

Characterizing common cause closed probability spaces

October 12, 2010

ZALÁN GYENIS

Department of Mathematics
Central European University
gyz@renyi.hu

MIKLÓS RÉDEI[†]

Department of Philosophy, Logic and Scientific Method
London School of Economics and Political Science
M.redei@lse.ac.uk

Abstract. A classical probability measure space was defined in earlier papers [14], [9] to be common cause closed if it contains a Reichenbachian common cause of every correlation in it, and common cause incomplete otherwise. It is shown that a classical probability measure space is common cause incomplete if and only if it contains more than one atom. Furthermore, it is shown that every probability space can be embedded into a common cause closed one; which entails that every classical probability space is common cause completable with respect to any set of correlated events. The implications of these results for Reichenbach's Common Cause Principle are discussed, and it is argued that the Principle is only falsifiable if conditions on the common cause are imposed that go beyond the requirements formulated by Reichenbach in the definition of common cause.

Keywords: Reichenbachian common cause, Common Cause Principle, causally closed theories

[†]Work supported in part by the Hungarian Scientific Research Found (OTKA), contract number: K68043

1 The problem of common cause closedness – informal definition and overview of paper

The aim of this paper is to prove new results on common cause closedness and common cause completability of classical, Kolmogorovian probability measure spaces. The propositions presented here give a complete characterization of common cause closed classical probability spaces and show that classical probability measure spaces are common cause completable in general in the very strong sense of being embeddable into a common cause closed probability space. These results solve several open problems that have been formulated in earlier papers, especially in [14], [15], [9].

Informally, a classical probability space (X, \mathcal{S}, p) (with the set X of elementary events, \mathcal{S} a Boolean algebra of some subsets of X and $p: \mathcal{S} \rightarrow [0, 1]$ a probability measure) is *common cause closed* (or common cause *complete*) if for every pair of events (a, b) that are correlated in p there exists an event $c \in \mathcal{S}$ that is a (Reichenbachian) common cause of the correlation. The latter means that c satisfies the conditions in Definition 2.1. If \mathcal{S} contains correlated events a, b for which there exists no common cause in \mathcal{S} , then (X, \mathcal{S}, p) is called common cause *incomplete*.

It is known that common cause incomplete probability spaces exist and it also is known that common cause closedness, while non-trivial, is not an impossible feature of probability spaces: It was shown (Proposition 2 in [9]) that no probability space with a Boolean algebra of finite cardinality can be common cause closed and that purely non-atomic probability spaces are common cause closed (Proposition 7 in [9]). (See Definition 2.2 for the measure-theoretic meaning of non-atomicity). Since probability spaces with a Boolean algebra of finite cardinality are atomic, these results indicate that the (lack of) atomicity of (X, \mathcal{S}, p) is directly relevant from the perspective of common cause closedness of (X, \mathcal{S}, p) ; however, the precise nature of the role of atomicity remained unclear in [9]. In this paper we clarify this role by showing (Theorem 3.8) that the probability space (X, \mathcal{S}, p) is common cause closed if and only if there is at most one atom in it.

If a probability space (X, \mathcal{S}, p) is common cause incomplete, then the question arises whether it can be common cause completed. Common cause completion of (X, \mathcal{S}, p) with respect to a set $\{(a_i, b_i) : i \in I\}$ of pairs of correlated elements is meant finding an extension (X', \mathcal{S}', p') of (X, \mathcal{S}, p) (Definition 4.1) such that (X', \mathcal{S}', p') contains a common cause $c_i \in \mathcal{S}'$ for every correlated pair (a_i, b_i) , $i \in I$. It was shown in [14] that every common cause incomplete probability space is common cause completable with respect to any *finite* set of correlations. It remained open in [14] however whether common cause completability holds with respect to an *infinite* set of correlations as well. We prove in this paper that this is the

case; in fact, we prove much more: it will be shown that every common cause incomplete probability space has an extension that is common cause *closed*.

Common cause (in)completeness of probability spaces is relevant from the perspective of what is called Reichenbach's Common Cause Principle: According to this Principle, which goes back to Reichenbach's work [25] where the notion of common cause is defined explicitly the first time, correlations are signs of causal relations; to be more precise, the Principle asserts that if two events are correlated then either it is the case that the correlation is due to a direct casual link between the correlated events, or, if a direct causal connection between the correlated entities is excluded for some reason, then there exists a common cause that explains the correlation. The Common Cause Principle has been extensively discussed in the literature (see the references in Section 5); in particular it has been debated whether the Principle holds in Nature. Common cause (in)completeness and common cause completability have ramifications for how Reichenbach's Common Cause Principle can(not) be falsified, and in Section 5 we will discuss how the presented technical results on common cause closedness and completability contribute to the further specification of the conditions under which such a falsification is feasible.

The structure of the paper is the following: In Section 2 we give the basic definitions and fix the notation. Section 3 characterizes common cause closedness in terms of atomicity. Section 4 presents the general common cause completability result. This result is based on a construction that provides extensions of classical probability spaces. Since existence of extensions of probability spaces are important for other problems in philosophy of science involving probability theory (e.g. in Bayesianism), this construction has some interest in its own right but it is somewhat technical and the technical details (together with the proofs of the mathematical claims involved in the construction) are collected in the Appendix.

2 Notation, definitions

In what follows, (X, \mathcal{S}, p) denotes a classical probability measure space (*probability space*, for short), where X is a set and \mathcal{S} is a Boolean algebra of some subsets of X and p is a σ -additive measure on \mathcal{S} . Typically, and throughout this paper, \mathcal{S} is assumed to be a σ -algebra. Whenever it is not important on which set X the σ -algebra \mathcal{S} is represented, we write (\mathcal{S}, p) and call (\mathcal{S}, p) a probability space.

Given a classical probability measure space, events $a, b \in \mathcal{S}$ are called correlated in p if

$$\text{Corr}_p(a, b) \doteq p(a \wedge b) - p(a)p(b) > 0 \tag{1}$$

Definition 2.1. Event $c \in \mathcal{S}$ is said to be the *common cause of the correlation* (1) if the following (independent) conditions hold:

$$p(a \wedge b|c) = p(a|c)p(b|c) \quad (2)$$

$$p(a \wedge b|c^\perp) = p(a|c^\perp)p(b|c^\perp) \quad (3)$$

$$p(a|c) > p(a|c^\perp) \quad (4)$$

$$p(b|c) > p(b|c^\perp) \quad (5)$$

where $p(x|y) = p(x \wedge y)/p(y)$ denotes the conditional probability of x on condition y , and it is assumed that none of the probabilities $p(x)$, $x = a, b, c, c^\perp$ is equal to zero.

Note that taking c to be either a or b , conditions (2)–(5) are satisfied, so, formally, both a and b are common causes of the correlation (1) between a and b ; intuitively however such a “common cause” is not a “proper” common cause. A common cause is *proper* if it differs from both a and b by more than a measure zero set (see below the remarks about treating measure zero sets). In what follows, a common cause will *always* mean a proper common cause.

In connection with every problem involving measure spaces, one has to deal with measure zero sets. Since we have in the background the Common Cause Principle in mind, especially its status in physics, and since from a physical point of view, measure zero sets are frequently irrelevant, we want to exclude measure zero sets. One way of doing this is to assume that p is *faithful* in the sense that $p(a) = 0$ ($a \in \mathcal{S}$) entails $a = \emptyset$. Faithfulness of p can always be achieved by passing from (X, \mathcal{S}, p) to the corresponding *probability algebra*: Denoting by \mathcal{N} the σ -ideal of \mathcal{S} consisting of measure-zero elements of \mathcal{S} , one can factorize \mathcal{S} by \mathcal{N} and extend p from \mathcal{S} to the quotient Boolean algebra $\hat{\mathcal{S}} \doteq \mathcal{S}/\mathcal{N}$ on which the extended \hat{p} is a faithful probability measure (see [10, p. 166] for details).

Note that $a, b \in \mathcal{S}$ are (positively) correlated if and only if their representatives \hat{a}, \hat{b} in the corresponding probability algebra $\hat{\mathcal{S}}$ are correlated; moreover $c \in \mathcal{S}$ is a common cause of the correlation between a and b if and only if \hat{c} in $\hat{\mathcal{S}}$ is a common cause of correlation between \hat{a} and \hat{b} . This means that investigating a probability space or its corresponding probability algebra (w.r.t common cause) are *equivalent*. This allows us to work with the probability algebra instead of the concrete probability space. A further advantage of working in the probability algebra rather than in the probability space is that the notions of atom in the algebraic and measure theoretic sense coincide in a probability algebra:

Definition 2.2. Let (\mathcal{S}, p) be a probability space with \leq as the non-strict (and $<$ as the corresponding strict) partial order arising from the Boolean lattice operations in \mathcal{S} .

1. Element $a \in \mathcal{S}$ is said to be a p -atom if $p(a) > 0$ and for all $b < a$ we have $p(b) = p(a)$ or $p(b) = 0$.
2. The space (\mathcal{S}, p) is *purely atomic* if for every $a \in A$ with $p(a) \neq 0$ there exists a p -atom $b \in A$ such that $b \leq a$.
3. The space is *purely non-atomic* if there are no p -atoms in it, that is, for all $a \in A$ with $p(a) \neq 0$ there exists $b < a$ such that $0 < p(b) < p(a)$.

Note that p -atoms and algebraic-atoms are different notions in general (see examples below); however, in the probability algebra these notions coincide: If a is a p -atom and $b < a$ with $p(b) = p(a)$ then a and b are equal modulo measure zero, hence in the corresponding probability algebra they are the same, so it follows that in a probability algebra every p -atom is also an algebraic-atom. (The converse is obvious.) Moreover a is a p -atom in (\mathcal{S}, p) if and only if \hat{a} is a \hat{p} -atom in $(\mathcal{S}/\mathcal{N}, \hat{p})$. This motivates the following definition:

Definition 2.3. For a probability space (\mathcal{S}, p) the *number of p -atoms in (\mathcal{S}, p)* is the cardinality of the set of \hat{p} atoms in the corresponding probability algebra $(\mathcal{S}/\mathcal{N}, \hat{p})$ (which is equal to the cardinality of the set of algebraic atoms in $\hat{\mathcal{S}}$).

We also note that every measure space decomposes into purely non-atomic and purely atomic parts, and so do probability measure spaces as well: for any probability measure space (X, \mathcal{S}, p) there exist mutually singular probability measures p_0 and p_1 such that p is a convex combination of p_0 and p_1 , and the spaces (X, \mathcal{S}, p_0) and (X, \mathcal{S}, p_1) are respectively purely non-atomic and purely atomic. Two measures p_0 and p_1 are mutually singular by definition if there exists $Z \subset X$ such that for all $a \in \mathcal{S}$ we have $(Z \cap a) \in \mathcal{S}$ and $(Z^\perp \cap a) \in \mathcal{S}$ and $p_0(Z \cap a) = p_1(Z^\perp \cap a) = 0$ (see [10, p. 126]).

Examples:

1. Every finite probability space is purely atomic.
2. Let $([0, 1], \mathfrak{L}, \lambda)$ denote the space of Lebesgue-measurable subsets of the interval $[0, 1]$ with the Lebesgue-measure λ . This probability space is purely non-atomic; note however that the one element sets $\{x\}$ ($x \in [0, 1]$) are algebraic atoms in \mathfrak{L} , all having measure zero.
3. If $r \in (0, 1)$ and $x \in [0, 1]$ are fixed numbers, and p' is the probability space on \mathfrak{L} defined by

$$p' \doteq (1 - r) \cdot \lambda + r \cdot \delta_{\{x\}}$$

where $\delta_{\{x\}}$ is the Dirac-measure concentrated on $\{x\}$, then the probability space $([0, 1], \mathfrak{L}, p')$ has exactly one atom, which is $\{x\}$, thus it is not purely non-atomic and not purely atomic.

From now on we assume that p in the probability space (\mathcal{S}, p) is faithful (equivalently: that (\mathcal{S}, p) is a probability algebra).

3 Characterization of common cause closed spaces

Definition 3.1. The probability space (X, \mathcal{S}, p) is *common cause incomplete* if there exist events $a, b \in \mathcal{S}$ such that a, b are correlated in p but \mathcal{S} does not contain a common cause of the correlation. The space (X, \mathcal{S}, p) is *common cause closed (complete)* if for every pair of correlated events $a, b \in \mathcal{S}$ there exists a common cause $c \in \mathcal{S}$.

There exist trivially common cause closed spaces: the ones that do not contain any correlated events (for instance if p is the probability measure concentrated at one single point in X then the probability space is trivially common cause closed). Since we are interested in non-trivial common cause closedness, in what follows, “common cause closed” will mean “non-trivially common cause closed”.

Both common cause incomplete and common cause closed probability spaces exist ([9]); however, a general characterization of common cause (in)complete spaces was not given in [9]. Our aim in this section is to give such a characterization in terms of the atomic structure of the probability spaces. The first result is:

Proposition 3.2. *Let (\mathcal{S}, p) be a probability algebra with two distinct atoms. Then (\mathcal{S}, p) is not common cause closed.*

In the proof we make use of the following lemma:

Lemma 3.3 ([9], Lemma 1). *Let (\mathcal{S}, p) be arbitrary and assume that $a, b \in \mathcal{S}$ are (positively) correlated. Then c can not be a proper common cause of the correlation between a and b if any of the following conditions hold.*

- (i) $c \leq (a \wedge b)^\perp$;
- (ii) $a \wedge b = c \wedge b$ and $a \leq c$;
- (iii) $c = a \wedge b$;
- (iv) $c = a \vee b$;
- (v) $a \leq b \leq c$.

Proof: (of Proposition 3.2) Let $a, b \in \mathcal{S}$ be two atoms. We may assume that $p(a \vee b) \neq 1$ because otherwise we are dealing essentially with the case of the Boolean algebra generated by two atoms, which is not common cause closed. Thus a and $a \vee b$ are correlated because

$$p(a \wedge (a \vee b)) = p(a) > p(a)p(a \vee b).$$

and $p(a \vee b) < 1$. We claim that this correlated pair has no common cause in \mathcal{S} . By (i) of Lemma 3.3 (applying to the pair a and $a \vee b$) if c is a common cause then $c \not\leq a^\perp$. Hence $c > a$ ($c < a$ is impossible because a is an atom). Consequently c is of form $a \vee d$ for some $d \in \mathcal{S}$. Since b is an atom there are two cases depending on whether $d \wedge b = 0$ or $d > b$.

Case 1: Suppose $d \wedge b = 0$. Then $a \wedge (a \vee b) = c \wedge (a \vee b)$ which contradicts (ii) of Lemma 3.3.

Case 2: Suppose $d > b$. Then $a \leq a \vee b \leq c$ which contradicts (v) of Lemma 3.3. ■

Our next aim is to prove that any probability algebra having at most one atom is common cause closed (Proposition 3.7). In order to do this we make use of the following lemmas and facts. First, observe that if $x, y \in [0, 1]$ then the inequality

$$x - x \cdot y + y \leq 1 \tag{6}$$

holds. One can see this by rearranging terms as $(1 - x)(1 - y) \geq 0$.

Lemma 3.4 ([9], Lemma 4). *Let (\mathcal{S}, p) be a probability algebra and $a, b \in \mathcal{S}$ be two correlated events such that $a \not\leq b$ and $b \not\leq a$. Then $c \in \mathcal{S}$ is a common cause of the correlation between a and b if one of the following two conditions holds.*

1. $c \leq (a \wedge b)$ and

$$p(c) = \frac{p(a \wedge b) - p(a)p(b)}{1 + p(a \wedge b) - (p(a) + p(b))} \tag{7}$$

2. $(a \vee b) \leq c$ and

$$p(c) = \frac{p(a)p(b)}{p(a \wedge b)} \tag{8}$$

Recall that every purely non-atomic probability algebra (space) possesses the *denseness property* (see e.g. [10]), i.e if (\mathcal{S}, p) is purely non-atomic then the following is true for all $\gamma \in (0, 1)$ and $a \in \mathcal{S}$:

$$\text{if } 0 \leq \gamma \leq p(a) \text{ then there exists } b \leq a \text{ such that } p(b) = \gamma \tag{9}$$

Replacing a and b with their complements we get the following equivalent (dual) form of the denseness property:

$$\text{if } p(a) \leq \gamma \leq 1 \text{ then there exists } b \geq a \text{ such that } p(b) = \gamma \quad (10)$$

If the space is not purely non-atomic then denseness holds in a limited sense:

Lemma 3.5. *Let (\mathcal{S}, p) be a probability algebra with exactly one atom, and let*

$$p = \alpha p_1 + (1 - \alpha) p_2 \quad (11)$$

be the decomposition of p with p_1 and p_2 being mutually singular probability measures which are respectively purely non-atomic and purely atomic. Then the following denseness properties hold:

For $x \in \mathcal{S}$ and $\gamma \in (0, 1)$, if either (a) or (b) below are fulfilled then there exists $y \leq x$ with $p(y) = \gamma$.

$$\text{(a) } p_1(x) \neq 0, p_2(x) = 0 \text{ and } 0 \leq \gamma \leq p(x).$$

$$\text{(b) } p_1(x) \neq 0, p_2(x) \neq 0 \text{ and } 1 - \alpha \leq \gamma \leq p(x).$$

Dually, for $x \in \mathcal{S}$ and $\gamma \in (0, 1)$ if one of (a') or (b') below are fulfilled then there exists $y \geq x$ with $p(y) = \gamma$.

$$\text{(a')} \quad p_1(x) \neq 1, p_2(x) = 1 \text{ and } p(x) \leq \gamma \leq 1.$$

$$\text{(b')} \quad p_1(x) \neq 1, p_2(x) \neq 1 \text{ and } p(x) \leq \gamma \leq \alpha.$$

Proof: By replacing x , y and γ with x^\perp , y^\perp and $1 - \gamma$, it is easy to see that (a') and (b') are equivalent to (a) and (b). Therefore it is enough to prove (a) and (b).

Suppose $p_1(x) \neq 0$, $p_2(x) = 0$ and $0 \leq \gamma < p(x)$. Then by definition $0 \leq \gamma < \alpha p_1(x) + (1 - \alpha) p_2(x) = \alpha p_1(x)$. Hence $0 \leq \frac{\gamma}{\alpha} < p_1(x)$ and by purely non-atomicity of p_1 and the denseness property (9) we can conclude that there is some $y < x$ such that $p_1(y) = \frac{\gamma}{\alpha}$. Then $p(y) = \alpha p_1(y) + (1 - \alpha) p_2(y) = \gamma$, as desired.

To prove (b) suppose $p_1(x) \neq 0$, $p_2(x) \neq 0$ and $1 - \alpha \leq \gamma < p(x)$. Then

$$0 \leq \frac{\gamma - (1 - \alpha)}{\alpha} < p_1(x).$$

Again by (9), there exists $y_0 < x$ such that $p_1(y) = \frac{\gamma - (1 - \alpha)}{\alpha}$. If we denote the unique atom of (\mathcal{S}, p) by q (thus $p_1(q) = 0$ and $p_2(q) = 1$) then $y \doteq y_0 \vee q$ is as desired. ■

Lemma 3.6. *Let (\mathcal{S}, p) be a probability algebra and suppose $a, b \in \mathcal{S}$ are correlated events. Then the following inequalities hold.*

$$p(a \vee b) \neq 1 \quad (12)$$

$$p(a \wedge b) \neq 0 \quad (13)$$

$$p(a \vee b) \leq \frac{p(a)p(b)}{p(a \wedge b)} \quad (14)$$

$$p(a \wedge b) \geq \frac{p(a \wedge b) - p(a)p(b)}{1 + p(a \wedge b) - (p(a) + p(b))} \quad (15)$$

Proof: (13) follows directly from the definition of correlation. To see (14) note that the following inequality holds:

$$p(a \vee b) \leq p(a) + p(b) \leq p(a) \cdot p(b) \leq \frac{p(a) \cdot p(b)}{p(a \wedge b)}.$$

where $p(a) + p(b) \leq p(a) \cdot p(b)$ follows from (6). To see (12) suppose, seeking a contradiction, that $p(a \vee b) = 1$. Then by the general additivity rule

$$p(a) + p(b) = p(a \wedge b) + p(a \vee b) \quad (16)$$

and by the correlation inequality we get $p(a) + p(b) - 1 > p(a) \cdot p(b)$. But this is impossible, again by (6). Finally, (15) can be obtained after rearranging terms and using (12), (14) and (6). ■

Proposition 3.7. *Assume that the probability algebra (\mathcal{S}, p) has at most one atom. Then (\mathcal{S}, p) is common cause closed.*

Proof: Let $p = \alpha \cdot p_1 + (1 - \alpha) \cdot p_2$ ($\alpha \in [0, 1]$) be the decomposition of p into mutually singular probability measures p_1 and p_2 which are respectively purely non-atomic and purely atomic. Since there is exactly one atom in \mathcal{S} (denote it by q), for all $b \in \mathcal{S}$ we have $p_2(b) \neq 0$ if and only if $q \leq b$ and then $p_2(b) = p_2(q) = 1$.

Let now $a, b \in \mathcal{S}$ be two correlated events (clearly $p(a), p(b) \neq 0$). We need to find a common cause $c \in \mathcal{S}$ for this pair of events. We proceed by distinguishing cases. There are basically four cases determined by the relation between a and b and some cases can be dealt with quickly: $a \wedge b = 0$ is impossible by (13) and by symmetry we do not have to differentiate between $a \leq b$ and $b \leq a$.

Case 1: Suppose $a \not\leq b$ and $b \not\leq a$. There are two subcases:

- (i) If $p_2(a \vee b) = 1$ then $p_1(a \vee b) \neq 1$ since otherwise $p(a \vee b)$ would be equal to 1 which contradicts (12). Let

$$\gamma = \frac{p(a)p(b)}{p(a \wedge b)}.$$

Then by (14) we get

$$p(a \vee b) \leq \gamma \leq 1,$$

and by (a') of Lemma 3.5 it follows that there exists $c \geq (a \vee b)$ with $p(c) = \gamma$. By Lemma 3.4 (2), c is a common cause.

(ii) If $p_2(a \vee b) = 0$ then $p_2(a \wedge b) = 0$ as well and hence $p_1(a \wedge b) \neq 0$ since otherwise we would get $p(a \wedge b) = 0$ which is impossible by (13). Let

$$\gamma = \frac{p(a \wedge b) - p(a)p(b)}{1 + p(a \wedge b) - (p(a) + p(b))}.$$

Then by (15) we get

$$0 \leq \gamma \leq p(a \wedge b),$$

and by (a) of Lemma 3.5 it follows that there exists $c < (a \wedge b)$ with $p(c) = \gamma$. By (1) of Lemma 3.4, c is a common cause.

Case 2: Suppose $a \leq b$ and $p_1(b) = 0$. Then q obviously satisfies equations (2)–(5) hence it is a common cause.

Case 3: Assume $a < b$ and $p_1(b) \neq 0$. Then any c satisfying $a \leq c \leq b$ is a common cause. ■

Theorem 3.8. *The probability algebra (\mathcal{S}, p) is common cause closed if and only if there is at most one atom in it.*

Proof: Straightforward by combining Proposition 3.7 and Proposition 3.2. ■

4 Common cause closure

Given a common cause incomplete probability space, it is natural to ask if it can be common cause completed in the sense that it can be embedded into a larger probability space that is common cause closed. To make this question precise we need the following definitions of extension and embedding:

Definition 4.1. The probability space (X', \mathcal{S}', p') is called an *extension* of (X, \mathcal{S}, p) if there exists a Boolean algebra σ -embedding h of \mathcal{S} into \mathcal{S}' such that

$$p(x) = p'(h(x)) \quad \text{for all } x \in \mathcal{S} \tag{17}$$

If (X', \mathcal{S}', p') is an extension of (X, \mathcal{S}, p) then we say that (X, \mathcal{S}, p) can be *embedded* into (X', \mathcal{S}', p') .

This definition, and in particular condition (17), implies that if (X', \mathcal{S}', p') is an extension of (X, \mathcal{S}, p) (with respect to the embedding h), then every single correlation $Corr_p(a, b)$ in (X, \mathcal{S}, p) is carried over intact by h into the correlation $Corr_{p'}(h(a), h(b))$ in (X', \mathcal{S}', p') because

$$\begin{aligned} p'(h(a) \cap h(b)) &= p'(h(a \cap b)) \\ &= p(a \cap b) > p(a)p(b) = p'(h(a))p'(h(b)) \end{aligned}$$

Hence, it does make sense to ask whether a correlation in (X, \mathcal{S}, p) has a Reichenbachian common cause in the extension (X', \mathcal{S}', p') .

The main result in this section is the following

Theorem 4.2. *Every probability space has a purely non-atomic common cause closed extension. In particular, every probability space is common cause completable with respect to any set of correlated events.*

To prove Theorem 4.2 all we need to show is that for any probability space (X, \mathcal{S}, p) there exists an extension (X', \mathcal{S}', p') which is purely non-atomic, since by Theorem 3.8 such an extension is common cause closed. We will prove such a theorem (Theorem 6.6). The main idea of the proof of Theorem 6.6 is the following: Given any (possibly common cause incomplete) probability space $\mathfrak{A} = (X, \mathcal{A}, \mu)$, take another probability space $\mathfrak{B} = (Y, \mathcal{B}, \nu)$ and form some sort of a product $\mathfrak{A} \otimes \mathfrak{B}$ of \mathfrak{A} and \mathfrak{B} in such a way that

1. $\mathfrak{A} \otimes \mathfrak{B}$ is an extension of both \mathfrak{A} and \mathfrak{B} .
2. If \mathfrak{B} is purely non-atomic then $\mathfrak{A} \otimes \mathfrak{B}$ is also purely non-atomic.

Taking then the purely non-atomic space $([0, 1], \mathfrak{L}, \lambda)$ the product $(X, \mathcal{A}, \mu) \otimes ([0, 1], \mathfrak{L}, \lambda)$ is the required extension.

We will give such a \otimes product construction in the Appendix. The construction consists of the following main steps.

Step 1. Using \mathcal{A} and \mathcal{B} in (X, \mathcal{A}, μ) and (Y, \mathcal{B}, ν) we construct a Boolean algebra \mathcal{D} into which both \mathcal{A} and \mathcal{B} can be σ -embedded (Lemma 6.1).

Step 2. We construct a measure ρ on \mathcal{D} in such a way that (\mathcal{D}, ρ) becomes an extension of both (\mathcal{A}, μ) and (\mathcal{B}, ν) (Lemma 6.2 and Lemma 6.3).

Step 3. Since we would like the \otimes product to be a probability space, i.e. a space with an underlying set representing the elementary random events, we find set-representation

for the algebra \mathcal{D} in such a way that this representation preserves infinite operations as well.

Step 4. We extend our algebra to a σ -algebra and we also extend our measure to a probability measure on the σ -algebra. We will show that the purely non-atomic character of (Y, \mathcal{B}, ν) entails that this extension is also purely non-atomic.

The details of this construction are presented in the Appendix.

5 The Common Cause Principle in the light of common cause completability

While the definition of common cause was given first by Reichenbach [25], he did not formulate the Common Cause Principle explicitly, this was done later, mainly by Salmon [26, 27, 28], and in the current research on causal modeling the Principle appears in the form of the notion of Causal Markov condition [33], [18]. The Common Cause Principle has been extensively discussed and debated in the literature, especially by Arntzenius [1], Butterfield [2, 3]; Cartwright [4]; Chang and Cartwright [5]; Henson [11], Hoover [17], Placek [20, 21, 19]; Salmon [26, 27, 28]; Sober [29, 30, 31, 32]; Spohn [34]; Suppes [35]; Uffink [37] and Van Fraassen [38, 7]. There seems to be consensus that the Principle is not universally valid. There are two types of arguments in favor of the claim that the Principle does not hold in general:

1. Informal examples of simple, classical correlations between events that are, intuitively, causally independent, and for which, again, intuitively, there should not exist a common cause. Sober's example of correlation between bread prices in England and Venetian sea levels [31] is a typical example in this category (see also [32], section 3.8).
2. Arguments showing that the EPR correlations cannot be explained by common causes.

The common cause completability result in the previous section (Theorem 4.2) shows however that the Common Cause Principle can always be defended by referring to "hidden common causes": A falsifying attempt based on displaying a correlation in a common cause incomplete probability space can be refuted by saying that the common cause incomplete probability space in question is just "too meager" and there exists a larger space that contains a common cause of the correlation in question. This strategy of saving the Principle was available already on the basis of earlier result stating that every common cause incomplete probability space is "locally" (i.e. with respect to a given, finite set of correlations) common cause completable [14]; however, this defence had two weaknesses:

1. One could object that the extended probability space was not necessarily common cause closed and hence it may violate the Principle for correlations that are not in the given, fixed set. This counterargument against the natural “hidden common cause” defence of the Principle against falsification is now neutralized by Theorem 4.2, because any common cause incomplete probability space can be embedded into a common cause closed one. One also can reject the above objection against the defence of the Principle by arguing that from the perspective of the Common Cause Principle what matters is only whether any particular correlation or *any* specific set of correlations (the ones we are interested in) have a satisfactory explanation in terms of common causes and that therefore a defence based on local common cause completability is sufficient, the status of the Principle does not hinge on whether or not common cause closed probability spaces can be achieved. On the basis of this reasoning Theorem 4.2 is an overkill. But an overkill is still a kill, and the claim that any probability space containing a correlation without a common cause in it can be extended into a common cause closed one does represent a defence of the Principle; this defence is just logically much stronger than the protection based on earlier, local extendability results.

2. The other weakness is that, since earlier common cause completability results were restricted to *finite* sets of correlations, those results did not allow one to conclude in full generality that *any* specific set of correlations we might be interested in can possibly have a common cause explanation – the strong common cause extendability result (Theorem 4.2) entitles us to make such a general claim, no matter what the cardinality of the set of correlations is for which we wish to find common causes. This means that Theorem 4.2 adds extra strength to the kind of defence of the Principle that was already available on the basis of previous local extendability results. This extra strength also can in principle turn out to be an overkill if one can prove that given *any* probability space with *any* set of correlations in it there exists an extension of the space that is not necessarily common cause closed but contains a common cause of every correlation in the given set. No result of this type is known at present however – the only known way of extending a probability space with respect to a set of correlations of arbitrary cardinality is the one ensured by Theorem 4.2

We wish to emphasize however that we do not interpret the strong common cause completability result (Theorem 4.2) as a proof that the Common Cause Principle holds; in our view, the theorem’s main message is that in order to falsify the Principle rigorously it is imperative that one requires some conditions that the common causes of correlations have to satisfy *in addition* to the ones in Reichenbach’s definition of common cause – without such

additional conditions the Principle is not falsifiable.

Those additional conditions can in principle be of two sorts:

Probabilistic: Formulated in terms of probabilities of the correlated events a, b and their common causes c – and possibly involving probabilities of additional events x beyond a, b and c .

Algebraic: Formulated in terms of the Boolean algebra operations between a, b, c – and possibly involving additional events x beyond a, b and c .

The additional conditions fix a certain probabilistic or algebraic *type* of the common cause, and it is a highly non-trivial problem whether a common cause incomplete probability space is common cause completable in such a way that the extension of the space contains common causes of a *certain fixed type* of the correlations in the common cause incomplete space. A technically more explicit analysis of the (lack of) common cause of the correlation between bread prices in Britain and Venetian sea levels would thus require the specification of the particular type of common cause that is claimed to be absent in this case. Whether a common cause closed space containing the required type of common cause exist remains an open question since Theorem 4.2 only shows that there exists a *particular type* of common cause in the common cause closed space, which might or might not be the type suiting the example. (One also can specify and analyze Sober’s example in terms of the theory of time series, as Hoover does [17]. Hoover’s conclusion is that the example is not a counterexample to the Principle.) It is to be expected that a purely non-atomic probability space, into which any probability space can be embedded by Theorem 6.6, will contain various types of common causes since such a space is extremely rich in events (cf. the denseness property), but one cannot expect even such probability spaces to be unconditionally common cause closed, i.e. common cause closed no matter what type of common causes one wishes to have. One can show for instance that one cannot in general require a common cause of a correlation to be a common cause of *other correlations* as well (“common causes are not *common* common causes” – see [16]) because the probabilistic constraints entailed by the assumption that common causes are *common* common causes are not satisfiable in general; hence common cause extensions of “*common* common cause” type do not always exist.

This is also the case with the EPR correlations: here one has more than one (typically three or four) correlations and in general they do not have a *common* common cause. But even if one does not require that the different correlations appearing in an EPR correlation experiment have a common common cause, one can impose additional probabilistic conditions on the common cause models of these correlations: in the EPR situation one has information about the spatio-temporal (hence causal) relation of the events involved and one can utilize

this information to motivate further probabilistic constraints on the correlated events and on the events in the hypothetical common causes. There are two sorts of extra requirements: “locality” and “no-conspiracy”.

“Locality” is the expression in terms of probabilistic independence conditions of the fact that the outcome of the measurement in one wing of the measurement apparatus in the experimental setup measuring EPR correlations cannot be causally influenced by the choice of the measurement direction in the opposite wing of the measurement setup since the acts of measurements are causally independent (they are spacelike separated).

The so-called “No-conspiracy” condition is intended to express that there is no direct causal dependence between the measurement choices and the hypothetical common causes. Whether the locality and no-conspiracy conditions are intuitively acceptable can be debated but it can be shown that local and non-conspiratorial common cause explanation of the EPR correlations are not possible. (See [36], [8], [22], [3], [13], [12] for proofs and further discussion). The significance of the strong common cause completability result in this paper (Theorem 4.2) from the perspective of the EPR correlations is that it is a “No-go” theorem on “No-go” theorems: it shows that the further conditions of locality and no-conspiracy are absolutely indispensable in the proofs of “No-go” propositions stating the impossibility of common cause explanations of EPR correlations: without these additional requirements even any countably infinite subset of all the correlations inherent in an entangled quantum state are in principle explainable by referring to hidden causes.

Note finally that one can raise the problem of common cause completability and common cause closedness of non-classical probability spaces [15], [23], [24] where the Boolean algebra is replaced by a more general orthocomplemented lattice and the classical probability measure by a general countably additive bounded measure. It would be interesting to see if the product construction described in this paper for classical probability spaces is extendible to general probability spaces and whether common cause closedness of non-classical probability spaces can be established along the lines presented in this paper.

6 Appendix

In this section we carry out the program described in Section 4

Step 1.

Let $\mathfrak{A} = (X, \mathcal{A}, \mu)$ and $\mathfrak{B} = (Y, \mathcal{B}, \nu)$ be two probability spaces. Let \mathcal{C} be the set of all functions mapping Y to \mathcal{A} :

$$\mathcal{C} \doteq \{f : f : Y \rightarrow \mathcal{A} \text{ is a function}\}$$

We make \mathcal{C} a Boolean-algebra in the following way. For $f, g \in \mathcal{C}$ let

$$f \vee g, f \wedge g, f^\perp : Y \rightarrow \mathcal{A}$$

be the functions defined for $y \in Y$ by

$$\begin{aligned} (f \wedge g)(y) &\doteq f(y) \wedge_{\mathcal{A}} g(y) \\ (f \vee g)(y) &\doteq f(y) \vee_{\mathcal{A}} g(y) \\ f^\perp(y) &\doteq f(y)^{\perp_{\mathcal{A}}} \end{aligned}$$

where $\wedge_{\mathcal{A}}, \vee_{\mathcal{A}}, \perp_{\mathcal{A}}$ are the Boolean operations in the algebra \mathcal{A} . Let the elements $\mathbb{0}, \mathbb{1}$ in \mathcal{C} be defined to be the following constant functions:

$$\mathbb{0}(y) \doteq 0_{\mathcal{A}} \quad \text{and} \quad \mathbb{1}(y) \doteq 1_{\mathcal{A}} \quad \forall y \in Y$$

It is routine to verify that \mathcal{C} is a Boolean-algebra with zero and unit elements $\mathbb{0}$ and $\mathbb{1}$. Furthermore, if \mathcal{A} is a σ -algebra, then so is \mathcal{C} : Given a countable set of functions $f_n \in \mathcal{C}$, the functions $\wedge f_i$ and $\vee f_i$ defined by

$$(\wedge f_i)(y) \doteq \wedge_{\mathcal{A}} f_i(y) \tag{18}$$

$$(\vee f_i)(y) \doteq \vee_{\mathcal{A}} f_i(y) \tag{19}$$

are well defined and belong to \mathcal{C} since the right hand sides of (18)-(19) exist by the σ property of \mathcal{A} .

We define two classes of special elements of \mathcal{C} : characteristic functions and simple functions. For $a \in \mathcal{A}$ and $B \subseteq Y$ write

$$\chi_B^a(y) \doteq \begin{cases} a & \text{if } y \in B \\ 0_{\mathcal{A}} & \text{otherwise} \end{cases}$$

Then χ_B^a is the a -valued characteristic function of the subset B . A function $s : Y \rightarrow \mathcal{A}$ is called *simple* if it is of the following form for some natural number n :

$$\bigvee_{i=0}^{n-1} \chi_{B_i}^{a_i}$$

where for all $i < n$ we have $B_i \in \mathcal{B}$ and $Y = \bigsqcup_{i < n} B_i$ (i.e. B_i 's are pairwise disjoint and cover Y), and $a_i \in \mathcal{A}$. Let \mathcal{D} denote the set of simple functions:

$$\mathcal{D} \doteq \left\{ \bigvee_{i=0}^{n-1} \chi_{B_i}^{a_i} : Y = \bigsqcup_{i < n} B_i, B_i \in \mathcal{B}, a_i \in \mathcal{A}, n \in \mathbb{N} \right\}.$$

Lemma 6.1. \mathcal{D} is a Boolean-algebra (a subalgebra of \mathcal{C}) and both \mathcal{A} and \mathcal{B} can be σ -embedded into \mathcal{D} .

Proof: Clearly, \mathbb{O} and \mathbb{I} are simple functions. If $s = \bigvee_{i=0}^{n-1} \chi_{B_i}^{a_i}$ is simple, then $s^\perp = \bigvee_{i=0}^{n-1} \chi_{B_i}^{a_i^\perp}$ is also simple. It is also not hard to see that join and meet of (finitely many) simple functions remain simple. Therefore \mathcal{D} is a Boolean-algebra although \mathcal{D} is not necessarily a σ -algebra in general; however, some infinite joins and meets may exist. Indeed, we show that both \mathcal{A} and \mathcal{B} can be embedded into \mathcal{D} : Let $h_{\mathcal{A}} : \mathcal{A} \rightarrow \mathcal{D}$ and $h_{\mathcal{B}} : \mathcal{B} \rightarrow \mathcal{D}$ be defined by

$$h_{\mathcal{A}}(a) \doteq \chi_Y^a \tag{20}$$

$$h_{\mathcal{B}}(B) \doteq \chi_B^1 \tag{21}$$

Then $h_{\mathcal{A}}$ and $h_{\mathcal{B}}$ are Boolean-algebra σ -homomorphisms obviously. ■

Step 2.

Next, we define a probability measure ρ on \mathcal{D} in such a way that (\mathcal{D}, ρ) becomes an extension of both (\mathcal{A}, μ) and (\mathcal{B}, ν) .

Let us call a function $f : Y \rightarrow \mathcal{A}$ ν -integrable if $\mu \circ f : Y \rightarrow [0, 1]$ is (\mathcal{B}, ν) -measurable. Then, because $\mu \circ f$ is non-negative, the integral $\int_Y \mu \circ f \, d\nu$ exists. For ν -integrable functions $f \in \mathcal{C}$ we define

$$\rho(f) \doteq \int_Y \mu \circ f \, d\nu.$$

Every simple function s in \mathcal{C} is ν -integrable, hence $\rho(s)$ is defined for every simple function s . The following lemma states that ρ is σ -additive on the set of ν -integrable functions (which is a broader class than simple functions, but in general the ν -integrable functions do not form a Boolean-algebra).

Lemma 6.2. Suppose that f_0, f_1, \dots are pairwise disjoint, ν -integrable functions. Then

$$\rho\left(\bigvee_i f_i\right) = \sum_i \rho(f_i).$$

Specifically, ρ is a σ additive probability measure on \mathcal{D} .

Proof: Pick $y \in Y$. Then $f_0(y), f_1(y), \dots$ are pairwise disjoint, hence $\mu(\bigvee_i f_i(y)) = \sum_i \mu(f_i(y))$. It follows that $\mu \circ (\bigvee_i f_i) = \sum_i \mu \circ f_i$. Then by definition

$$\rho\left(\bigvee_i f_i\right) = \int \mu \circ \left(\bigvee_i f_i\right) = \int \sum_i \mu \circ f_i \stackrel{!}{=} \sum_i \int \mu \circ f_i = \sum_i \rho(f_i),$$

where the equality marked with ! is a consequence of the monotone convergence theorem. $\rho(\mathbb{I}) = 1$ and $\rho(\mathbb{O}) = 0$, so ρ is normalized. ■

Lemma 6.3. (\mathcal{D}, ρ) is an extension of both (\mathcal{A}, μ) and (\mathcal{B}, ν) .

Proof: We need to show that ρ extends $\mu \circ h_{\mathcal{A}}$ and $\nu \circ h_{\mathcal{B}}$, where $h_{\mathcal{A}}$ and $h_{\mathcal{B}}$ are defined by (20) and (21), respectively:

$$\rho(h_{\mathcal{A}}(a)) = \int_Y \mu \circ \chi_Y^a d\nu = \int_Y \mu(a) d\nu = \mu(a) \cdot \int_Y 1 d\nu = \mu(a)$$

$$\rho(h_{\mathcal{B}}(B)) = \int_Y \mu \circ \chi_B^1 d\nu = \int_Y 1 d\nu = \nu(B)$$

■

We show next that (\mathcal{D}, ρ) inherits the purely non-atomic structure of (Y, \mathcal{B}, ν) :

Lemma 6.4. If (Y, \mathcal{B}, ν) is purely non-atomic then (\mathcal{D}, ρ) is also purely non-atomic.

Proof: We have to prove that for all $f \in \mathcal{D}$ with $\rho(f) > 0$ there exists $f' \in \mathcal{D}$ with $f' < f$ and $0 < \rho(f') < \rho(f)$. Suppose $\rho(f) > 0$ for some $f \in \mathcal{D}$. Because f is a simple function, there exists a measurable subset A of Y (i.e. $A \in \mathcal{B}$) with $\nu(A) > 0$ and $\rho(f \upharpoonright A) > 0$. Hence $f \upharpoonright A > \mathbb{0} \upharpoonright A$. By non-atomicity of ν , there is a smaller set $B \subset A$ with $0 < \nu(B) < \nu(A)$. Let

$$f'(y) = \begin{cases} 0 & \text{if } y \in A \setminus B \\ f(y) & \text{otherwise} \end{cases}$$

Then $0 < \rho(f') < \rho(f)$ as desired. ■

Step 3.

In this step we give a set-representation of \mathcal{C} which gives rise to a set-representation of \mathcal{D} . By Stone's representation theorem every Boolean algebra is isomorphic to a set-algebra, but this is not sufficient for our purposes because it may happen that the isomorphism coming from the Stone-representation does not preserve infinite operations, which we must preserve. So we give a set-representation of \mathcal{C} in a way that the representation preserves countable joins and meets.

We define first the set $X \otimes Y$ by

$$X \otimes Y \doteq \bigcup_{y \in Y} (X \times \{y\}) \quad (22)$$

Set theoretically, $X \otimes Y$ is just the Descartes product $X \times Y$ but writing it in the manner of eq. (22) will be useful below. Now for any $E \subseteq X \otimes Y$ and for any fixed $y \in Y$ let

$$E_y = \{x : (x, y) \in E\}$$

Defining Σ by

$$\Sigma \doteq \{E \subseteq X \otimes Y : E_y \in \mathcal{A} \text{ for all } y \in Y\}$$

we have $\Sigma \subseteq \mathcal{P}(X \otimes Y)$, and the fact that \mathcal{A} is a Boolean σ algebra entails that Σ is a Boolean σ -algebra as well. We claim that \mathcal{C} can be identified with Σ . We give the σ -isomorphism $\iota : \mathcal{C} \rightarrow \Sigma$. For $f \in \mathcal{C}$ write

$$\iota(f) \doteq \bigcup_{y \in Y} (f(y) \times \{y\}).$$

Note that $f(y) \subseteq X$, $f(y) \in \mathcal{A}$ for all $y \in Y$, hence $\iota(f) \in \Sigma$. Conversely, if $E \in \Sigma$ then $E_y \in \mathcal{A}$; hence the function $f(y) = E_y$ is mapped to E by ι . Checking that ι is a σ -isomorphism is routine and omitted.

Now we find the image of \mathcal{D} as

$$\mathfrak{D} \doteq \iota[\mathcal{D}] = \{\iota(f) : f \in \mathcal{D}\}.$$

Then ι is a σ -isomorphism between \mathcal{D} and \mathfrak{D} . Further let

$$\tau \doteq \rho \circ \iota^{-1}.$$

Then τ is a measure on \mathfrak{D} , moreover the isomorphism ι between the two spaces (\mathcal{D}, ρ) and (\mathfrak{D}, τ) preserves the measures. It follows that (\mathfrak{D}, τ) is an extension of both (X, \mathcal{A}, μ) and (Y, \mathcal{B}, ν) .

To summarize what we have obtained so far: Given two probability spaces $\mathfrak{A} = (X, \mathcal{A}, \mu)$ and $\mathfrak{B} = (Y, \mathcal{B}, \nu)$ we constructed an algebra \mathfrak{D} of subsets of $X \otimes Y$ with a probability measure τ on it with the property that both \mathcal{A} and \mathcal{B} can be embedded into \mathfrak{D} by a σ -homomorphism $\iota \circ h$ and τ being an extension of both $\mu \circ \iota \circ h$ and $\nu \circ \iota \circ h$.

Step 4.

It is a well-known fact from measure theory that if \mathcal{S} is an algebra of subsets of a set and p is a σ -finite measure on it, then p can be uniquely extended to a σ -finite measure \bar{p} on the σ algebra $\bar{\mathcal{S}}$ generated by \mathcal{S} in such a way that $\bar{p} \upharpoonright \mathcal{S} = p$ (see e.g. [10, Theorem A, Section 13]).

In our situation it follows that there exists a σ -algebra $\mathcal{A} \otimes \mathcal{B}$ of subsets of $X \otimes Y$ and a probability measure $\mu \otimes \nu$ on it such that \mathfrak{D} is a subalgebra of $\mathcal{A} \otimes \mathcal{B}$ and $\mu \otimes \nu \upharpoonright \mathfrak{D} = \tau$. The only thing we have to prove is that the space $(X \otimes Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$ inherits the purely non-atomic structure of (Y, \mathcal{B}, ν) . This we prove in the next lemma.

Lemma 6.5. *Let \mathcal{S} be an algebra of subsets of a set and let p be a probability measure on \mathcal{S} such that the space (\mathcal{S}, p) is purely non-atomic. Let \bar{p} be the extension of p to the generated σ -algebra $\bar{\mathcal{S}}$. Then $(\bar{\mathcal{S}}, \bar{p})$ is also purely non-atomic.*

Proof: To simplify notation, since $\bar{p} \upharpoonright \mathcal{S} = p$, from now on we write p instead of \bar{p} . What we have to prove is that for any $x \in \bar{\mathcal{S}}$ with $p(x) > 0$ there exists $y \in \bar{\mathcal{S}}$ such that $y \subseteq x$ and

$$0 < p(y) < p(x) \quad (23)$$

The idea of the proof is the following: By a standard result of measure theory related to the extension of measures from algebras to σ algebras [10, Theorem D, Section 13], for any $x \in \bar{\mathcal{S}}$ and any $\varepsilon > 0$ there is an element a in the original algebra \mathcal{S} such that the symmetric difference of x and a is smaller than ε . The purely non-atomic property of (\mathcal{S}, p) entails that for any real numbers r and s such that $0 < r < s < p(a)$ we can find an element $a' \subseteq a$ with $r < p(a') < s$. We will show that r, s and ε can be chosen in such a way that the probability of $x \cap a'$ is both non-zero and smaller than $p(x)$; so $y = x \cap a'$ is the element satisfying (23). Let's see this in detail.

Let $x \in \bar{\mathcal{S}}$ be any element of positive measure and let $\varepsilon > 0$ be small. By [10, Theorem D, Section 13] there exists an $a \in \mathcal{S}$ with the property $p(x \Delta a) < \varepsilon$, where $x \Delta a = (x \setminus a) \cup (a \setminus x)$ is the symmetric difference of sets x and a . The following inequalities hold:

$$p(x) + \varepsilon > p(a) \quad (24)$$

$$p(a) > p(x) - \varepsilon \quad (25)$$

they are immediate consequences of $a \cup (x \Delta a) = a \cup x = x \cup (x \Delta a)$ and of the general additivity rule for p (eq. (16)). Since (\mathcal{S}, p) is purely non-atomic, by the denseness property it follows that for every $a \in \mathcal{S}$ with $p(a) > 0$ there is $a' \in \mathcal{S}$ such that $a' \subseteq a$ and

$$0 < \alpha < p(a') < \beta < p(a), \quad (26)$$

where $0 < \alpha < \beta < p(a)$ are arbitrary real numbers. Let $\alpha = \frac{p(a)}{2}$ and let $\beta > \alpha$ (the value of β will be given later). By (26), we can choose an element $a' \subseteq a$, $a' \in \mathcal{S}$ such that $0 < \alpha < p(a') < \beta < p(a)$. So we have

$$p(a') \geq \frac{p(a)}{2} \quad (27)$$

Our aim is to find a suitable ε to be able to prove $0 < p(x \cap a') < p(x)$. If we find such ε then taking $y = x \cap a'$ the proof is completed.

First we find conditions on ε to ensure $p(a' \cap x) < p(x)$. Clearly, it is enough to have $p(a') < p(x)$. This we can achieve if we can choose β to be smaller than $p(x)$. Since $\beta > \alpha = \frac{p(a)}{2}$ to find such β it is enough to have $\frac{p(a)}{2} < p(x)$. By (24), $p(a) < p(x) + \varepsilon$ consequently if $p(x) + \varepsilon < 2 \cdot p(x)$ then we are done. So we need the following condition on ε to hold:

$$\varepsilon < p(x) \quad (28)$$

Now we are looking for a condition on ε to ensure $0 < p(x \cap a')$. We proceed as follows. Since $p(a') = p(a' \setminus x) + p(a' \cap x)$ and $a' \setminus x \subseteq a \Delta x$ the inequality $p(a' \setminus x) < \varepsilon$ holds. Putting this together with (27) and (25) we obtain

$$p(a' \cap x) \geq \frac{p(a)}{2} - \varepsilon \geq \frac{p(x) - \varepsilon}{2} - \varepsilon \quad (29)$$

Consequently if

$$\varepsilon < \frac{p(x)}{3} \quad (30)$$

then $p(a' \cap x) > 0$.

To summarize if both (28) and (30) hold then $0 < p(x \cap a') < p(x)$ is true. It is clear that for all $x \in \overline{\mathcal{S}}$ with $0 < p(x)$ we can find such ε . ■

We summarize the results:

Theorem 6.6. *Let $\mathfrak{A} = (x, \mathcal{A}, \mu)$ and $\mathfrak{B} = (y, \mathcal{B}, \nu)$ be two probability spaces. Then $\mathfrak{A} \otimes \mathfrak{B} = (X \otimes Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$ defined above is a probability space extending both \mathfrak{A} and \mathfrak{B} . If \mathfrak{B} is purely non-atomic then $\mathfrak{A} \otimes \mathfrak{B}$ is also purely non-atomic.* ■

A natural question that arises at this point is whether for any probability space an extension exists that is common cause closed but is *not* purely non-atomic but has one single atom. Concerning this question we state the following open problem.

Problem 6.7. Is it true that every probability space can be embedded into another probability space which has exactly one atom?

Our conjecture is that the answer is “yes” for all not purely non-atomic spaces. We also note that it would be enough to solve the problem for purely non-atomic spaces, in particular, for (\mathcal{L}, λ) .

References

- [1] F. Arntzenius. The common cause principle. *PSA 1992*, 2:227–237, 1993.
- [2] J. Butterfield. A space-time approach to the Bell inequality. In J. Cushing and E. McMullin, editors, *Philosophical Consequences of Quantum Theory*, pages 114–144. University of Notre Dame Press, Notre Dame, 1989.
- [3] J. Butterfield. Stochastic Einstein Locality revisited. *The British Journal for the Philosophy of Science*, 58:805–867, 2007.

- [4] N. Cartwright. How to tell a common cause: Generalization of the conjunctive fork criterion. In J. H. Fetzer, editor, *Probability and Causality*, pages 181–188. Reidel Pub. Co., 1987.
- [5] H. Chang and N. Cartwright. Causality and realism in the EPR experiment. *Erkenntnis*, 38:269–290, 1993.
- [6] J. Cushing and E. McMullin, editors. *Philosophical Consequences of Quantum Theory*, Notre Dame, 1989. University of Notre Dame Press.
- [7] B.C. Van Fraassen. Rational belief and the common cause principle. In R. McLaughlin, editor, *What? Where? When? Why?*, pages 193–209. D. Reidel Pub. Co., 1982.
- [8] G.S. Grasshoff, S. Portmann, and A. Wüthrich. Minimal assumption derivation of a Bell-type inequality. *The British Journal for the Philosophy of Science*, 56:663–680, 2005.
- [9] B. Gyenis and M. Rédei. When can statistical theories be causally closed? *Foundations of Physics*, 34:1285–1303, 2004.
- [10] P. Halmos. *Measure Theory*. D. Van Nostrand, New York, 1950.
- [11] J. Henson. Comparing causality principles. *Studies in the History and Philosophy of Modern Physics*, pages 519–543, 2005.
- [12] G. Hofer-Szabó. Separate common cause systems and Bell(δ) inequalities. *Submitted*.
- [13] G. Hofer-Szabó. Separate- versus *common*-common-cause-type derivations of the Bell inequalities. *Synthese*, 163:199–215, 2008.
- [14] G. Hofer-Szabó, M. Rédei, and L.E. Szabó. On Reichenbach’s Common Cause Principle and Reichenbach’s notion of common cause. *The British Journal for the Philosophy of Science*, 50:377–398, 1999.
- [15] G. Hofer-Szabó, M. Rédei, and L.E. Szabó. Common cause completability of classical and quantum probability spaces. *International Journal of Theoretical Physics*, 39:913–919, 2000.
- [16] G. Hofer-Szabó, M. Rédei, and L.E. Szabó. Common-causes are not common common-causes. *Philosophy of Science*, 69(20):623–636, 2002.
- [17] Kevin D. Hoover. Non-stationary time series, cointegration and the Principle of the Common Cause. *The British Journal for the Philosophy of Science*, 54:527–551, 2003.

- [18] J. Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, New York, 2000.
- [19] T. Placek. *Is Nature Deterministic?* Jagellonian University Press, Cracow, 2000.
- [20] T. Placek. Outcomes in branching space-time and GHZ-Bell theorems. *The British Journal for the Philosophy of Science*, 50:349–375, 2000.
- [21] T. Placek. Stochastic outcomes in branching space-time. An analysis of Bell theorems. *The British Journal for the Philosophy of Science*, 51:445–475, 2000.
- [22] S. Portmann and A. Wüthrich. Minimal assumption derivation of a weak Clauser–Horne inequality. *Studies in History and Philosophy of Modern Physics*, 38:844–862, 2007. electronic preprint: <http://www.arxiv.org/quant-ph/0604216>.
- [23] M. Rédei and S.J. Summers. Local primitive causality and the common cause principle in quantum field theory. *Foundations of Physics*, 32:335–355, 2002.
- [24] M. Rédei and S.J. Summers. Remarks on causality in relativistic quantum field theory. *International Journal of Theoretical Physics*, 46:2053–2062, 2007.
- [25] H. Reichenbach. *The Direction of Time*. University of California Press, Los Angeles, 1956.
- [26] W.C. Salmon. Why ask “why?”? In *Proceedings and Addresses of the American Philosophical Association*, volume 51, pages 683–705, 1978.
- [27] W.C. Salmon. Probabilistic causality. *Pacific Philosophical Quarterly*, 61:50–74, 1980.
- [28] W.C. Salmon. *Scientific Explanation and the Causal Structure of the World*. Princeton University Press, Princeton, 1984.
- [29] E. Sober. Common cause explanation. *Philosophy of Science*, 51:212–241, 1984.
- [30] E. Sober. The principle of the common cause. In J.H. Fetzer, editor, *Probability and Causality*, pages 211–228. Reidel Pub. Co., Boston, 1988.
- [31] E. Sober. Venetian sea levels, British bread prices, and the principle of common cause. *The British Journal for the Philosophy of Science*, 52:331–346, 2001.
- [32] E. Sober. *Evidence and Evolution*. Cambridge University Press, Cambridge, 2008.

- [33] P. Spirtes, C. Glymour, and P. Scheines. *Causation, Prediction and Search*. Lecture Notes in Statistics. Springer, New York, 1993. Second edition: MIT Press, Cambridge, MA, 2001.
- [34] W. Spohn. On Reichenbach's Principle of the Common Cause. In W. Salmon and G. Wolters, editors, *Logic, Language and the Structure of Scientific Theories*. University of Pittsburgh Press, Pittsburgh, 1991.
- [35] P. Suppes. *A Probabilistic Theory of Causality*. North-Holland, Amsterdam, 1970.
- [36] L.E. Szabó. Attempt to resolve the EPR-Bell paradox via Reichenbach's concept of common cause. *International Journal of Theoretical Physics*, 39:901–911, 2000.
- [37] J. Uffink. The principle of the common cause faces the Bernstein paradox. *Philosophy of Science, Supplement*, 66:512–525, 1999.
- [38] B.C. van Fraassen. The Charybdis of realism: Epistemological implications of Bell's inequalities. *Synthese*, 52:97–113, 1982. reprinted in [6] pages 97-113.