy have. It could be seen that these conditions allow for the required joint distribution of stochastic observables. However, since this is a rather lengthy and somewhat technical topic, I postpone the investigation to a later paper.

To sum up the results of the current paper, I extended the models of outcomes in branching space-time so that they could allow for the analysis of stochastic subluminal causation, fulfilling the promise given in (Kowalski, Placek 1999). Thus, I assigned a probability space to each event in the model, the probability of an outcome of the event being interpreted as a "measure of possibility" of that outcome. Technically, the models of stochastic OBST are of the form $\langle W, \leq, \mathcal{E}, \Upsilon\rangle$, where W is non-empty set, $\leq$—partial ordering, $\mathcal{E}$—the set of events in W, and Υ—the collection of probability measures, where these probability measures correspond one-to-one to events in W. Next I defined the notion of common cause of correlated outcomes of an event. The quest for a common cause of correlated outcomes of an event in a model amounts to appropriately enlarging this initial model, so that the new model contains a properly located event, i.e, a common cause, and specific probabilistic constraints hold. I showed that in general a non-deterministic common cause for outcomes of a single event $E \cup F$ does not exist, but a deterministic common cause is always a mathematically possible option. A structure that one postulates in the Bell-Aspect setup, however, is not a common cause but a common common cause that screens-off correlated outcomes of many events, where 'events' are events of measuring spin projections on pairs of particles, and multiplicity of these events is due to many possible directions of polarization. I argued that in this particular set-up the search for such a common common cause is entirely reasonable. The conclusion is that the models of stochastic OBST rigorously show that no causal story can be told about the Bell-Aspect correlations, where causality is subluminal and subject to screening-off and locality conditions. Turning finally to assessing the implications of the violation of the Bell-CH inequalities, the diagnosis of the present framework is in agreement with the standard one. All the conditions given in the definition of common cause, with the sole exception of Screening-off, are purely constructional and can be easily satisfied. Also, as was observed in Section 8, Locality' is trivially satisfiable once Locality holds, and both these constraints are entailed by our REQUIREMENT. This last condition entails as well No Conspiracy. Further, the issue of holism reduces to the problem of satisfying the conditions of common cause (see Section 8). Thus, we are left with this choice: REQUIREMENT, Screening-Off, and Subluminal Causality cannot be jointly true.

In this final paragraph, let me mention a task and comment on prospects for future development. A question that comes to mind is the precise relation between models of OBST and models of stochastic OBST. Since this is a relatively easy problem, I leave it to the reader, suggesting perhaps that to this end it is worth drawing the stochastic model for GHZ (the only probabilities being 1 and 0) and then comparing it with the non-stochastic model for GHZ.[13] A more serious issue is a relation between quantum logic and the models of OBST. The framework implies that if a measurement can be performed, the outcomes of it form a Boolean algebra. The set of outcomes of more than one event, however, is not even a lattice. Thus, could some elements of the component Boolean algebras be glued together so that the resulting structure will turn out to be a partial Boolean algebra? Given that there is a systematic way of doing this (a procedure known as "pasting" among algebraists), it will provide a link between OBST and quantum logic, proving thus that models of stochastic OBST provide a semantics for quantum logic.

Department of Philosophy
Jagiellonian University, Kraków, Poland
e-mail: uzplacek@cyf-kr.edu.pl

[13] See (Kowalski, Placek 1999).

Appendixes

Facts

FACT 1
If e and f are outcomes of E and F, respectively, then $e \cap f$ is an outcome of $E \cup F$. If a is an outcome of $E \cup F$, then there are outcomes e of E and f of F such that $a = e \cap f$.

PROOF. We begin with the first implication. Since $A \subseteq A^{\perp_B \perp_B}$ for any $A, B \subset W$ and $B \prec A$,[14] it remains to be shown that $(e \cap f)^{\perp_{E\cup F} \perp_{E\cup F}} \subseteq e \cap f$. Thus, suppose (indirectly) that

$$(1)\ x \in (e \cap f)^{\perp_{E\cup F} \perp_{E\cup F}} \ \text{ but } (2)\ x \notin e \cap f.$$

Note that (1) is equivalent to:

$$\forall_{y \in \Omega_{E\cup F}} \exists_{z \in e\cap f} [z \perp_{E\cup F} y \Rightarrow x \perp_{E\cup F} y]. \tag{*}$$

(1) means that $x \in \Omega_{E\cup F}$. Next, (2) entails $x \notin e$, and this in turn—$x \in e^{\perp_E}$.[15] Since $z \in e$, we have $z \perp_E x$. From this $z \perp_{E\cup F} x$ follows. Now, instantiating x for y in Eq. *, we have

$$x \perp_{E\cup F} x,$$

which is absurd.

To show that the second implication of Fact 1 holds, we explicitly construct e and f such that:

$$e = a^{\perp_E \perp_E} \quad f = a^{\perp_F \perp_F}$$

By Lemma 1 (ii) of (Kowalski, Placek 1999) e and f are indeed outcomes of E and F, respectively. Moreover, since $E \prec a$ and $F \prec a$, $a \subseteq a^{\perp_E \perp_E} = e$ and $a \subseteq a^{\perp_F \perp_F} = f$,[16] and hence $a \subseteq e \cap f$. Thus, we need only to show that

$$e \cap f = a^{\perp_E \perp_E} \cap a^{\perp_F \perp_F} \subseteq a$$

Suppose (indirectly) that for some x, (1) $x \in a^{\perp_E \perp_E} \cap a^{\perp_F \perp_F}$, but (2) $x \notin a$. Condition (1) means that $E \prec x$ and $F \prec x$. (1) is also equivalent to:

$$\forall_{y \in \Omega_E} \exists_{z \in a} [z \perp_E y \Rightarrow x \perp_E y] \ \text{ and } \ \forall_{y \in \Omega_F} \exists_{z \in a} [z \perp_F y \Rightarrow x \perp_F y] \tag{**}$$

From (2) we have $x \in a^{\perp_{E\cup F}}$.[17] Hence, for $z \in a$:

$$C(z, x) \cap (E \cup F) \neq 0.$$

Thus, $C(z, x) \cap E \neq 0$ or $C(z, x) \cap F \neq 0$, which means $z \perp_E x$ or $z \perp_F x$. Instantiating now x for y in Eq. (**), we have $x \perp_E x$ or $x \perp_F x$, which is absurd.

□

FACT 2

[14] See Lemma 1 (i) of (Kowalski, Placek 1999).

[15] This follows from the fact proved in the proof of Lemma 5 of (Kowalski, Placek 1999) which says that if $E \prec x$ and $x \notin e$, then $x \in e^{\perp_E}$, where e is an outcome of E.

[16] See Lemma 1 (i) of (Kowalski, Placek 1999).

[17] See the proof of Lemma 5 of (Kowalski, Placek 1999).

If e and f are atomic outcomes of E and F, respectively, then $e \cap f$ is either empty or an atomic outcome of $E \cup F$.

PROOF
Observe first that if $e \cap f = \emptyset$, then (trivially) $e \cap f$ is an empty outcome of $E \cup F$. Thus, assume that $e \cap f \neq \emptyset$ and (indirectly) that it is not atomic. Thus, two different atomic outcomes a and b of $E \cup F$ can be found such that $a \subset e \cap f$ and $b \subset e \cap f$. Since these are atoms, we have:

$$\forall_{x \in a, y \in b}\; x \perp_{E \cup F} y$$

and hence

$$\forall_{x \in a, y \in b}\; C(x,y) \cap (E \cup F) \neq \emptyset$$

This means that one can find either (1) some x, y such that $C(x,y) \cap E \neq \emptyset$, or (2) some x, y such that $C(x,y) \cap F \neq \emptyset$. Taking case (1), it entails $x \perp_E y$. Now we build outcomes $\{x\}^{\perp_E \perp_E}$ and $\{y\}^{\perp_E \perp_E}$ that are different. Since $\{x\} \subset a$ and $a \subset e \cap f \subset e$, it follows that two different outcomes $\{x\}^{\perp_E \perp_E}$ and $\{y\}^{\perp_E \perp_E}$ are contained in e. The argument for case (2) goes analogously. Thus, contrary to the assumption of the Fact, we have that at least one of e, f is not atomic.

□

Discussion of Eqs. (1) and (3)

To begin with, note that if $M = 1$ or $N = 1$, the formula (3), i.e.,

$$\sum_{m \neq i, n \neq j}^{M,N} (p_{ijk} p_{mnk} - p_{ink} p_{mjk}) = 0$$

yields only a trivial constraint, i.e., satisfied by all p_{ijk}. Thus, since Eqs. (1) imposes $M \cdot N$ constraints, the number of variables p_{ijk} being $M \cdot N \cdot K$, it is possible to treat some p_{ijk} as parameters and find solutions such that for any k at least two p_{ijk} are non-zero. This can be done for any $K \geq 2$.

A similar result is true for the case (1) with $M = N = 2$ and the case (2) with $M = 2, N = 3$ or $M = 3, N = 2$. In case (1), formula (3) yields K constraints, and accordingly there are truly stochastic solutions for $K \geq 2$. In case (2), this formula imposes $3 \cdot K$ non-trivial constraints, there being truly stochastic solution only if $K \geq 3$.

Now, if $M = N = 3$, the number of independent equations formula (3) determines is $N \cdot M \cdot K$, and for a given k, the equations have only two kinds of solutions:

1. for all $i, j \leq 3$, $p_{ijk} = a$;
2. $p_{ijk} = a \neq 0$ and $p_{mnk} = 0$ if $m \neq i$ or $n \neq j$.

Now, given that some outcomes of $E \cup F$ are correlated, surface probabilities q_{ij} cannot be all equal. Thus, to satisfy Eqs. (1) one needs to include in the sum there at least one set $\{p_{ijk}\}$ (for some k) that is a solution of the second kind of Eqs. (3). Consequently, at least one outcome ω_k of C must work deterministically.

An alternative definition of common cause

As before, let us consider space-like events E and F, $e_1, \ldots, e_N$ and $f_1, \ldots f_M$ being atomic outcomes of E and F, respectively. Outcomes e_i correspond one-to-one to results r^{ei} and outcomes f_j correspond one-to-one to results r^{fj}. Some (or even all) atomic outcomes $e_i \cap f_j$ of $E \cup F$ are correlated. To view these correlations as resulting from an underlying common causes C one enlarges the initial model by multiplying event E with its future as well as event F with its future. One may consider as well the multiplication of an event preceding E and an event preceding F. As a result, you obtain (in the enlarged model) K events E_k, K multiplied families of atomic outcomes $\{e_{1k} \ldots e_{nk}\}$, and K families of corresponding results $\{r_k^{e1} \ldots r_k^{en}\}$. The same happens to event F, its atomic outcomes, and corresponding results. Next, you require that any E_k and E_l ($k \neq l$) have nothing in common in future, i.e.

$$\neg \exists_{x \in W, y \in E_k, z \in E_l} \; x \geq y \;\wedge\; x \geq z$$

and analogously for events F_k. You also demand that intersections of only specific e_{ik} and f_{jk} are non-empty. If $e_i \cap f_j \neq \emptyset$ ($i \leq N, j \leq M$), then

$$e_{ik} \cap f_{jl} \neq \emptyset \text{ iff } l = k.$$

To introduce common cause C, you require that numbers k code all atomic outcomes of C:

$$E_n \subset \omega_k \Leftrightarrow n = k \text{ and } F_n \subset \omega_k \Leftrightarrow n = k \text{ for all } k \leq K.$$

Note that this principle entails both Causal Precedence and Conservation. Screening-off now reads:,

$$\forall_{i \leq N, j \leq M, k \leq K} \; \overset{*}{\mu}_{E_k \cup F_k}(e_{ik} \cap f_{jk}) = \overset{*}{\mu}_{E_k \cup F_k}(e_{ik} \cap \Omega_{F_k}) \cdot \overset{*}{\mu}_{E_k \cup F_k}(\Omega_{E_k} \cap f_{jk}),$$

Finally, Adequacy identifies:

$$\sum_{k \leq K} \overset{*}{\mu}_{E_k \cup F_k}(e_{ik} \cap f_{jk}) = \mu_{E \cup F}(e_i \cap f_j).$$

We leave to the reader a construction of the definition of common common cause along the lines sketched here.

Common common cause

We say that two events E and F are *incompatible* if there is no history that contains both of them. Now, the definition reads:

DEFINITION 13 (OF COMMON COMMON CAUSE)
Let $\{E^1, F^1\}, \{E^2, F^2\} \ldots \{E^N, F^N\}$ be a collection of space-like sets in model $\langle W, \leq, \mathcal{E}, \Upsilon \rangle$. Let $E^n \cup F^n$ and $E^m \cup F^m$ be incompatible for all $n \neq m$ ($n, m \leq N$) and some event S exist such that $S \prec E^n$ and $S \prec F^n$ for all $n \leq N$. Further let $e_1^n, \ldots, e_I^n$ be atomic outcomes of E^n and $f_1^n, \ldots f_J^n$ be atomic outcomes of F^n, which corresponds to results $r^{e1n}, \ldots, r^{eIn}$ and $r^{f1n}, \ldots r^{fJn}$, respectively. Let also,

for each $n \leq N$ some outcomes of $E^n \cup F^n$ be correlated, i.e., for some (possibly all) $i \leq I$ and $j =\leq J$:

$$\mu_{E^n \cup F^n}(e_i^n \cap f_j^n) \neq \mu_{E^n \cup F^n}(e_i^n) \cdot \mu_{E^n \cup F^n}(f_j^n).$$

A common common cause C that accounts for all these correlations is an event with K atomic outcomes ω_k in a model $\langle W^*, \leq^*, \mathcal{E}^*, \Upsilon^* \rangle$ that is enlarged in respect to $\langle W, \leq, \mathcal{E}, \Upsilon \rangle$ by:

1. *K-multiplying results r^{ein} and r^{fjn} to obtain r_k^{ein} and r_k^{fjn}, respectively, for all $k \leq K, i \leq I, j \leq J$, and $n \leq N$;*
2. *extending outcomes e_i^n and f_j^n into e_i^{n*} and f_j^{n*}, respectively, for all $i \leq I, j \leq J$, and $n \leq N$;*
3. *requiring that for each atomic outcome $e_i^n \cap f_j^n$ of $E^n \cup F^n$*

$$(r_k^{ein})^{\perp_E n \perp_E n} \cap (r_l^{fjn})^{\perp_F n \perp_F n} \neq \emptyset \;\text{ iff }\; k = l \;\wedge\; k \leq K.$$

4. *adding an event C with K atomic outcomes.*

Common common cause C should satisfy:

1. *For all $i \leq I,\ j \leq J,\ k \leq K,\ n \leq N$,*

$$C \prec r_k^{ein} \;\text{ and }\; C \prec r_k^{fjn}, \qquad \text{(Causal Precedence)}$$

   where r_k^{ein} and r_k^{fjn} are K-multiplied results r^{ein} and r^{fjn}, respectively;

2. *Every history that contains C, contains E^n and F^n for some $n \leq N$;*
   (Conservation)

3. *For all $i \leq I, j \leq J, k \leq K, n \leq N$,*

$$\overset{*}{\mu}_{E^n \cup F^n \cup C}(e_i^{n\,*} \cap f_j^{n\,*} \mid \omega_k) = \overset{*}{\mu}_{E^n \cup F^n \cup C}(e_i^{n\,*} \mid \omega_k) \cdot \overset{*}{\mu}_{E^n \cup F^n \cup C}(f_j^{n\,*} \mid \omega_k),$$

   where $\overset{*}{\mu}_{E^n \cup F^n \cup C}(\omega_k) = \overset{*}{\mu}_{E^n \cup F^n \cup C}(\Omega_{E^n} \cap \Omega_{F^n} \cap \omega_k)$; (Screening-off)

4. *For all $i \leq I, j \leq J, k \leq K, n \leq N$,*

$$\mu_{E^n \cup F^n}(e_i^n \cap f_j^n) = \overset{*}{\mu}_{E^n \cup F^n}(e_i^{n\,*} \cap f_j^{n\,*}). \qquad \text{(Adequacy)}$$

Events E^n (F^N) are assumed to be incompatible since they represent alternative measurements. Although they are alternative, there is some event at which one or another setting is chosen, and this translates into requiring that event S exists. Note that Causal Precedence and Conservation guarantee that if only (in the initial model) $e_i^n \cap f_j^n \neq \emptyset$, then (in the enlarged model) for any ω_k, $e_i^{n\,*} \cap f_j^{n\,*} \cap\ \omega_k \neq \emptyset$. Thus, the two conditions permit the application of REQUIREMENT, which yields

$$\overset{*}{\mu}_{E^n \cup F^n}(e_i^{n\,*} \cap f_j^{n\,*}) = \overset{*}{\mu}_{E^n \cup F^n \cup C}(e_i^{n\,*} \cap f_j^{n\,*}).$$

This is what allows for applications of Adequacy.

References

Aspect, A., Dalibard, J. and Roger, G. [1982]: 'Experimental Test of Bell's Inequalities using Time-Varying Analyzers', *Phys. Rev. Lett.* **49**, pp. 1804–1807.

Bell, J. [1964]: 'On the Einstein-Podolski-Rosen paradox', *Physics* **1**, pp. 195–200.

Belnap, N. [1992]: 'Branching Space-time', *Synthese* **92**, pp. 385–434.

Cartwright, N. [1989]: *Nature's Capacity and Their Measurement*, New York: Oxford University Press.

Cartwright, N. and Chang, H. [1993]: 'Causality and Realism in the EPR Experiment', *Erkenntniss* **3**8, pp. 169–190.

Clauser, J.F. and Horne, M.A. [1974]: 'Experimental consequences of objective local theories', *Phys. Rev. D* **10**, pp. 526–535.

Fine, A. [1982]: 'Joint distributions, quantum correlations, and commuting observables', *Journal of Mathematical Physics* **23** (7), pp. 1306–1310.

Fine, A. [1982a]: 'Hidden Varibles, Joint Probability, and the Bell Inequalities', *Phys. Rev. Lett.* **48** (5), pp. 291–295.

Greenberger, D., Horne, M.A., and Zeilinger, A. [1989]: 'Going Beyond Bell's Theorem', in M. Kafatos (*ed.*), *Bell's Theorem, Quantum Theory, and Conceptions of Universe*, Dordrecht: Kluwer Academic Press, pp. 69–72.

Greenberger, D., Horne, M.A., Shimony, A., and Zeilinger, A. [1990]: 'Bell's theorem without inequalities', *Am. J. Phys* **5**8 (12), Dec. 1990, p. 1131–1143.

Hofer-Szabó, G., Rédei, M., and Szabó, L. [1999]: 'On Reichenbach's Common Cause Principle and Reichenbach's Notion of Common Cause', forthcoming in *Brit. J. Phil. Sci.*

Kowalski, T. and Placek, T. [1999]: 'Outcomes in branching space-time and GHZ-Bell theorems', forthcoming in *Brit. J. Phil. Sci.*

Reichenbach, H. [1956]: *The Direction of Time* (*ed.* by M. Reichenbach), Berkeley: University of California Press.

Spohn, W. [1990]: 'Direct and Indirect Causes', *Topoi* **9**, pp. 125-145.

Spohn, W. [1994]: 'On Reichenbach's Principle of the Common Cause', in W.C. Salmon and G. Wolters: *Logic, Language, and the Structure of Scientific Theories*, Pittsburgh: University of Pittsburgh Press, pp. 211–235.

van Fraassen, B. [1991]: *Quantum Mechanics. An Empiricist View*, Oxford: Clarendon Press.