On the relation between the probabilistic characterization of the common cause and Bell's notion of local causality

*Gábor Hofer-Szabó**

Abstract

In the paper the relation between the standard probabilistic characterization of the common cause (used for the derivation of the Bell inequalities) and Bell's notion of local causality will be investigated. It will be shown that the probabilistic common cause follows from local causality if one accepts, as Bell did, two assumptions concerning the common cause: first, the common cause is localized in the *intersection* of the past of the correlating events; second, it provides a *complete specification* of the 'beables' of this intersection. However, neither assumptions are *a priori* requirements. In the paper the logical role of these assumptions will be studied and it will be shown that only the second assumption is necessary for the derivation of the probabilistic common cause from local causality.

1 Introduction

There has been a long tradition going back to Hans Reichenbach (1956) to characterize the notion of the common cause in probabilistic terms. This probabilistic characterization of the common cause turned out to be a fruitful mathematical tool to study causal problems in physics, among them the possibility of hidden variable models for quantum theory. In its full-fledged form a probabilistic common causal explanation contained not only the condition expressing Reichenbach's characterization of the common cause as a screener-off, but also such probabilistic requirements as locality and no-conspiracy. Since these latter requirements had spatiotemporal connotations, the question arose as to whether there exists a 'spatiotemporal justification' of the probabilistic requirements imposed on the notion of the common cause. The first step in such a justification is to establish a mathematically well-defined and physically well-motivated relation connecting *events* undestood as elements of a probability space and *regions* understood as subsets of a spacetime. Only after having such a relation can we ask whether a certain probabilistic equation can be derived from a certain spacetime localization of the common cause.

What kind of spacetime localizations do we have in mind? Obviously, the common cause is an event C happening somewhere in the past of two correlating events, say A and B. But in which past? Relativistically two spacetime separated events can have (at least) two different pasts. Let V_A and V_B denote the regions where A and B, respectively are localized. Then one can define the *weak past* of A and B as $\mathcal{P}^W(V_A, V_B) := I_-(V_A) \cup I_-(V_B)$ and the *strong past* of A and B as $\mathcal{P}^S(V_A, V_B) := I_-(V_A) \cap I_-(V_B)$ where $I_-(V)$ denotes the union of the causal pasts $I_-(x)$ of every point x in V. Let us call the appropriate common causes *weak* and *strong common causes*, respectively (see Fig. 1).

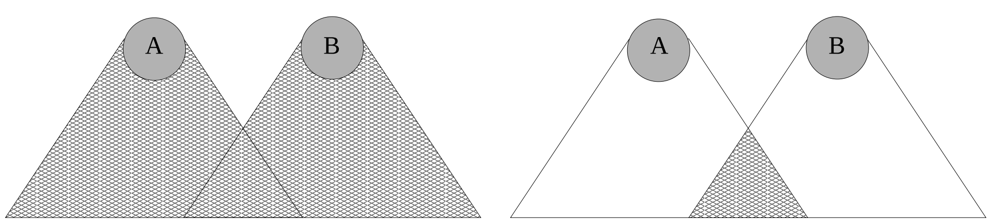

Figure 1: Weak and the strong past of the correlating events A and B.

*Institute of Philosophy, Research Center for the Humanities, Budapest, email: szabo.gabor@btk.mta.hu

Now, one might consider the *strong* past as a more natural candidate for the localization of the common cause, and indeed plenty of classical examples attest that the strong past is a reasonable choice. (But see (Butterfield, 1989) and the debate in (Henson, 2005), (Rédei and Sanpedro, 2012), (Henson, 2013).) The correlation between two fans' shouting at the same time at a football match is explained by the goals scored, that is by events localized in the strong past of the shouts. Curiously enough, however, in algebraic quantum field theory common causes are typically understood as *weak* common causes. It is not difficult to see why.

In algebraic quantum field theory observables are represented by (C^*-)algebras associated to bounded regions of a spacetime. This association is called a *net*. A *state* ϕ is defined as a normalized positive linear functional on the quasilocal algebra $\mathcal{A}$ which is the inductive limit of the net. From our perspective, the two important axioms of the *net*, are *isotony* and *local primitive causality*. *Isotony* requires that if a region V_1 is contained in another region V_2, then the local algebra $\mathcal{A}(V_1)$ associated to V_1 be a (unital C^*-)subalgebra of $\mathcal{A}(V_2)$. *Local primitive causality* is the requirement that for any region V, $\mathcal{A}(V) = \mathcal{A}(V'')$, where V'' is the causal completion of V.

Now, suppose that there is a (superluminal) correlation, $\phi(AB) \neq \phi(A)\phi(B)$, between events $A \in \mathcal{A}(V_A)$ and $B \in \mathcal{A}(V_B)$ such that V_A and V_B are spatially separated. Consider the local algebra $\mathcal{A}((V_A \cup V_B)'')$ associated to the causal completion of $V_A \cup V_B$ and suppose that we find a common cause C of the correlation in $\mathcal{A}((V_A \cup V_B)'')$. In which past of V_A and V_B can C be located? Consider a region V_C in the *weak* past $P^W(V_A, V_B)$ which is 'wide' enough to ensure that $(V_A \cup V_B) \subset V_C''$. Due to isotony, $\mathcal{A}(V_A \cup V_B)$ will be a subalgebra of $\mathcal{A}(V_C'')$ which, due to local primitive causality, is identical to $\mathcal{A}(V_C)$. Thus, C will be in V_C and hence in the weak past of V_A and V_B. To sum up, isotony and local primitive causality ensures that if a superluminal correlation has a common cause, then it can be localized in the weak past.

Can the common cause be localized also in the strong past? It may, but not simply due to the axioms of algebraic quantum field theory. If V_C is in $P^S(V_A, V_B)$, then isotony and local primitive causality does not help to relate $\mathcal{A}(V_C)$ to $\mathcal{A}((V_A \cup V_B)'')$. One also needs to know about the dynamics of the system. The axioms of algebraic quantum field theory are completely silent about whether one can locate the common cause in the strong past. As a consequence, weak common causes *cannot* be excluded *a priori* from our explanatory arsenal. (For more on common causal explanation in algebraic quantum field theory see (Rédei 1997), (Rédei and Summers, 2002), (Butterfield 2007) and (Hofer-Szabó and Vecsernyés, 2012a,b, 2013a,b).)

So we have (at least) two options to localize the common cause in the past of the correlating events. What else can we use in the derivation of the probabilistic common cause? Some principles regulating the possible causal connection of events in accordance with the special theory of relativity. An analogy might help. The theory of Bayesian nets consists of two components: a causal graph representing the causal relations among certain events and a probability space with random variables. How these two parts of the theory are related to one another? The bridge relating the two components is called the Causal Markov Condition. It says that if the nods on the graph are related to one another in such-and-such a way, the variables pertaining to the nods should satisfy such-and-such probabilistic independencies. So the role of the Causal Markov Condition in the theory of Bayesian nets is to syncronize the probabilistic and the graphic description of causal relations.

A principle playing a similar syncronizing role in the philosophy of physics has been introduced by John S. Bell (1975/87) and has been called *local causality*. Local causality is a relativistic principle tailor-made to study probabilistic relations between events localized in different spacetime regions, among them the relation between the common cause and the correlating events. From the influential writings of Bell on, the probabilistic notion of the common cause has been regarded as an expression of probabilistic constraints between certain events in the spacetime imposed on by relativistic considerations. In what follows we will show that the link between the spatiotemporal and the probabilistic characterization of the common cause is very sensitive to two essential assumptions concerning the common cause, both rightly emphasized by Bell himself. The first assumption is that the common cause is localized in the *strong past*, the second is that it provides a *complete specification* of the causal past of the correlating events.

In the paper we intend to investigate the role of these assumptions in the derivation of the probabilistic common cause from local causality. In Section 2 the standard requirements of the probabilistic common causal explanation will be recalled. In Section 3 Bell's original idea of local causality will be delineated with the emphasis on the role of the two above assumptions. In order to proceed in a more picturesque way, in Section 4 and 5 classical toy models will be introduced which will help us in scrutinizing the role of the two assumptions in the derivation of probabilistic common cause from local causality. We conclude the paper in Section 6. Some technicalities are put in the Appendices.

2 Common causal explanation

As mentioned above, the first probabilistic characterization of the common cause is due to Reichenbach. There is a long route leading from Reichenbach's original idea of the common cause to the sophisticated probabilistic requirements used today in the philosophy of quantum physics. For the sake of brevity, we do not repeat here all the intermediate steps of the entire definitional process (for this see (Hofer-Szabó and Vecsernyés, 2012a)), but jump directly to the full-fledged probabilistic characterization of the common cause and give a brief motivation of the requirements thereafter.

Let $\{a_m\}$ and $\{b_n\}$ ($m \in M, n \in N$) be two sets of measurement procedures (thought as happening in two spatially separated spacetime regions). Suppose that each measurement can have two outcomes and denote the 'positive' outcomes by A_m and B_n and the 'negative' outcomes by $\overline{A}_m$ and $\overline{B}_n$, respectively. Let all these events be accomodated in a classical probability space (Σ, p). Suppose that there is a conditional correlation between the measurement outcomes in the sense that for any $m \in M$ and $n \in N$

$$p(A_m \wedge B_n | a_m \wedge b_n) \neq p(A_m|a_m)\, p(B_n|b_n) \tag{1}$$

representing that if we set to measure the pair a_m and b_n, the appropriate outcomes will correlate.

The standard probabilistic characterization of a common causal explanation of this correlation is the following. A partition $\{C_k\}$ in Σ (that is a set of mutually exclusive events adding up to unity) is said to be a *local, non-conspiratorial joint common causal explanation* of the correlations (1) if for any $m, m' \in M$ and $n, n' \in N$ the following requirements hold:

$$p(A_m \wedge B_n | a_m \wedge b_n \wedge C_k) = p(A_m|a_m \wedge b_n \wedge C_k)\, p(B_n|a_m \wedge b_n \wedge C_k) \qquad \text{(screening-off)} \tag{2}$$

$$p(A_m|a_m \wedge b_n \wedge C_k) = p(A_m|a_m \wedge b_{n'} \wedge C_k) \qquad \text{(locality)} \tag{3}$$

$$p(B_n|a_m \wedge b_n \wedge C_k) = p(B_n|a_{m'} \wedge b_n \wedge C_k) \qquad \text{(locality)} \tag{4}$$

$$p(a_m \wedge b_n \wedge C_k) = p(a_m \wedge b_n)\, p(C_k) \qquad \text{(no-conspiracy)} \tag{5}$$

The motivation behind requirements (2)-(5) is the following. *Screening-off* (2) (also called as *outcome independence* (Shimony, 1986), *completeness* (Jarrett, 1984) and *causality* (Van Fraassen, 1982)) is simply the application of Reichenbach's original characterization of the common cause as a screener-off to conditional correlations: although A_m and B_n are correlating conditioned on a_m and b_n, they will cease to do so, if we further condition on C_k. *Locality* (3)-(4) (also called as *parameter independence* (Shimony, 1986), *locality* (Jarrett, 1984) and *hidden locality* (Van Fraassen, 1982)) is the constraints that the measurement outcome on the one side can depend only on the measurement choice on the same side and the value of the common cause, but not on the measurement choice on the opposite side (for more on that see below). Finally, no-conspiracy (5) is the requirement that the common cause system and the measurement choices should not influence each other, they should be probabilistically independent.

Now, it is a well known fact that if a set of correlations has a local, non-conspiratorial joint common causal explanation in the above sense, then the set of correlations has to satisfy various Bell inequalities. (For the derivation of one of the simplest Bell inequality, the Clauser–Horne inequality see Appendix A.) In the EPR situation (if quantum correlations are interpreted as classical conditional correlation *alá* (1)) these Bell inequalities are violated excluding a local, non-conspiratorial joint common causal explanation of EPR correlations.

Thus, in the EPR-Bell literature (2)-(5) is regarded as the correct probabilistic characterization of the common cause. But observe that the above relativistic motivations for the probabilistic independence relations (2)-(5) are completely meaningless until we do not localize the common cause on the spacetime, or more generally, until we have no principled way to associate *events* understood as elements of the probability space (Σ, p) to *regions* of a given spacetime.

Suppose that we do have such an association, that is suppose we have an isotone net $\mathfrak{N}$ associating bounded regions of the Minkowski spacetime to σ-subalgebras of Σ. We do not assume that local primitive causality also holds. (For more on the relation of local primitive causality and Bell's local causality see (Hofer-Szabó and Vecsernyés, 2014).) What else is needed for (2)-(5) to represent a legitimate probabilistic characterization of a common cause? Does Bell's notion of local causality, for instance, help us to arrive at (2)-(5)? Or turning the question around, do the probabilistic constraints imposed on the notion of common cause restrict also the possible spacetime localization of the common cause? Do we need to choose between weak and strong common causes for example? To address these questions first recall the notion of local causality.

3 Local causality

As mentioned in the Introduction, there is an influential tradition according to which equations (2)-(5) are consequences of the requirement that a certain set of correlations are to be accomodated in a *locally causal* theory. The clearest formulation of such a theory is due to Bell himself:

> "Consider a theory in which the assignment of values to some beables Λ implies, not necessarily a particular value, but a probability distribution, for another beable A. Let $p(A|\Lambda)$ denote[1] the probability of a particular value A given particular values Λ. Let A be localized in a space-time region A. Let B be a second beable localized in a second region B separated from A in a spacelike way. (Fig. 2.) Now my intuitive notion of local causality is that events in B

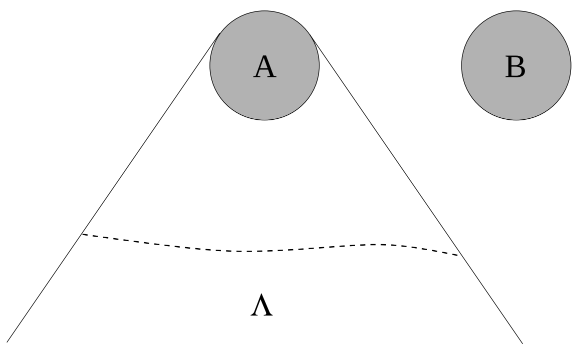

Figure 2: Local causality I.

> should not be 'causes' of events in A, and vice versa. But this does not mean that the two sets of events should be uncorrelated, for they could have common causes in the overlap of their backward light cones. It is perfectly intelligible then that if Λ in (6) does not contain a complete record of events in that overlap, it can be usefully supplemented by information from region B. So in general it is expected that
>
> $$p(A|\Lambda, B) \neq p(A|\Lambda) \tag{6}$$
>
> However, in the particular case that Λ contains already a *complete* specification of beables in the overlap of the light cones, supplementary information from region B could reasonably be expected to be redundant."

And here comes the definition of a locally causal theory.

> "Let C denote a specification of *all* beables, of some theory, belonging to the overlap of the backward light cones of spacelike regions A and B. Let a be a specification of some beables from the remainder of the backward light cone of A, and B of some beables in the region B. (See Fig. 3.) Then in a *locally causal theory*
>
> $$p(A|a, C, B) = p(A|a, C) \tag{7}$$
>
> whenever both probabilities are given by the theory." (Bell, 1987, p. 54)

[1] For the sake of uniformity throughout the paper I slightly changed Bell's denotation and figures.

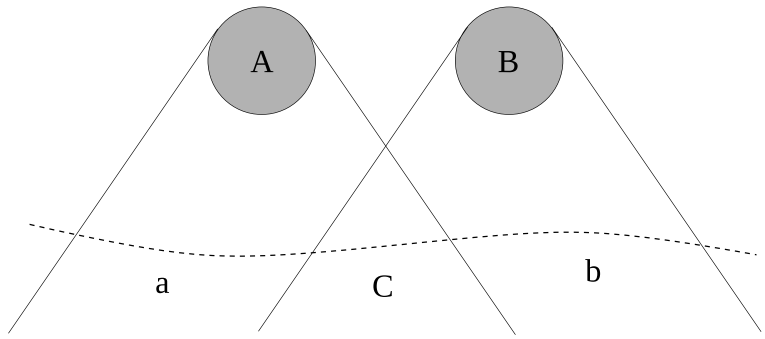

Figure 3: Local causality II.

Now, let us spell out Bell's characterization of local causality using the notion of *net* introduced above. To do this first we should translate Bell's language using random *variables* in (7) into a language using *events*. Second, the term 'beables in a certain spacetime region' is to be replaced by 'events in an algebra supported in a certain spacetime region' and 'complete specification' by 'set of atoms of the algebra in question' (assuming that local algebras are atomic). (For more on Bell's local causality and the role of 'beables' see (Norsen 2011); for the translation of 'complete specification' into atomicity see (Henson, 2013).) Finally, instead of considering the *whole causal past* of an event we will consider only a suitable *Cauchy segment* of this past.

Then Bell's notion of local causality can be paraphrased as follows.

Definition 1. An isotone net $\mathfrak{N}$ associating bounded regions of the Minkowski spacetime to σ-subalgebras of Σ is called *locally causal*, if for any classical probability measure p (or, more generally, state ϕ) on Σ, and for any two events $A_m \in \mathcal{A}(V_A)$ and $B_n \in \mathcal{A}(V_B)$ localized in the spatially separated regions V_A and V_B and correlating in the probability measure p, and for every Cauchy surface $\mathcal{S}$ (lying past to V_A and V_B), the following is true:

Let V_a, V_C and V_b be three open neighborhoods of $\mathcal{S}\cap(I_-(V_A)\setminus I_-(V_B))$, $\mathcal{S}\cap P^S(V_A, V_B)$ and $\mathcal{S}\cap(I_-(V_B)\setminus I_-(V_A))$, respectively (see Fig. 4) and let $\mathcal{A}(V_a)$, $\mathcal{A}(V_C)$ and $\mathcal{A}(V_b)$ the associated local algebras. Let a_m

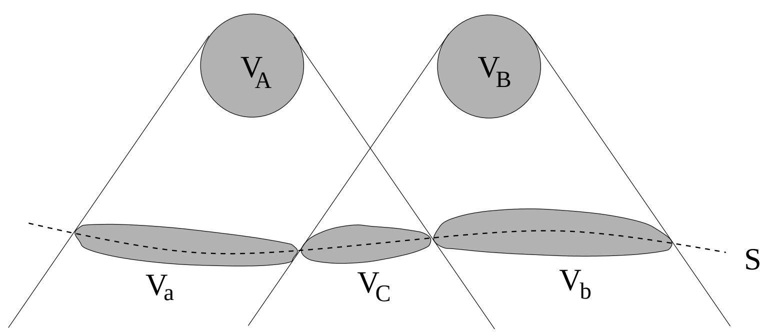

Figure 4: Local causality III.

and b_n be events in $\mathcal{A}(V_a)$ and $\mathcal{A}(V_b)$, respectively and let C_k be an atom in $\mathcal{A}(V_C)$. Then the following conditional probabilistic independencies hold:

$$p(A_m|a_m \wedge C_k \wedge B_n) = p(A_m|a_m \wedge C_k) \tag{8}$$
$$p(B_n|A_m \wedge C_k \wedge b_n) = p(B_n|b_n \wedge C_k) \tag{9}$$
$$p(A_m|a_m \wedge C_k \wedge b_n) = p(A_m|a_m \wedge C_k) \tag{10}$$
$$p(B_n|a_m \wedge C_k \wedge b_n) = p(B_n|b_n \wedge C_k) \tag{11}$$

Why four equations instead of Bell's single (7)? Observe that (9) is just the symmetric version of (8) where A_m and a_m are interchanged with B_n and b_n. Equations (10)-(11), however, are slight extensions of Bell's formulation. Observe that V_A is spacelike separated not only from V_B but also from V_b, and therefore the same conditional independence should hold between A_m and b_n as between A_m and B_n. Thus (10) is the application of Bell's idea to algebras $\mathcal{A}(V_A)$ and $\mathcal{A}(V_b)$, and (11) to algebras $\mathcal{A}(V_b)$ and $\mathcal{A}(V_A)$. There are no more spatially separated regions in Fig. 4 to which local causality could be applied.

How the above considerations relate to the probabilistic characterization (2)-(5) of the common cause delineated in the previous Section?

First observe that (10)-(11) are equivalent to *locality* (3)-(4) and from (8)-(11) *screening-off* (2) follows directly. This proves that the probabilistic characterization of the common cause by the requirements of screening-off and locality can be 'derived' from Bell's notion of local causality imposed on an isotone net associating spacetime regions and local albegras. We note, however, that the third requirement in the definition of a local, non-conspiratorial joint common causal explanation, namely no-conspiracy (5) *cannot* be 'derived' from Bell's notion of local causality in a similar way. No-conspiracy is an independent assumption stating that the events $a_m \in \mathcal{A}(V_a)$, $C_k \in \mathcal{A}(V_C)$ and $b_n \in \mathcal{A}(V_b)$ are probabilistically independent.

So far, so good. But here comes the point. To obtain this deductive relation between the probabilistic characterization of the common cause and Bell's notion of local causality the following two assumptions have been made: the common cause system provides "a *complete* specification of beables", and it is located in the "overlap of the light cones". In other words, one assumed that (i) C_k is located in the *strong past* of the correlating events, and (ii) it is an *atom* of the appropriate algebra. As we saw, Bell explicitly stresses both assumptions, and in all the subsequent papers of Van Fraassen (1982), Jarrett (1984), Shimony (1986) etc. trying to turn spacetime considerations into probabilistic independencies these two assumptions have been (explicitly or implicitly) made.

However, neither assumptions are *a priori* requirements concerning the common cause. One can easily make up common causes which are either *non-atomic* or *not located in the strong past* of the correlating events. How these common causes relate to Bell's notion of local causality? In the following two Sections the relation between local causality and probabilistic characterization of the common cause will be studied in the lack of these two assumptions. First toy models will be introduced in which the two assumptions are violated, then the formal results will be gathered.

4 Non-atomic common causes

Example 1. Consider the following toy model. There are five lighthouses on the ocean in a line of equal distance from each other. (See Fig. 5.) Let us count them from left to right. In the middle one, that is

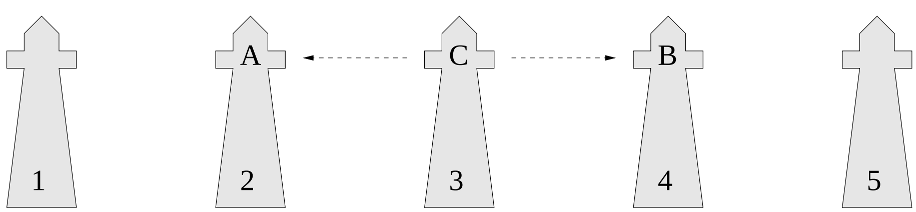

Figure 5: Lighthouses I.

in lighthouse 3 the lighthouse keeper C has three lamps, C', C'' and C'''. He has the following strategy to turn the lamps on: either he turns on only the lamp C', or only lamp C''', or all three lamps, or none. He never turns on the lamps in any other combination. He chooses between these four options with equal probability (say tossing two coins). Let us denote that a given lamp is turned on and off by C and $\overline{C}$, respectively. Using this notation the four possible state of the lamps are the following:

$$C_1 \equiv C' \wedge \overline{C}'' \wedge \overline{C}''' \tag{12}$$

$$C_2 \equiv \overline{C}' \wedge \overline{C}'' \wedge C''' \tag{13}$$

$$C_3 \equiv C' \wedge C'' \wedge C''' \tag{14}$$

$$C_4 \equiv \overline{C}' \wedge \overline{C}'' \wedge \overline{C}''' \tag{15}$$

each with probability

$$p(C_k) = \frac{1}{4} \tag{16}$$

Now, in the left neighboring lighthouse, that is in lighthouse 2, there is another lighthouse keeper, A and his role is simply to watch the light signals arriving from either the left or from the right that is from either lighthouse 1 or lighthouse 3. He does not know that lighthouse 1 is empty, therefore he spends equal time watching both neighboring lighthouses. Suppose furthermore that if he is watching left, he will miss the light signals coming from the right. This means that with probability $\frac{1}{2}$ he observes the signals coming from lighthouse 3 and with probability $\frac{1}{2}$ he will miss them. Denoting the event that the lighthouse keeper A is watching to the left and to the right by a_L and a_R, respectively and denoting by A the event that he observes a light signal (disregarding from which lamp), one obtains the following conditional probabilities:

$$p(A|a_m \wedge C_k) = \begin{cases} 1 & \text{if } m = R, k = 1,2,3 \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

In other words, the lighthouse keeper A observes the light signal only if he is watching right and there is a signal sent from C.

Suppose that the same thing happens also in lighthouse 4. The lighthouse keeper B is watching in both directions with equal probability, but since lighthouse 5 is empty, he misses the light signal coming from lighthouse 3 with probability $\frac{1}{2}$. Denoting again the events that the lighthouse keeper B is watching to the left and to the right by b_L and b_R, respectively and denoting by B the event that he observes a signal, one obtains the following conditional probabilities for B's observing a light signal:

$$p(B|b_n \wedge C_k) = \begin{cases} 1 & \text{if } n = L, k = 1,2,3 \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

This situation completely characterizes a probability space. The event algebra is generated by the following events:

$$A,\ \overline{A},\ B,\ \overline{B},\ a_m,\ b_n,\ C_k$$

with $m, n = L, R$ and $k = 1,2,3,4$. The event algebra has 64 atoms, 16 of which have non-zero probability:

$$\begin{aligned}
p(A \wedge B \wedge a_m \wedge b_n \wedge C_k) &= \frac{1}{16} && \text{if } m = R,\ n = L,\ k = 1,2,3 \\
p(A \wedge \overline{B} \wedge a_m \wedge b_n \wedge C_k) &= \frac{1}{16} && \text{if } m, n = R,\ k = 1,2,3 \\
p(\overline{A} \wedge B \wedge a_m \wedge b_n \wedge C_k) &= \frac{1}{16} && \text{if } m, n = L,\ k = 1,2,3 \\
p(\overline{A} \wedge \overline{B} \wedge a_m \wedge b_n \wedge C_k) &= \frac{1}{16} && \text{if } \begin{cases} m = L,\ n = R,\ k = 1,2,3, \\ \text{or } k = 4 \end{cases}
\end{aligned}$$

and the remaining 48 are of probability zero. By means of the probability of the atoms one can easily calculate the probability of any events of the algebra.

Now, it is easy to see that there is a correlation between events A and B that is between the lighthouse keepers' observing a light signal, both in the non-conditional and conditional sense:

$$\frac{3}{16} = p(A \wedge B) \quad \neq \quad p(A)\, p(B) = \frac{3}{8} \cdot \frac{3}{8} \tag{19}$$

$$\frac{3}{4} = p(A \wedge B|a_m \wedge b_n) \quad \neq \quad p(A|a_m)\, p(B|b_n) = \frac{3}{4} \cdot \frac{3}{4} \qquad \text{if } m = R, n = L \tag{20}$$

As one expects, the correlation is due to C's signaling: C_k is a local, (non-conspiratorial) joint common causal explanation of the correlation (20) in the sense of (2)-(5):

$$
\begin{aligned}
p(A \wedge B|a_m \wedge b_n \wedge C_k) = p(A|a_m \wedge b_n \wedge C_k)\, p(B|a_m \wedge b_n \wedge C_k) &= \begin{cases} 1 & \text{if } m = R, n = L, k = 1,2,3 \\ 0 & \text{otherwise} \end{cases} \\
p(A|a_m \wedge b_n \wedge C_k) = p(A|a_m \wedge b_{n'} \wedge C_k) &= \begin{cases} 1 & \text{if } m = R, k = 1,2,3 \\ 0 & \text{otherwise} \end{cases} \\
p(B|a_m \wedge b_n \wedge C_k) = p(B_n|a_{m'} \wedge b_n \wedge C_k) &= \begin{cases} 1 & \text{if } n = L, k = 1,2,3 \\ 0 & \text{otherwise} \end{cases} \\
p(a_m \wedge b_n \wedge C_k) = p(a_m \wedge b_n)\, p(C_k) &= \frac{1}{4} \cdot \frac{1}{4}
\end{aligned}
$$

Example 2. Suppose we take a coarser clustering of the switching of the lamps, say $D_1 \equiv C_1 \vee C_2 \vee C_3$ and $D_2 \equiv C_4$. Physically, D_1 is the event that *any* light is on in lighthouse 3, and D_2 is the event that *no* light is on. As one expects, for this coarser partition (2)-(5) will hold just as good as for the partition $\{C_k\}$:

$$
\begin{aligned}
p(A \wedge B|a_m \wedge b_n \wedge D_k) = p(A|a_m \wedge b_n \wedge D_k)\, p(B|a_m \wedge b_n \wedge D_k) &= \begin{cases} 1 & \text{if } m = R, n = L, k = 1 \\ 0 & \text{otherwise} \end{cases} \\
p(A|a_m \wedge b_n \wedge D_k) = p(A|a_m \wedge b_{n'} \wedge D_k) &= \begin{cases} 1 & \text{if } m = R, k = 1 \\ 0 & \text{otherwise} \end{cases} \\
p(B|a_m \wedge b_n \wedge D_k) = p(B_n|a_{m'} \wedge b_n \wedge D_k) &= \begin{cases} 1 & \text{if } n = L, k = 1 \\ 0 & \text{otherwise} \end{cases} \\
p(a_m \wedge b_n \wedge D_k) = p(a_m \wedge b_n)\, p(D_k) &= \begin{cases} \frac{1}{4} \cdot \frac{3}{4} & \text{if } n = L, k = 1 \\ \frac{1}{4} \cdot \frac{1}{4} & \text{otherwise} \end{cases}
\end{aligned}
$$

Thus, $\{D_k\}$ is also a local, (non-conspiratorial) joint common causal explanation of the correlation (20).

Example 3. Now, consider a coarser clustering of the switchings 'in the wrong way': $D'_1 \equiv C_1 \vee C_2 \vee C_4$ and $D'_2 \equiv C_3$ mixing together light ons and light offs. Contrary to the previous case, for this coarser partition the requirement of screening-off is violated. For example:

$$\frac{2}{3} = p(A \wedge B|a_R \wedge b_L \wedge D'_1) \quad \neq \quad p(A|a_R \wedge b_L \wedge D'_1)\, p(B|a_R \wedge b_L \wedge D'_1) = \frac{2}{3} \cdot \frac{2}{3}$$

(Locality and no-conspiracy will hold even in this case.) Hence $\{D'_k\}$ is *not* a local, (non-conspiratorial) joint common causal explanation of the correlation (20).

Now, let us consider the spacetime diagram of the above examples which is depicted in Fig. 6. Let $\mathfrak{N}$

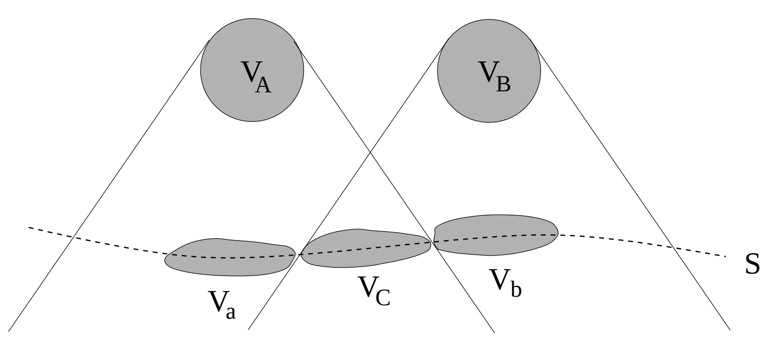

Figure 6: Spacetime diagram of Examples 1, 2 and 3.

be a *locally causal* net associating bounded spacetime regions to local algebras such that $A \in \mathcal{A}(V_A)$, $B \in \mathcal{A}(V_B)$, $a_m \in \mathcal{A}(V_a)$, $b_n \in \mathcal{A}(V_b)$ and $C_k, D_k, D'_k \in \mathcal{A}(V_C)$ for all m, n and k. As shown in Section 2, local causality of the net implies that the set $\{C_k\}$—being an atomic partition localized in the strong past $\mathcal{P}^S(V_A, V_B)$—satisfies (2)-(4), hence providing a local, joint common causal explanation of the correlation (20). (No-conspiracy (5), as already stressed above, is not a consequence of local causality but assumed in the model.) Thus, $\{C_k\}$ is an *atomic*, *strong*, local, non-conspiratorial joint common cause system.

What about *non-atomic* partitions localized in the strong past? As Examples 2 and 3 attest local causality has no bearing on this case. $\{D_k\}$ and $\{D'_k\}$ are all localized in $\mathcal{P}^S(V_A, V_B)$, but whereas $\{D_k\}$ is a common cause system of the correlation (20), $\{D'_k\}$ is *not*. This leads to the following

Moral 1. The probabilistic characterization of a local, joint common cause system $\{C_k\}$ *via* (2)-(4) *cannot* be justified by Bell's local causality applied to a net associating spacetime regions to local algebras, if $\{C_k\}$ is a *non-atomic* partition of $\mathcal{A}(V_C)$.

Thus, a coarse-grained (non-atomic) probabilistic common causal explanation of a correlation cannot be backed by Bell's spatiotemporal considerations on local causality. In the next Section we turn to the role of the other premise, namely the localization of the common cause in the strong past.

5 Weak common causes

Example 4. Now, let us modify the population of the lighthouses. Let A and B remain in their places that is in lighthouse 2 and 4, respectively, but suppose that lighthouses 1, 3 and 5 are inhabitated by three lighthouse keepers C', C'' and C''', respectively, each having the appropriate one of the three lamps introduced in the previous Section. (See Fig. 7.) That is suppose that now lighthouse keeper C' in

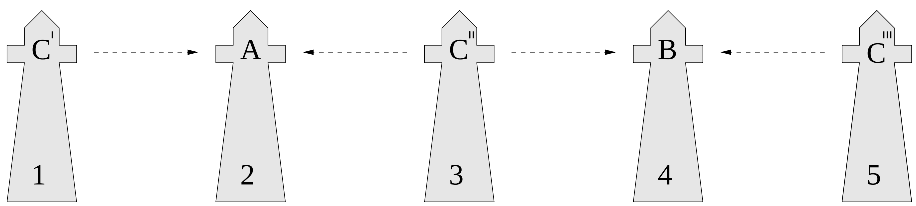

Figure 7: Lighthouses II.

lighthouse 1 operates lamp C', lighthouse keeper C'' in lighthouse 3 operates lamp C'' and lighthouse keeper C''' in lighthouse 5 operates lamp C'''. Suppose furthermore that the ons and offs of the different lamps follow just the same statistics defined in (12)-(16), that is $p(C_k) = \frac{1}{4}$ for every $k = 1, 2, 3, 4$ (only lamp C' is on, only lamp C''', all three lamps are on, none is on).

Now, the role of lighthouse keepers A and B is just as above to watch the light signals arriving at lighthouse 2 and 4, respectively. But now both can obtain a signal from both directions. Suppose that both A and B can only see the light signal sent from a neighboring lighthouse that is A cannot see the signal sent from C''' (say, it is two far or the lighthouses hide each other) and B cannot see the signal sent from C'. Now, again the event algebra has 16 atoms with non-zero probability:

$$
\begin{aligned}
p(A \wedge B \wedge a_m \wedge b_n \wedge C_k) &= \frac{1}{16} && \text{if } k = 3 \\
p(A \wedge \overline{B} \wedge a_m \wedge b_n \wedge C_k) &= \frac{1}{16} && \text{if } m = L,\ k = 1 \\
p(\overline{A} \wedge B \wedge a_m \wedge b_n \wedge C_k) &= \frac{1}{16} && \text{if } n = R,\ k = 2 \\
p(\overline{A} \wedge \overline{B} \wedge a_m \wedge b_n \wedge C_k) &= \frac{1}{16} && \text{if } \begin{cases} m = R,\ k = 1, \\ \text{or } n = L,\ k = 2, \\ \text{or } k = 4 \end{cases}
\end{aligned}
$$

and there is a conditional and non-conditional correlation between event A and B, the detections of light signals in lighthouse 2 and 4, respectively both in the non-conditional and conditional sense:

$$\frac{1}{4} = p(A \wedge B) \quad \neq \quad p(A)\, p(B) = \frac{3}{8} \cdot \frac{3}{8} \tag{21}$$

$$\frac{1}{4} = p(A \wedge B | a_m \wedge b_n) \quad \neq \quad p(A|a_m)\, p(B|b_n) = \begin{cases} \frac{1}{4} \cdot \frac{1}{4} & \text{if } m = R, n = L, \\ \frac{1}{4} \cdot \frac{1}{2} & \text{if } m, n = R, \\ \frac{1}{2} \cdot \frac{1}{4} & \text{if } m, n = L. \end{cases} \tag{22}$$

As one expects, $\{C_k\}$ is a local, (non-conspiratorial) joint common causal explanation of the correlation:

$$\begin{aligned} p(A \wedge B | a_m \wedge b_n \wedge C_k) = p(A|a_m \wedge b_n \wedge C_k)\, p(B|a_m \wedge b_n \wedge C_k) \quad &= \quad \begin{cases} 1 & \text{if } m = R, n = L, k = 3 \\ 0 & \text{otherwise} \end{cases} \\ p(A|a_m \wedge b_n \wedge C_k) = p(A|a_m \wedge b_{n'} \wedge C_k) \quad &= \quad \begin{cases} 1 & \text{if } m = L, k = 1 \\ 1 & \text{if } k = 3 \\ 0 & \text{otherwise} \end{cases} \\ p(B|a_m \wedge b_n \wedge C_k) = p(B_n|a_{m'} \wedge b_n \wedge C_k) \quad &= \quad \begin{cases} 1 & \text{if } m = R, k = 2 \\ 1 & \text{if } k = 3 \\ 0 & \text{otherwise} \end{cases} \\ p(a_m \wedge b_n \wedge C_k) = p(a_m \wedge b_n)\, p(C_k) \quad &= \quad \frac{1}{4} \cdot \frac{1}{4} \end{aligned}$$

Now, consider again the spacetime diagram of Example 4 depicted in Fig. 8. Here $\{C_k\}$ is localized *not*

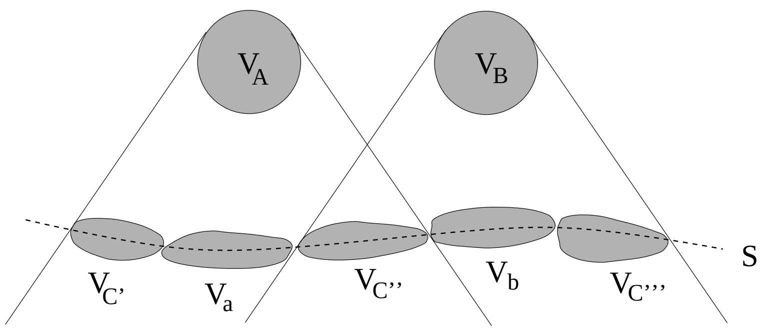

Figure 8: Spacetime diagram of Example 4.

in the *strong* past but in the *weak* past of the correlating events. How these weak common causes relate to Bell's local causality? This question is answered in the following

Proposition 1. Let $\mathfrak{N}$ be again a *locally causal* net associating bounded spacetime regions to local algebras such that $A \in \mathcal{A}(V_A)$, $B \in \mathcal{A}(V_B)$, $a_m \in \mathcal{A}(V_a)$, $b_n \in \mathcal{A}(V_b)$ for all m and n, and for the partition

$$\{C_{ijk}\} \equiv \{C'_i \wedge C''_j \wedge C'''_l\}$$

(where $C'_1 \equiv C'$ and $C'_2 \equiv \overline{C}'$ and similarly for C''_j and C'''_l) $C'_i \in \mathcal{A}(V_{C'})$, $C''_j \in \mathcal{A}(V_{C''})$ and $C'''_l \in \mathcal{A}(V_{C'''})$ for all i, j and l. Then $\{C_{ijk}\}$ is a *weak*, local, joint common cause of the conditional correlations

$$p(A \wedge B | a_m \wedge b_n) \neq p(A|a_m)\, p(B|b_n) \tag{23}$$

in the sense that the following equations hold:

$$p(A \wedge B_n | a_m \wedge b_n \wedge C_{ijk}) \quad = \quad p(A|a_m \wedge b_n \wedge C_{ijk})\, p(B|a_m \wedge b_n \wedge C_{ijk}) \tag{24}$$

$$p(A|a_m \wedge b_n \wedge C_{ijk}) \quad = \quad p(A|a_m \wedge b_{n'} \wedge C_{ijk}) \tag{25}$$

$$p(B|a_m \wedge b_n \wedge C_{ijk}) \quad = \quad p(B|a_{m'} \wedge b_n \wedge C_{ijk}) \tag{26}$$

Proof. Since $\{C''_j\}$ is an *atomic* partition localized in the strong past $\mathcal{P}^S(V_A, V_B)$, local causality of the net implies that for any event $a'_{im} \equiv C'_i \wedge a_m \in \mathcal{A}(V_{C'} \cup V_a)$, $b'_{nl} \equiv b_n \wedge C'''_l \in \mathcal{A}(V_b \cup V_{C'''})$ and atomic event C''_j the following will hold:

$$\begin{aligned} p(A \wedge B_n | a'_{im} \wedge b'_{nl} \wedge C''_j) &= p(A | a'_{im} \wedge b'_{nl} \wedge C''_j)\, p(B | a'_{im} \wedge b'_{nl} \wedge C''_j) && (27) \\ p(A | a'_{im} \wedge b'_{nl} \wedge C''_j) &= p(A | a'_{im} \wedge b'_{n'l'} \wedge C''_j) && (28) \\ p(B | a'_{im} \wedge b'_{nl} \wedge C''_j) &= p(B | a'_{i'm'} \wedge b'_{nl} \wedge C''_j) && (29) \end{aligned}$$

In other words, $\{C''_j\}$ is a *strong*, local, joint common cause of the conditional correlations

$$p(A \wedge B | a'_{im} \wedge b'_{nl}) \neq p(A | a'_{im})\, p(B | b'_{nl}) \tag{30}$$

with the *new* conditions a'_{im} and b'_{nl}. (Again, no-conspiracy

$$p(a'_{im} \wedge b'_{nl} \wedge C''_j) = p(a'_{im} \wedge b'_{nl})\, p(C''_j) \tag{31}$$

does *not* follow from local causality of the net.) But (27)-(29) are just equivalent to (32)-(34) proving that $\{C_{ijk}\}$ is a *weak*, local, joint common cause of the conditional correlations (23). ■

This leads to

Moral 2. The probabilistic characterization of a local, joint common cause system $\{C_{ijk}\}$ *via* (2)-(4) can be justified by Bell's local causality applied to a net associating spacetime regions to local algebras, if $\{C_{ijk}\} \equiv \{C'_i \wedge C''_j \wedge C'''_l\}$ is a *weak* common cause ($C'_i \in \mathcal{A}(V_{C'})$, $C''_j \in \mathcal{A}(V_{C''})$ and $C'''_l \in \mathcal{A}(V_{C'''})$) and C''_j is an *atomic* partition of $\mathcal{A}(V_{C''})$.

In the Example 4 one might have found it peculiar that although the common cause $\{C_{ijk}\}$ was non-conspiratorial (it was probabilistically independent of a_m and b_n), still there was a 'conspiracy' *within* the common cause: C'_i, C''_j and C'''_l were *not* probabilistically independent. For example it never happened that only lamp C'' was switched on. This fact does not raise any problem until the common cause is localized at one place, as in Example 1, where all the three lamps were localized in lighthouse 3. But in Example 4 the common cause was scattered around in three different locations. It was located in three different lighthouses. The problem with such a common cause that it may well question our whole project to provide a common causal explanation for a correlation. If the *explanans* itself has a built-in correlation, then what is the point of using it for explaining correlations? Can we not come up with a common causal model in which C'_i, C''_j and C'''_l are spatially separated but still independent, say, regulated by three independent coin tossings in lighthouse 1, 3 and 5, respectively. Can one obtain a *weak* common cause for a given correlation without a *built-in correlation*?

Let $\{C_{ijk}\} \equiv \{C'_i \wedge C''_j \wedge C'''_l\}$ be a weak common cause of a given correlation. (Here $\{C'_i\}, \{C''_j\}$ and $\{C'''_l\}$ are general partitions of $\mathcal{A}(V_{C'})$, $\mathcal{A}(V_{C''})$ and $\mathcal{A}(V_{C'''})$, respectively, and not those specified in the above Examples.) Let us call $\{C_{ijk}\}$ a *genuine weak common cause*, iff $\{C''_j\}$—the 'middle part' of $\{C_{ijk}\}$—is *not* a strong common cause. In what follows we will show that the above mentioned 'built-in correlation' is a necessary condition to explain a correlation by a genuine weak common cause. In other words, we will show that if $\{C_{ijk}\} \equiv \{C'_i \wedge C''_j \wedge C'''_l\}$ is a common cause of the correlation (23) and C'_i, C''_j and C'''_l are probabilistically independent, then also $\{C''_j\}$ will be a common cause of the correlation.

Proposition 2. Suppose that $\{C'_i \wedge C''_j \wedge C'''_l\}$ is a common cause of the correlation between A_m and B_n in the sense that the following equations hold:

$$\begin{aligned} p(A_m \wedge B_n | a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) &= p(A_m | a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l)\, p(B_n | a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) && (32) \\ p(A_m | a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) &= p(A_m | a_m \wedge b_{n'} \wedge C'_i \wedge C''_j \wedge C'''_l) && (33) \\ p(B_n | a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) &= p(B_n | a_{m'} \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) && (34) \\ p(a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) &= p(a_m \wedge b_n)\, p(C'_i \wedge C''_j \wedge C'''_l) && (35) \end{aligned}$$

and suppose that C'_i, C''_j and C'''_l are independent that is

$$p(C'_i \wedge C''_j \wedge C'''_l) = p(C'_i)\, p(C''_j)\, p(C'''_l) \tag{36}$$

then $\{C''_j\}$ is also a common cause the correlation:

$$p(A_m \wedge B_n | a_m \wedge C''_j) = p(A_m | a_m \wedge b_n \wedge C''_j)\, p(B_n | a_m \wedge b_n \wedge C''_j) \tag{37}$$
$$p(A_m | a_m \wedge b_n \wedge C''_j) = p(A_m | a_m \wedge b_{n'} \wedge C''_j) \tag{38}$$
$$p(B_n | a_m \wedge b_n \wedge C''_j) = p(B_n | a_{m'} \wedge b_n \wedge C''_j) \tag{39}$$
$$p(a_m \wedge b_n \wedge C''_j) = p(a_m \wedge b_n)\, p(C''_j) \tag{40}$$

For the proof see Appendix B. Since in Example 4 $\{C_{ijk}\} \equiv \{C'_i \wedge C''_j \wedge C'''_l\}$ was localized in the weak past and $\{C''_j\}$ was localized in the strong past, we can interpret Proposition 2 as follows: a weak common cause with no 'built-in correlation' is always parasitic on a strong common cause in the sense that there is no other way to provide a genuine weak common cause for a given correlation, then to make the spatially separated parts of the common cause probabilistically dependent. In brief, there is no genuine weak common cause without 'built-in correlation'.

6 Conclusion

The probabilistic characterization of the common cause can be justified *via* Bell's notion of local causality if two assumptions concerning the common cause are made: first, the common cause is localized in the strong past of the correlating events; second, it provides a complete specification of the 'beables' of this past. In the paper it was argued that only the second assumption, that is complete specification, is necessary for the derivation of the probabilistic common cause from local causality. Thus, coarse-grained common causal explanations cannot be rationalized in this way. (Whether it can be justified in other ways, based on non-spatiotemporal considerations, is not investigated here. For a justification *via* Causal Markov Condition see (Glymour 2006).)

Concerning the first assumption, namely localization in the strong past, it was shown that genuine weak common causes can be provided for a given correlation only at the cost of introducing a 'built-in correlation' between the spatially separated parts of the common cause.

We conclude the paper with a highly speculative question. As it was shown in the Introduction, the common causes that naturally arise in algebraic quantum field theory are *weak* and *not strong* common causes.

Question: Is this fact somehow related to or a consequence of the following two facts? (If they are facts at all.)

1. In algebraic quantum field theory quantum states establishing a superluminal correlation between two spacelike separated events, also establish (or 'typically' establish) a 'built-in correlation' between the spacelike separated parts of the weak common causes of this correlation.

2. An analoge of Proposition 2 holds in algebraic quantum field theory stating that a 'built-in correlation' is a necessary condition to explain a correlation by a genuine weak common cause.

Were these two facts to hold, one could understand why weak common causes in algebraic quantum field theory are geniune common causes that is why they do not reduce to strong common causes. (For more on this see (Hofer-Szabó and Vecsernyés, 2014).)

Acknowledgements. I wish to thank the Reading Group of Fellows in the Center for Philosophy of Science at the University of Pittsburgh and the Castle Hill Reading Group in Philosophy of Science in the Institute of Philosophy at the Hungarian Academy of Sciences for helpful comments on an earlier version of the paper. This work has been supported by the Hungarian Scientific Research Fund, OTKA K-100715 and by the Fulbright Research Grant while G. H-Sz. was a Visiting Fellow in the Center for Philosophy of Science at the University of Pittsburgh.

Appendix A

Here we will show that if a set of correlations $\{(A_m, B_n) | m, n = 1, 2\}$ has a local, non-conspiratorial joint common causal explanation in the sense of (2)-(5), then the following Clauser–Horne inequalities have to hold for any $m, m', n, n' = 1, 2$; $m \neq m', n \neq n'$:

$$-1 \leqslant p(A_m \wedge B_n | a_m \wedge b_n) + p(A_m \wedge B_{n'} | a_m \wedge b_{n'}) + p(A_{m'} \wedge B_n | a_{m'} \wedge b_n) \\ -p(A_{m'} \wedge B_{n'} | a_{m'} \wedge b_{n'}) - p(A_m | a_m \wedge b_n) - p(B_n | a_m \wedge b_n) \leqslant 0 \tag{41}$$

The derivation of (41) from (2)-(5) is simple. It is an elementary fact of arithmetic that for any $\alpha, \alpha', \beta, \beta' \in [0, 1]$ the number

$$\alpha\beta + \alpha\beta' + \alpha'\beta - \alpha'\beta' - \alpha - \beta \tag{42}$$

lies in the interval $[-1, 0]$. Now let $\alpha, \alpha', \beta, \beta'$ be the following conditional probabilities:

$$\alpha \equiv p(A_m | a_m \wedge b_n \wedge C_k) \tag{43}$$

$$\alpha' \equiv p(A_{m'} | a_{m'} \wedge b_{n'} \wedge C_k) \tag{44}$$

$$\beta \equiv p(B_n | a_m \wedge b_n \wedge C_k) \tag{45}$$

$$\beta' \equiv p(B_{n'} | a_{m'} \wedge b_{n'} \wedge C_k) \tag{46}$$

Plugging (43)-(46) into (42) and using locality (3)-(4) one obtains

$$-1 \leqslant p(A_m | a_m \wedge b_n \wedge C_k) p(B_n | a_m \wedge b_n \wedge C_k) + p(A_m | a_{m'} \wedge b_n \wedge C_k) p(B_{n'} | a_{m'} \wedge b_n \wedge C_k) \\ +p(A_{m'} | a_{m'} \wedge b_n \wedge C_k) p(B_n | a_{m'} \wedge b_n \wedge C_k) - p(A_{m'} | a_{m'} \wedge b_{n'} \wedge C_k) p(B_{n'} | a_{m'} \wedge b_{n'} \wedge C_k) \\ -p(A_m | a_m \wedge b_n \wedge C_k) - p(B_n | a_m \wedge b_n \wedge C_k) \leqslant 0 \tag{47}$$

Using screening-off (2) one obtains

$$-1 \leqslant p(A_m \wedge B_n | a_m \wedge b_n \wedge C_k) + p(A_m \wedge B_{n'} | a_{m'} \wedge b_n \wedge C_k) \\ +p(A_{m'} \wedge B_n | a_{m'} \wedge b_n \wedge C_k) - p(A_{m'} \wedge B_{n'} | a_{m'} \wedge b_{n'} \wedge C_k) \\ -p(A_m | a_m \wedge b_n \wedge C_k) - p(B_n | a_m \wedge b_n \wedge C_k) \leqslant 0 \tag{48}$$

Finally, multiplying the above inequality by $p(C_k)$, then summing up for the indices k and using no-conspiracy (5) one arrives at (41).

Appendix B

Here we prove Proposition 1. Suppose that $\{C'_i \wedge C''_j \wedge C'''_l\}$ is a common cause of the correlation between A_m and B_n in the sense of (32)-(35) and suppose that C'_i, C''_j and C'''_l are independent in the sense of (36). First, observe that (35) and (36) together entail that:

$$p(a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) = p(a_m \wedge b_n)\, p(C'_i) p(C''_j) p(C'''_l) \tag{49}$$

Then C''_j is a *strong* common cause that is (37)-(40) hold:

$$
\begin{aligned}
p(A_m \wedge B_n | a_m \wedge b_n \wedge C''_j) &= \frac{p(A_m \wedge B_n \wedge a_m \wedge b_n \wedge C''_j)}{p(a_m \wedge b_n \wedge C''_j)} \\
&\overset{(49)}{=} \frac{\sum_{il} p(A_m \wedge B_n | a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) p(a_m \wedge b_n) p(C'_i) p(C''_j) p(C'''_l)}{p(a_m \wedge b_n) p(C''_j)} \\
&\overset{(32)}{=} \sum_{il} p(A_m | a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) p(B_n | a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) p(C'_i) p(C'''_l) \\
&\overset{(33)(34)}{=} \sum_{il} p(A_m | a_m \wedge b_n \wedge C'_i \wedge C''_j) p(B_n | a_m \wedge b_n \wedge C''_j \wedge C'''_l) p(C'_i) p(C'''_l) \\
&\overset{(49)}{=} p(A_m | a_m \wedge b_n \wedge C''_j)\, p(B_n | a_m \wedge b_n \wedge C''_j) \\
p(A_m | a_m \wedge b_n \wedge C''_j) &= \frac{p(A_m \wedge a_m \wedge b_n \wedge C''_j)}{p(a_m \wedge b_n \wedge C''_j)} \\
&\overset{(49)}{=} \frac{\sum_{il} p(A_m | a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) p(a_m \wedge b_n) p(C'_i) p(C''_j) p(C'''_l)}{p(a_m \wedge b_n) p(C''_j)} \\
&\overset{(33)}{=} \sum_{il} p(A_m | a_m \wedge b_{n'} \wedge C'_i \wedge C''_j \wedge C'''_l) p(C'_i) p(C'''_l) \\
&= \frac{\sum_{il} p(A_m | a_m \wedge b_{n'} \wedge C'_i \wedge C''_j \wedge C'''_l) p(a_m \wedge b_{n'}) p(C'_i) p(C''_j) p(C'''_l)}{p(a_m \wedge b_{n'}) p(C''_j)} \\
&\overset{(49)}{=} \frac{p(A_m \wedge a_m \wedge b_{n'} \wedge C''_j)}{p(a_m \wedge b_{n'} \wedge C''_j)} = p(A_m | a_m \wedge b_{n'} \wedge C''_j) \\
p(B_n | a_m \wedge b_n \wedge C''_j) &= \frac{p(B_n \wedge a_m \wedge b_n \wedge C''_j)}{p(a_m \wedge b_n \wedge C''_j)} \\
&\overset{(49)}{=} \frac{\sum_{il} p(B_n | a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) p(a_m \wedge b_n) p(C'_i) p(C''_j) p(C'''_l)}{p(a_m \wedge b_n) p(C''_j)} \\
&\overset{(34)}{=} \sum_{il} p(B_n | a_{m'} \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) p(C'_i) p(C'''_l) \\
&= \frac{\sum_{il} p(B_n | a_{m'} \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) p(a_{m'} \wedge b_n) p(C'_i) p(C''_j) p(C'''_l)}{p(a_{m'} \wedge b_n) p(C''_j)} \\
&\overset{(49)}{=} \frac{p(B_n \wedge a_{m'} \wedge b_n \wedge C''_j)}{p(a_{m'} \wedge b_n \wedge C''_j)} = p(B_n | a_{m'} \wedge b_n \wedge C''_j) \\
p(a_m \wedge b_n \wedge C''_j) &= \sum_{il} p(a_m \wedge b_n \wedge C'_i \wedge C''_j \wedge C'''_l) \\
&\overset{(49)}{=} \sum_{il} p(a_m \wedge b_n \wedge C'_i \wedge C'''_l) p(C''_j) = p(a_m \wedge b_n)\, p(C''_j)
\end{aligned}
$$

where the numbers over the equation signs refer to the equation used at that step.

References

Bell, J.S., "Beables for quantum field theory," (TH-2053-CERN, presented at the Sixth GIFT Seminar, Jaca, 2–7 June 1975) reprinted in J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, (Cambridge: Cambridge University Press, 1987, 52-62.).

J. Butterfield, "A Spacetime approach to the Bell inequality," in: J. Cushing and E. McMullin (eds.), *Philosophical Consequences of Quantum Theory*, Notre Dame, 114-144, (1989).

J. Butterfield, "Stochastic Einstein Locality Revisited," *British Journal for The Philosophy of Science*, **58**, 805-867, (2007).

C. Glymour, "Markov properties and quantum experiments," in W. Demopoulos and I. Pitowsky (eds.) *Physical Theory and its Interpretation*, (Springer, 117-126, 2006).

J. Henson, "Comparing causality principles," *Studies in the History and Philosophy of Modern Physics*, **36**, 519-543 (2005).

J. Henson, "Comfounding causality principles: Comment on Rédei and San Pedro's "Distinguishing causality principles"," *Studies in the History and Philosophy of Modern Physics*, **44**, 17-19 (2013).

J. Henson, "Non-separability does not relieve the problem of Bell's theorem," URL = http://arxiv.org/abs/1302.7188. (2013).

G. Hofer-Szabó and P. Vecsernyés, "Reichenbach's Common Cause Principle in AQFT with locally finite degrees of freedom," *Found. Phys.*, **42**, 241-255 (2012a).

G. Hofer-Szabó, P. Vecsernyés, "Noncommuting local common causes for correlations violating the Clauser–Horne inequality," *Journal of Mathematical Physics*, **53**, 12230 (2012b).

G. Hofer-Szabó, P. Vecsernyés, "Noncommutative Common Cause Principles in AQFT," *Journal of Mathematical Physics*, **54**, 042301 (2013a).

G. Hofer-Szabó, P. Vecsernyés, "Bell inequality and common causal explanation in algebraic quantum field theory," *Studies in the History and Philosophy of Modern Physics*, (forthcoming) (2013b).

G. Hofer-Szabó, P. Vecsernyés, "On the concept of local causality in local classical and quantum theory," (in preparation) (2014).

J. Jarrett, "On the Physical Significance of the Locality Conditions in Bell Arguments," *Nous*, **18**, 569-589, (1984).

T. Norsen, "J.S. Bell's concept of local causality," *Am. J. Phys*, **79**, 12, (2011).

H. Reichenbach, *The Direction of Time*, (University of California Press, Los Angeles, 1956).

M. Rédei, "Reichenbach's Common Cause Principle and quantum field theory," *Found. Phys.*, **27**, 1309–1321 (1997).

M. Rédei and J. S. Summers, "Local primitive causality and the Common Cause Principle in quantum field theory," *Found. Phys.*, **32**, 335-355 (2002).

M. Rédei and I. San Pedro, "Distinguishing causality principles," *Studies in the History and Philosophy of Modern Physics*, **43**, 84-89 (2012).

Shimony, A., "Events and processes in the quantum world," in: C. Isham and R. Penrose (eds.), *Quantum Concepts in Space and Time*, Oxford University Press, London/New York, 1986.

Van Fraassen, B. C., "The Charybdis of Realism: Epistemological Implications of Bell's Inequality," *Synthese*, **52**, 25-38 (1982).