

A New Notion of Causal Closedness

Leszek Wroński · Michał Marczyk

Received: 3 February 2013 / Accepted: 3 February 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract In recent years part of the literature on probabilistic causality concerned notions stemming from Reichenbach's idea of explaining correlations between not directly causally related events by referring to their common causes. A few related notions have been introduced, e.g. that of a "common cause system" (Hofer-Szabó and Rédei in *Int J Theor Phys* 43(7/8):1819–1826, 2004) and "causal (N-)closedness" of probability spaces (Gyenis and Rédei in *Found Phys* 34(9):1284–1303, 2004; Hofer-Szabó and Rédei in *Found Phys* 36(5):745–756, 2006). In this paper we introduce a new and natural notion similar to causal closedness and prove a number of theorems which can be seen as extensions of earlier results from the literature. Most notably we prove that a finite probability space is causally closed in our sense iff its measure is uniform. We also present a generalisation of this result to a class of non-classical probability spaces.

1 Introduction

The so-called Principle of the Common Cause is usually taken to say that any surprising correlation between two factors which are believed not to directly influence one another is due to their (possibly hidden) common cause. The original version of the Principle as introduced by Hans Reichenbach in his book *The Direction of time* (1956) includes precise mathematical conditions connected to the notion (see definition 4 below) and became a hot topic for philosophers of science in the last decades of the previous century, after van Fraassen (1982) had linked it with the issues regarding causality in the context of EPR correlations. The Principle was widely criticised (see e.g. Arntzenius 1992 for a collection of its difficulties), but in recent years a number of researchers explored various mathematical questions

L. Wroński (✉) · M. Marczyk
Department of Philosophy, Jagiellonian University, Grodzka 52, 31-044 Kraków, Poland
e-mail: leszek.wronski@uj.edu.pl

regarding it, in at least one case leading even to the statement that the principle is “unfalsifiable” (Hofer-Szabó et al. 2000). This paper contributes to the discussion about the mathematical notions relevant to the Reichenbachian approach to explaining correlations. We prove a number of results concerning the (types of) probability spaces in which one can find Reichenbach-style explanations for correlations between events given an independence relation.

Suppose a probability space contains a correlation between two events we believe to be causally independent. Does the space contain a common cause for the correlation? If not, can the probability space be extended to contain such a cause but ‘preserving’ the old measure? This question has been asked and answered in the positive in Hofer-Szabó et al. (1999), where the notion of *common cause completability* was introduced: speaking a bit informally, a probability space S is said to be common cause completable with respect to a set A of pairs of correlated events iff there exists an extension of the space containing statistical common causes of all the correlated pairs in A . Gyenis and Rédei (2004) introduced the notion of *common cause closedness*, which (in our slightly different terminology) is equivalent to the following: a probability space S is common cause closed (or “causally closed”) with respect to a relation of independence $R_{ind} \subseteq S^2$ iff it contains statistical common causes (see definition 4 below) for all pairs of correlated events belonging to R_{ind} . The authors have proven therein that a finite classical probability space with no atoms of probability 0 is non-trivially common cause closed w.r.t. the relation of logical independence iff it is the space consisting of a Boolean algebra with 5 atoms and the uniform probability measure.¹ In other words, finite classical probability spaces (big enough to contain correlations between logically independent events) are in general *not* common cause closed w.r.t. the relation of logical independence, i.e. they contain a correlation between logically independent events for which no statistical common cause in the space exists; the only exception to this rule is the space with precisely 5 atoms of probability $\frac{1}{5}$ each. More spaces are common cause closed w.r.t. a more stringent relation of logical independence modulo measure zero event (“ L_{ind}^+ ”, see definition 6 below): they are the spaces with 5 atoms of probability $\frac{1}{5}$ each and any number of atoms of probability 0.

Still, a (statistical) common cause is not the only entity which could be used as an explanation for a correlation. Hofer-Szabó and Rédei (2004) generalized the idea of a statistical common cause, arriving at *statistical common cause systems* (“SCCSs”; see definition 5 below). SCCSs may have any countable size greater than 1;² the special case of size 2 reduces to the usual notion of common cause.

It was natural for corresponding notions of causal closedness to be introduced; a probability space is said to be *causally n -closed*³ w.r.t. a relation of independence R_{ind} iff it contains an SCCS of size n for any correlation between A, B such that

¹ The phrasing of the paper was in fact stronger, omitting the assumption about non-0 probabilities on the atoms (due to a missed special sub-case in the proof of case 3 of proposition 4 on p. 1299). The issue is connected to the distinction between proper and improper common causes and is discussed below in Sect. 4.

² See Wroński and Marczyk (2010) and Hofer-Szabó and Rédei (2006).

³ The notion was introduced in Hofer-Szabó and Rédei (2006).

$\langle A, B \rangle \in R_{ind}$. It is one of the results of the present paper that with the exception of the 5-atom uniform distribution probability space, no finite probability spaces without 0 probability atoms are causally n -closed w.r.t. the relation of logical independence, for any $n \geq 2$. Similarly, with the exception of the spaces with 5 atoms of probability $\frac{1}{5}$ each and any number of atoms of probability 0, no finite probability spaces with 0 probability atoms are causally n -closed w.r.t. L_{ind}^+ , for any $n \geq 2$.

We are interested in a slightly different version of causal closedness. If the overarching goal is to find explanations for correlations, why should we expect all explanations to be SCCSs of the same size? Perhaps some correlations are explained by common causes and others by SCCSs of a bigger size. We propose to explore the idea of *causal up-to- n -closedness*—a probability space is causally up-to- n -closed w.r.t. a relation of independence R_{ind} iff it contains an SCCS of size at most n for any correlation between events A, B such that $\langle A, B \rangle \in R_{ind}$.

It turns out that, in the class of finite classical probability spaces with no atoms of probability 0, just as the space with 5 atoms and the uniform measure is unique with regard to common cause closedness, the whole class of spaces with uniform distribution is special with regard to causal up-to-3-closedness—see theorem 2: a finite classical probability space with no atoms of probability 0 is causally up-to-3-closed w.r.t. the relation of logical independence iff it has the uniform distribution. We provide a method of constructing a statistical common cause or an SCCS of size 3 for any correlation between logically independent events in any finite classical probability space with the uniform distribution.

We require (following Gyenis and Rédei) of a causally closed probability space that all correlations be explained by means of proper—that is, differing from both correlated events by a non-zero measure event—statistical common causes. This has the consequence that a space causally closed w.r.t. the relation of logical independence can be transformed into a space which is not causally closed w.r.t. this relation just by adding a 0-probability atom. Perhaps, to avoid this unfortunate consequence, the notion of logical independence modulo measure zero event should be required? We discuss the matter in Sect. 4.

In this paper we also briefly consider other independence relations, and a generalisation of our results to finite non-classical probability spaces.

2 Causal (up-to- n -)closedness

2.1 Preliminary Definitions

Throughout this paper the sample spaces of the probability spaces involved are irrelevant. The crucial elements are the Boolean algebra (of which, due to Stone's theorem, we always think as of a field of sets and therefore compatible with set theoretical operators) containing the events and the measure defined on that algebra. This motivates the phrasing of the following definition in terms of pairs, instead of triples:

Definition 1 (*Probability space*) A (classical) *probability space* is a pair $\langle S, P \rangle$ such that S is a Boolean algebra and P is a function from S to $[0, 1] \subseteq \mathbb{R}$ such that

- $P(\mathbf{1}_S) = 1$;
- P is countably additive: for a countable family \mathcal{G} of pairwise disjoint members of S , $\cup \mathcal{G} \in S$ and $P(\cup \mathcal{G}) = \sum_{A \in \mathcal{G}} P(A)$.

In the following the context will usually be that of a *finite* classical probability space, i.e., a space $\langle S, P \rangle$ in which S is finite. By Stone's representation theorem, in such a case S is isomorphic—and will be identified with—the algebra of all subsets of the set $\{0, \dots, n-1\}$ for some $n \in \mathbb{N}$. In such a case the requirement of countable additivity reduces to the simple condition that for two disjoint events $A, B \in S$, $P(A \cup B) = P(A) + P(B)$. In Sect. 6 nonclassical spaces are considered, in which the Boolean algebra is exchanged for a nondistributive orthomodular lattice. The required definitions are presented therein.

In the sequel we will sometimes consider spaces of the form $\langle S^+, P^+ \rangle$, where S^+ and P^+ are as defined below:

Definition 2 Let $\langle S, P \rangle$ be a finite classical probability space. S^+ is the unique Boolean algebra whose set of atoms consists of all the non-zero probability atoms of S . P^+ is the restriction of P to S^+ .

This paper concerns a certain approach to explaining correlations; loosely speaking, this is to be done by events which screen off the correlated events and are positively statistically relevant for them. We introduce all these important notions in the following definition:

Definition 3 (*Correlation, screening off, statistical relevance*) Let $\langle S, P \rangle$ be a probability space and let $A, B \in S$. We say that:

- A and B are (positively) *correlated* whenever $P(AB) > P(A)P(B)$;
- event $C \in S$ *screens off* A and B whenever $P(AB | C) = P(A|C)P(B|C)$;
- an event $C \in S$ is *positively statistically relevant* for A if $P(A|C) > P(A|C^\perp)$;
- a partition of $\mathbf{1}_S \{C_i\}_{i \in I}$ is *statistically relevant* for A and B if, whenever $i \neq j$,

$$(P(A | C_i) - P(A | C_j))(P(B | C_i) - P(B | C_j)) > 0.$$

Notice that, according to the above definition, if C is positively statistically relevant for both A and B , then $\{C, C^\perp\}$ is statistically relevant for A and B .

In *The direction of time* (1971) Hans Reichenbach offered a causal theory of time in which a central role was played by “conjunctive forks”—triples of events A, B, C in which C is positively statistically relevant for both A and B and both C and C^\perp screen off A and B (see def. 4 below). A part of the literature refers to events defined as meeting Reichenbach's conditions for the “ C ” in such a conjunctive fork as (“Reichenbachian”) “common causes”; see e.g. Hofer-Szabó and Rédei (2004). Hofer-Szabó et al. (2000) and Hofer-Szabó and Rédei (2006) even go so far as to state that Reichenbach himself defined common causes as the middle elements of conjunctive forks with correlated extreme elements; in other words, that fulfilling

the statistical requirements for being the middle element of a conjunctive fork is sufficient to be a common cause for the correlated events. This is unfortunate since Reichenbach himself noticed that common *effects* could also meet his probabilistic requirements (Reichenbach 1971, p. 161–162) and also suggested that if there is more than one common cause for a given correlation the conditions are to be met by their disjunction, not by the causes themselves (p. 159). Reichenbach’s “Principle of the Common Cause” maintains simply that in the case of a correlation between A and B there is a common cause C such that A , B and C meet the statistical requirements for a conjunctive fork (p. 163). Nevertheless, the main results of this work pertain to problems posed in various papers by the above-cited authors. Therefore, some slight terminological changes are in order.

Definition 4 (*Statistical common cause*) Let $\langle S, P \rangle$ be a probability space. Let $A, B \in S$. Any $C \in S$ different from both A and B such that

- C screens off A and B ;
- C^\perp screens off A and B ;
- C is positively statistically relevant for both A and B ;

is called a *statistical common cause* of A and B .

Statistical common causes (henceforth “SCCs”) have at least two features relevant from the perspective of explaining correlations. First, the screening off conditions mean the correlation disappears after conditionalisation on the SCC. Second (as noted by Reichenbach), from the fact that there exists an SCC for A and B one can derive the correlation between A and B .

It is intuitive that a similar notion could be considered, with the difference that it would permit the cause to be more complicated than a simple “yes” / “no” event. This is indeed the path taken without further comment by van Fraassen (1982), but only the screening off requirement is retained. A generalisation which also takes into account the conditions of statistical relevance was developed by Hofer-Szabó and Rédei (2004); the resulting constructs were originally called “Reichenbachian common cause systems”, but, for reasons given above, we will abstain from the adjective “Reichenbachian”.

Definition 5 (*Statistical common cause system*) Let $\langle S, P \rangle$ be a probability space. A partition of $\mathbf{1}_S$ is said to be a *statistical common cause system* (SCCS) for A and B iff:

- all its members are different from both A and B ;
- all its members screen off A and B ;
- it satisfies the statistical relevance condition w.r.t. A and B .

The cardinality of the partition is called the *size* of the statistical common cause system.

As remarked above, statistical common cause systems (henceforth “SCCSs”) come in different cardinalities; they may have any countable size greater than 1.

SCCSs share the “deductive” explanatory feature of SCCs: from the assumption that one exists for A and B , the correlation between A and B is derivable.⁴

Throughout this paper, by a “common cause” we always mean a “statistical common cause”. At the beginning we usually supply the additional adjective, but then sometimes refrain from using it to conserve space, as the arguments unfortunately become rather cluttered even without the additional vocabulary.

We will now define two relations of independence. Intuitively, we will regard two events as logically independent if, when we learn that one of the events occurs (or does not occur), we cannot infer that the other occurs (or does not occur), for all four Boolean combinations.

Definition 6 (*Logical independence*) We say that events $A, B \in S$ are *logically independent* ($\langle A, B \rangle \in L_{ind}$) iff all of the following sets are nonempty: $A \cap B$, $A \cap B^\perp$, $A^\perp \cap B$ and $A^\perp \cap B^\perp$.

We say that events $A, B \in S$ are *logically independent modulo measure zero event* ($\langle A, B \rangle \in L_{ind}^+$) iff all of the following numbers are positive: $P(A \cap B)$, $P(A \cap B^\perp)$, $P(A^\perp \cap B)$ and $P(A^\perp \cap B^\perp)$.

Equivalently, two events are logically independent if neither of the events is contained in the other one, their intersection is non-empty and the union of the two is less than the whole space. Two events are logically independent modulo measure zero event if every Boolean combination of them has a non-zero probability of occurring. It is always true that $L_{ind}^+ \subseteq L_{ind}$; if there are 0-probability atoms in the space, the inclusion may be strict.

The following definition is a refinement of the SCC idea, expressing the requirement that a common cause should be meaningfully different from both correlated events.

Definition 7 (*Proper SCC(S)*) A statistical common cause C of events A and B is a *proper* statistical common cause of A and B if it differs from both A and B by more than a measure zero event. It is an *improper* SCC of these events otherwise.

An SCCS $\{C_i\}_{i \in I}$ of events A and B is a *proper* SCCS of A and B if all its elements differ from both A and B by more than a measure zero event. It is an *improper* SCCS of these events otherwise.

We will sometimes say that a probability space *contains* an SCCS, which means that the SCCS is a partition of unity of the event algebra of the space.

We now come to the main topic of this paper. Should someone prefer it, the following definition could be phrased in terms of SCCSs only.

Definition 8 (*Causal (up-to- n)-closedness*) We say that a classical probability space is *causally up-to- n -closed* w.r.t. to a relation of independence R_{ind} if all pairs of correlated events independent in the sense of R_{ind} possess a proper statistical common cause or a proper statistical common cause system of size at most n .

⁴ See Hofer-Szabó and Rédei (2004).

A classical probability space is *causally n -closed* w.r.t. to a relation of independence R_{ind} if all pairs of correlated events independent in the sense of R_{ind} possess a proper statistical common cause system of size n .

If the space is causally up-to-2-closed, in other words causally 2-closed, we also say that it is *causally closed* or *common cause closed*.

Note that, in terms of providing explanation for correlations, a space which is causally up-to-3-closed (or up-to- n -closed for any other finite n) is as good (or as bad) as a causally closed space. Namely, any correlation is provided with something that screens the correlation off and from the existence of which the given correlation can be deduced. This is the reason for which it is interesting to check whether we might have luck finding up-to-3-closed spaces, as opposed to searching “just” for causally closed spaces. Forgetting about the measure zero-related issues for a second, it turns out that while among finite classical probability spaces there is only one that is (non-trivially) causally closed, infinitely many are causally up-to-3-closed.

2.2 Summary of Results

Theorem 1 will be our main tool in proving the lemmas featured in Table 1.

Theorem 1 *Let $\langle S, P \rangle$ be a finite classical probability space with S^+ having at least 4 atoms of non-zero probability. Then P^+ is uniform if and only if $\langle S^+, P^+ \rangle$ is causally up-to-3-closed w.r.t. L_{ind}^+ .*

Lemmas 1-3 tie uniformity of P and P^+ with causal up-to-3-closedness of $\langle S, P \rangle$ with respect to the two notions of independence introduced above.

Lemma 1 *Let $\langle S, P \rangle$ be a finite classical probability space with Shaving at least 4 atoms. If P is uniform, then $\langle S, P \rangle$ is causally up-to-3-closed w.r.t. L_{ind} and L_{ind}^+ .*

Lemma 2 *Let $\langle S, P \rangle$ be a finite classical probability space with S^+ having at least 4 atoms. If P^+ is not uniform, then $\langle S, P \rangle$ is not causally up-to-3-closed w.r.t. either L_{ind} or L_{ind}^+ .*

Lemma 3 *Let $\langle S, P \rangle$ be a finite classical probability space with S^+ having at least 4 atoms. If P^+ is uniform, then $\langle S, P \rangle$ is causally up-to-3-closed w.r.t. L_{ind}^+ . All correlated pairs from $L_{ind} \setminus L_{ind}^+$ have statistical common causes, but some only have improper ones.*

Table 1 The main results of the paper

	$\langle S, P \rangle$ is up-to-3-closed w.r.t.	
	L_{ind}	L_{ind}^+
P is uniform	\Rightarrow (1)	\Rightarrow (1)
P^+ is uniform	\Leftarrow (2)	\Leftrightarrow (2,3)
	\Rightarrow^* (3)	

The numbers in parentheses correspond to lemmas below. The implications take the row heading as their left argument and the column heading as their right argument

3 Proofs

3.1 Some Useful Parameters

For expository reasons, we will not prove theorem 1 directly, but rather demonstrate its equivalent, theorem 2 (p. 9). Before proceeding with the proof, we shall introduce a few useful parameters one may associate with a pair of events A, B in a finite classical probability space $\langle S, P \rangle$.

Let n be the number of atoms in the Boolean algebra S . The size of the set of atoms lying below A in the lattice ordering of S will from now on be referred to as a , and likewise for B and b . The analogous parameter associated with the conjunction of events A and B is just the size of the intersection of the relevant sets of atoms and will be called k .

It will soon become apparent that while a and b have some utility in the discussion to follow, the more convenient parameters describe A and B in terms of the number of atoms belonging to one, but not the other. Thus we let $a' = a - k$ and $b' = b - k$. In fact, if we set $z = n - (a' + k + b')$, we obtain a set of four numbers precisely describing the blocks of the partition of the set of atoms of S into the four classes which need to be non-empty for A and B to be logically independent. It is clear that in the case of logically independent events a', b', k and z are all non-zero.

Lastly, before we begin the proof of the main result of this paper, let us state the following important lemma: when searching for statistical common causes, screening off is enough. If both an event and its complement screen off a correlation, then one of them is a statistical common cause for the correlation.

Lemma 4 *Let $\langle S, P \rangle$ be a probability space. Let $A, B, C \in S$. Suppose A and B are positively correlated. If both C and C^\perp screen off A from B , then either C or C^\perp is a statistical common cause of A and B .*

Proof As the reader may check, if events A and B are correlated, then for all events C such that $0 < P(C) < 1$

$$\frac{P(AB|C) - P(A|C)P(B|C)}{P(\neg C)} + \frac{P(AB|\neg C) - P(A|\neg C)P(B|\neg C)}{P(C)} \quad (1)$$

$$> - [P(A|C) - P(A|\neg C)][P(B|C) - P(B|\neg C)].$$

Then, if both C and C^\perp screen off A from B , the left-hand side of inequality 1 is 0. Therefore $[P(A|C) - P(A|\neg C)][P(B|C) - P(B|\neg C)]$ is positive, which means that both differences have the same sign—so either C or C^\perp meets the conditions for being a statistical common cause for A and B . \square

3.2 Proof of Theorem 1

In this section we will provide a proof of the main tool in this paper—theorem 1, formulated on p. 7. The form in which it was stated in that section is dictated by its use in the proofs of lemmas 1-3. However, when treated in isolation, it is better phrased in the following way:

Theorem 2 (Equivalent to theorem 1) *Let $\langle S, P \rangle$ be a finite classical probability space with no atoms of probability 0. Suppose S has at least 4 atoms.⁵ The following conditions are equivalent:*

Measure uniformity: P is the uniform probability measure on S ;

Causal up-to-3-closedness w.r.t. L_{ind} : $\langle S, P \rangle$ is causally up-to-3-closed w.r.t. the relation of logical independence.

Before proceeding with the proof we will provide a sketch of the construction and some requisite definitions. Instead of focusing on a particular n -atom algebra, we will show how the problem presents itself while we ‘move’ from smaller to bigger algebras. We assume without loss of generality that the set of atoms of an n -atom Boolean algebra is $\{0, 1, \dots, n - 1\}$ and that each event is a set of atoms. Consider the sequence of all finite classical probability spaces with the uniform probability measure, in which the number of atoms of the underlying Boolean algebra of the space increases by 1 at each step, beginning with the algebra with a single atom. We use the shorthand expression “at stage n ” to mean “in the probability space with uniform distribution whose underlying Boolean algebra has n atoms”. Observe that due to our convention whereby events are identified with sets of atoms, an event present at stage m (one found in the algebra from that stage) is also present at all further stages. In other words, a set of atoms defining an event at stage m can also be interpreted as defining an event at any stage m' , with $m' > m$. Thus we can naturally say that a certain event belongs to many different probability spaces; e.g. the event $\{1, 2, 11\}$ is present at stages 12, 13 and so on. Similarly, pairs of events can be present at many stages—and be correlated at some, but not at others. If they are correlated at stage m , they are correlated at all stages n , for $n > m$ (see below). The same is true of logical independence: a pair may not consist of logically independent events at stage n , because their union is the whole set of n atoms, but may become a pair of logically independent events at stage $n + 1$, when an additional atom is introduced, which does not belong to either of the events in question.⁶

Some remarks on the shape of events considered are in order. We will always be talking about pairs of events A, B , with numbers a, a', b, b', k, z and n defined as above (see Sect. 3.1). We assume (without loss of generality) $a \geq b$. Also, since we are dealing with the uniform measure, all relevant characteristics of a pair of events A, B are determined by the numbers a', b', k , and z ; therefore, for any combination of these numbers it is sufficient only to consider a single example of a pair displaying them. The rest is just a matter of renaming the atoms. For example, if we are looking for an explanation for the pair $\{\{8, 7, 3, 5\}, \{2, 8, 7\}\}$ at stage 10, or the pair $\{\{1, 3, 5, 6\}, \{1, 6, 4\}\}$ at the same stage, we shall search for an

⁵ It is easy to verify that if S has 3 atoms or less, then $\langle S, P \rangle$ contains no correlations between logically independent events.

⁶ Note that the space at stage $n + 1$ is not to be thought of as an extension of the space at stage n in the sense of the latter being embeddable in the former; we propose no measure-preserving homomorphism providing such an embedding (and indeed no such homomorphism exists between any two adjacent stages except stages 1 and 2). Thus when we speak of “the same” events being present at different stages, we simply mean that they are equal as sets of natural numbers—a property useful in the proofs to follow.

explanation for the pair $\{\{0, 1, 2, 3\}, \{2, 3, 4\}\}$ at stage 10 and then just appropriately ‘translate’ the result (explicit examples of this follow in Sect. 3.2.1). In general: the convention we adopt is for A to be a set of consecutive atoms beginning with 0, and B a set of consecutive atoms beginning with $a - k$

For illustrative purposes we propose to examine the situation at the early stages. The proof proper begins with definition 9 below. For the remainder of Sect. 3.2, by “common cause” we will always mean “proper common cause”; similarly with “common cause system”.

There are no correlated pairs of logically independent events at stage 1; similarly for stages 2, 3 and 4. (Remember the measure is uniform and so at stage 4 e.g. the pair $\{\{0, 1\}, \{1, 2\}\}$, while composed of logically independent events, is not correlated.)

First correlated pairs of logically independent events appear at stage 5. These are of one of the two following types: either $a' = b' = k = 1$, or $a' = b' = 1$ and $k = 2$. Proposition 3 from Gyenis and Rédei (2004) says that all pairs of these types have statistical common causes at stage 5. As noted above, we can without loss of generality consider just two tokens of these types—the pairs $\{\{0, 1\}, \{1, 2\}\}$ and $\{\{0, 1, 2\}, \{1, 2, 3\}\}$. In the first case, the events already formed a logically independent pair at stage 4, but were not correlated—we will say that the pair *appears from below at stage 5* (see definition 9 below). In the second case, stage 5 is the first stage where the events form a logically independent pair, and they are already correlated at that stage. We will say that the pair $\{\{0, 1, 2\}, \{1, 2, 3\}\}$ *appears from above at stage 5*. There are no other correlated pairs of logically independent events at stage 5. It will turn out that we can always find statistical common causes for pairs which appear from above or from below at a given stage.

Let us move to stage 6. A new (type of) pair appears from above— $\{\{0, 1, 2, 3\}, \{1, 2, 3, 4\}\}$. No pairs appear from below, but both pairs which appeared at stage 5 are still correlated and logically independent at stage 6 (as well as at all later stages), so they are again in need of an explanation at this higher stage. It turns out that if a correlated pair of logically independent events at stage n is ‘inherited’ from the earlier stages, i.e. it appears neither from above nor from below at stage n , we can modify the common cause which we know how to supply for it at the stage where it originally appeared to provide it with an explanation adequate at stage n . This takes the form of a statistical common cause or, in some cases, an SCCS of size 3.

Definition 9 (*Appearing from above or below*) A pair $\{A, B\}$ of events of the form $\{0, \dots, a - 1\}$, $\{a - k, \dots, a - k + b - 1\}$ *appears from above at stage n* if it is (1) logically independent at stage n , (2) not logically independent at stage $n - 1$ and (3) correlated at stage n .

A pair $\{A, B\}$ of events of the same form *appears from below at stage n* if it is (1) logically independent at stage n , (2) logically independent at stage $n - 1$ and (3) correlated at stage n , but (4) not correlated at stage $n - 1$.

We will divide common causes into types depending on whether the occurrence of a given common cause makes the occurrence of at least one member of the

correlation it explains necessary, impossible or possible with probability less than 1.⁷

Definition 10 (*1-, 0-, and #-type statistical common causes*) A proper statistical common cause C for a correlated pair of logically independent events A, B is said to be:

- *1-type* iff $P(A | C) = 1$ or $P(B | C) = 1$;
- *0-type* iff $P(A | C^\perp) = 0$ or $P(B | C^\perp) = 0$;
- *#-type* iff it is neither 1-type nor 0-type.

Notice that no proper statistical common cause C for some two logically independent, correlated events A and B can be both 1-type and 0-type at the same time.

Definition 11 (*0-type statistical common cause system*) A proper statistical common cause system of size n $\{C_i\}_{i \in \{0, \dots, n-1\}}$ is a *0-type statistical common cause system* (*0-type SCCS*) for the correlation iff $P(A | C_{n-1}) = 0$ or $P(B | C_{n-1}) = 0$.

We do not need to worry about the fact that rearranging the elements of a 0-type SCCS necessarily makes it lose the 0-type status, because during the proof the SCCSs will be explicitly construed so that their “last” element gives conditional probability 0 to both correlated events to be explained. Were this notion to be used in general, its definition should be rephrased as an existential condition: “there exists $m \leq n - 1$ such that $P(A | C_m) = 0$ and $P(B | C_m) = 0$ ”.

We will prove the following:

- if a pair appears from above at stage n , it has a statistical common cause at that stage (lemma 6);
- if a pair appears from below at stage n , it has a statistical common cause at that stage (lemma 7);
- if a pair of logically independent events is correlated at stage n and has a statistical common cause or a 0-type SCCS of size 3 at that stage, it has a statistical common cause or a 0-type SCCS of size 3 at stage $n + 1$ (lemma 8).

It should be straightforward to see that this is enough to prove theorem 2 (p. 9) in its ‘downward’ direction. Consider a correlated pair of logically independent events A, B at stage n . If it appears from above, we produce a common cause using the technique described in lemma 6. If it appears from below, we use the method from lemma 7. If it appears neither from above nor from below, it means that it was logically independent at stage $n - 1$ and was correlated at that stage, and we repeat the question at stage $n - 1$. This descent terminates at the stage where our pair first appeared, which clearly must have been either from below or from above. This allows us to apply either lemma 6 or lemma 7, as appropriate, followed by lemma 8 to move back up to stage n , where we will now be able to supply the pair with an SCC or an SCCS of size 3. As said before, the SCCs and SCCSs we will construct will always be *proper* SCCs and SCCSs.

⁷ We believe the conceptual difference between necessity and probability 1 is not important for the present topic.

Put $\text{Corr}(A, B) := P(AB) - P(A)P(B)$. $\text{Corr}(A, B)$ can always be expressed as a fraction with denominator n^2 . Of special interest to us will be the numerator of this fraction. Let us call this number $SC_n(A, B)$. (For example, if $A = \{0, 1, 2\}$ and $B = \{2, 3\}$, $SC_5(A, B) = -1$.) If $SC_n(A, B) \leq 0$, the events are not correlated at stage n . If $SC_n(A, B) > 0$, A and B are correlated at stage n and we need to find either a common cause or a common cause system of size 3 for them. The following lemma will aid us in our endeavour (remember the definitions from Sect. 3.1):

Lemma 5 *Let $\langle S_n, P \rangle$ be a finite classical probability space, S_n being the Boolean algebra with n atoms and P the uniform measure on S_n . Let $A, B \in S_n$. Then $SC_n(A, B) = kz - a'b'$.*

Proof $\text{Corr}(A, B) = P(AB) - P(A)P(B) = \frac{k}{n} - \frac{k+a'}{n} \frac{k+b'}{n} = \frac{k(n-k-a'-b')-a'b'}{n^2} = \frac{kz-a'b'}{n^2}$.
Therefore $SC_n(A, B) = kz - a'b'$. \square

An immediate consequence of this lemma is that any pair of logically independent events will eventually (at a high enough stage) be correlated—it is just a matter of injecting enough atoms into z . For example, consider events $A = \{0, 1, 2, 3, 4, 5, 6\}$, $B = \{6, 7, 8, 9, 10, 11\}$. At any stage n , $SC_n(A, B)$ is equal to $z - 30$. This means that the pair is correlated at all stages in which $z > 30$; in other words, at stages 43 and up. At some earlier stages (from 13 to 42) the pair is logically independent but not correlated; at stage 12 it is not logically independent; and the events constituting it do not fit in the algebras from stages lower than that.

Notice that since for any A, B : $SC_{n+1}(A, B) = SC_n(A, B) + k$, it follows that at the stage m where the pair first appears (either from above or from below) $SC_m(A, B)$ is positive but less than or equal to k .

We now have all the tools we need to prove theorem 2.

Proof (of theorem 2) **Measure uniformity \Rightarrow Causal up-to-3-closedness w.r.t. \mathbf{L}_{ind}**

Lemma 6 *Suppose a pair A, B appears from above at stage n . Then there exists a 1-type common cause for the correlation at that stage.*

Proof We are at stage n . Since the pair A, B appears from above at this stage, $z = 1$ and so (by lemma 5) $SC_n(A, B) = k - a'b'$. (If z was equal to 0, the events would not be logically independent at stage n ; if it was greater than 1, the events would be logically independent at stage $n - 1$ too, and so the pair would not appear from above at stage n .) Notice that since A, B are logically independent (so both a' and b' are non-zero) but correlated at stage n , $0 < SC_n(A, B) = k - a'b' < k$. Let C consist of exactly $SC_n(A, B)$ atoms from the intersection $A \cap B$. Such a C will be a screener-off for the correlation, since $P(AB | C) = 1 = P(A | C)P(B | C)$. What remains is to show that C^\perp is a screener-off as well. This follows from the observation that $P(AB | C^\perp) = \frac{k-(k-a'b')}{n-(k-a'b')} = \frac{a'b'}{n-k+a'b'} = \frac{a'b'(n-k+a'b')}{(n-k+a'b')^2} = \frac{a'b'(1+a'+b'+k)-a'b'k+a'^2b'^2}{(n-k+a'b')^2} = \frac{a'b'+a'b'^2+a'^2b'+a'^2b'^2}{(n-k+a'b')^2} = \frac{a'+a'b'}{n-k+a'b'} \cdot \frac{b'+a'b'}{n-k+a'b'} = \frac{k+a'-(k-a'b')}{n-k+a'b'}$.
 $\frac{k+b'-(k-a'b')}{n-k+a'b'} = \frac{k+a'-SC_n(A,B)}{n-k+a'b'} \cdot \frac{k+b'-SC_n(A,B)}{n-k+a'b'} = P(A | C^\perp)P(B | C^\perp)$. \square

Lemma 7 Suppose a pair A, B appears from below at stage n . Then there exists a 1-type common cause or a 0-type common cause for the correlation at that stage.

Proof **Case 1:** $k > b'$ and $a' > z$.

In this case we will construct a 1-type common cause. Let C consist of $k - b'$ atoms from $A \cap B$ and $a' - z$ atoms from $A \setminus B$. Since $C \subset A$, it screens off the correlation: $P(AB | C) = P(B | C) = 1 \cdot P(B | C) = P(A | C)P(B | C)$. We need to show that C^\perp screens off the correlation as well. This follows from the fact that
$$P(AB | C^\perp) = \frac{b'}{n - (k - b') - (a' - z)} = \frac{b'}{2b' + 2z} = \frac{2b'^2 + 2zb'}{(2b' + 2z)^2} = \frac{(b' + z)2b'}{(2b' + 2z)^2} = \frac{b' + z}{2b' + 2z} \cdot \frac{2b'}{2b' + 2z} = \frac{b' + z}{n - (k - b') - (a' - z)} \cdot \frac{2b'}{n - (k - b') - (a' - z)} = P(A | C^\perp)P(B | C^\perp).$$

Case 2: $z > b'$ and $a' > k$.

In this case we will construct a 0-type common cause. Let C^\perp consist of $a' - k$ atoms from $A \setminus B$ and $z - b'$ atoms from $(A \cup B)^\perp$. Since $C^\perp \subset B^\perp$, it screens off the correlation: $P(AB | C^\perp) = 0 = P(A | C^\perp) \cdot 0 = P(A | C^\perp)P(B | C^\perp)$. We need to show that C too screens off the correlation. This follows from the fact that
$$P(AB | C) = \frac{k}{n - (a' - k) - (z - b')} = \frac{k}{2k + 2b'} = \frac{2k^2 + 2kb'}{(2k + 2b')^2} = \frac{2k(k + b')}{(2k + 2b')^2} = \frac{2k}{2k + 2b'} \cdot \frac{k + b'}{2k + 2b'} = \frac{2k}{n - (a' - k) - (z - b')} \cdot \frac{k + b'}{n - (a' - k) - (z - b')} = P(A | C)P(B | C).$$

Case 3a: $z \geq a', k \geq a'$ and $a' > b'$.

As can be verified easily, in this case $k = z = a'$ and $b' = a' - 1$. We can construct both a 0-type common cause and a 1-type common cause. Suppose we choose to produce the former. An appropriate C^\perp would consist just of a single atom from $(A \cup B)^\perp$. C^\perp screens off the correlation because $P(AB | C^\perp) = 0 = P(A | C^\perp)P(B | C^\perp)$. That C is also a screener-off is guaranteed by the fact that
$$P(AB | C) - P(A | C)P(B | C) = \frac{k}{k + a' + b' + z - 1} - \frac{k + a'}{k + a' + b' + z - 1} \cdot \frac{k + b'}{k + a' + b' + z - 1} = \frac{k}{4k - 2} - \frac{2k}{2(2k - 1)} \cdot \frac{2k - 1}{4k - 2} = 0.$$
 To produce a 1-type common cause instead, let C consist just of a single atom from $A \cap B$. C screens off the correlation because $P(AB | C) = 1 = P(A | C)P(B | C)$. That C^\perp is also a screener-off follows from the fact that
$$P(AB | C^\perp) = \frac{k - 1}{k - 1 + a' + b' + z} = \frac{b'}{2b' + 2a'} = \frac{2b'^2 + 2a'b'}{(2b' + 2a')^2} = \frac{(a' + b')2b'}{(2b' + 2a')^2} = \frac{a' + b'}{2b' + 2a'} \cdot \frac{2b'}{2b' + 2a'} = \frac{k - 1 + a'}{2b' + 2a'} \cdot \frac{k - 1 + b'}{2b' + 2a'} = P(A | C^\perp)P(B | C^\perp).$$

Case 3b: $z = a' + 1$ and $k = a' = b'$.

In this case we will construct a 0-type common cause. Let C^\perp consist of just a single atom from $(A \cup B)^\perp$. C^\perp screens off the correlation because $P(AB | C^\perp) = 0 = P(A | C^\perp)P(B | C^\perp)$. C screens off the correlation because
$$P(AB | C) = \frac{k}{4k} = \frac{4k^2}{16k^2} = \frac{2k}{4k} \cdot \frac{2k}{4k} = \frac{k + a'}{k + a' + b' + z - 1} \cdot \frac{k + b'}{k + a' + b' + z - 1} = P(A | C)P(B | C).$$

Case 3c: $k = a' + 1$ and $z = a' = b'$.

In this case we will construct a 1-type common cause. Let C consist of just a single atom from $(A \cap B)$. As in case 3a, C screens off the correlation. That

C^\perp is also a screener-off follows from $P(AB \mid C^\perp) = \frac{a'}{4a'} = \frac{4a'^2}{16a'^2} = \frac{2a'}{4a'} \cdot \frac{2a'}{4a'} = \frac{k-1+a'}{k-1+a'+b'+z} \cdot \frac{k-1+b'}{k-1+a'+b'+z} = P(A \mid C^\perp)P(B \mid C^\perp)$. \square

Notice that the five cases used in the proof above are exhaustive. To see this, consider that $a' \geq b'$ (by our convention) and $SC_n(A, B) = kz - a'b' > 0$ (because A and B are correlated). The latter inequality rules out the possibility that $k, z \leq a', b'$. Also, if $b' \leq k \leq z \leq a'$, then the leftmost inequality must be strict, since $b' = k \leq z \leq a'$ clearly violates the condition on $SC_n(A, B)$. The remaining possibilities are as follows:

1. $k \leq b' \leq a' < z$,
2. $z \leq b' \leq a' < k$,
3. $b' < k \leq z < a'$,
4. $b' < k \leq z = a'$.

1. is further subdivided into the following cases:

- $k = b' = a' < z$ —this is Case 3b (if additionally $z > a' + 1$, then the pair A, B would have been already logically independent and correlated at the prior stage and would not appear from below at stage n),
- $k = b' < a' < z$ —this matches the conditions in Case 2,
- $k < b' \leq a' < z$ —likewise.

2. is further subdivided into the following cases:

- $z = b' = a' < k$ —this is Case 3c (a remark similar to that on the first subcase of 1. applies),
- $z = b' < a' < k$ —this matches the conditions in Case 1,
- $z < b' \leq a' < k$ —likewise.

3. matches the conditions in Case 2.

4. is further subdivided into two cases depending on whether the inequality $k \leq z$ is strict:

- $k < z$ —this matches the conditions in Case 2,
- $k = z$ —this matches the conditions in Case 3a.

Lemma 8 Suppose A, B form a pair of logically independent events correlated at stage n . Suppose further that they have a common cause or a 0-type SCCS of size 3 at that stage. Then they have a common cause or a 0-type SCCS of size 3 at stage $n + 1$.

Proof (Note that the cases are not exclusive; they are, however, exhaustive, which is enough for the present purpose.)

Case 1: A, B have a 0-type common cause at stage n .

Let C be a 0-type common cause for the correlation. When moving from stage n to $n + 1$, a new atom ($\{n + 1\}$) is added. Let $C'^\perp = C^\perp \cup \{n + 1\}$. Notice that C and C'^\perp form a partition of unity of the algebra at stage $n + 1$. C contains exclusively atoms from the algebra at stage n and so continues to be a screener off.

Notice that since C was a 0-type common cause at stage n , at that stage $P(A | C^\perp) = 0$ or $P(B | C^\perp) = 0$. Since the atom $n + 1$ lies outside the events A and B , at stage $n + 1$ we have $P(A | C'^\perp) = 0$ or $P(B | C'^\perp) = 0$, and so C'^\perp is a screener-off too. Thus C and C'^\perp are both screener-offs and compose a partition of unity at stage $n + 1$. By lemma 4 (p. 8), this is enough to conclude that A, B have a 0-type common cause at stage $n + 1$.

Case 2: A, B have a common cause which is not a 0-type common cause at stage n .

Let C be a non-0-type common cause for the correlation at stage n . Notice that both $P(AB | C)$ and $P(AB | C^\perp)$ are non-zero. In this case the ‘new’ atom cannot be added to C or C^\perp without breaking the corresponding screening-off condition. However—as we remarked in the previous case—the atom $n + 1$ lies outside the events A and B , so the singleton $\{n + 1\}$ is trivially a screener-off for the pair.

Since conditioning on $\{n + 1\}$ gives probability 0 for both A and B , the statistical relevance condition is satisfied. Therefore our explanation of the correlation at stage $n + 1$ will be a 0-type SCCS of size 3: $C' = \{C, C^\perp, \{n + 1\}\}$.⁸

Case 3: A, B have a 0-type SCCS of size 3 at stage n .

Let the partition $C = \{C_i\}_{i \in \{0,1,2\}}$ be a 0-type SCCS of size 3 at stage n for the correlation, with C_2 being the zero element (that is $P(A | C_2) = 0$ or $P(B | C_2) = 0$ (or possibly both), with the conditional probabilities involving C_0 and C_1 being positive). Let $C' = \{C_0, C_1, C_2 \cup \{n + 1\}\}$. Appending the additional atom to C_2 does not change any conditional probabilities involved, so the statistical relevance condition is satisfied. Since $n + 1 \notin A \cup B$, $C_2 \cup \{n + 1\}$ screens off the correlation at stage $n + 1$ and C' is a 0-type SCCS of size 3 at stage $n + 1$ for the correlation. \square

As mentioned above, lemmas 6–8 complete the proof of this direction of the theorem since a method is given for obtaining a statistical common cause or an SCCS of size 3 for any correlation between logically independent events in any finite probability space with uniform distribution.

We proceed with the proof of the ‘upward’ direction of theorem 2.

Causal up-to-3-closedness w.r.t. $L_{ind} \Rightarrow$ Measure uniformity

In fact, we will prove the contrapositive: if in a finite probability space with no 0-probability atoms the measure is not uniform, then there exist logically independent, correlated events A, B possessing neither a common cause nor an SCCS of size 3.⁹ In the remainder of the proof we extend the reasoning from case 2 of proposition 4 of Gyenis and Rédei (2004), which covers the case of common causes.

Consider the space with n atoms; arrange the atoms in the order of decreasing probability and label them as numbers $0, 1, \dots, n - 1$. Let $A = \{0, n - 1\}$ and $B = \{0, n - 2\}$. Gyenis and Rédei (2004) prove that A, B are correlated and do not

⁸ The fact that a correlation has an SCCS of size 3 does not necessarily mean it has no statistical common causes.

⁹ Recall that we assume the probability space under consideration has at least 4 atoms.

have a common cause. We will now show that they do not have an SCCS of size 3 either.

Suppose $C = \{C_i\}_{i \in \{0,1,2\}}$ is an SCCS of size 3 for the pair A, B . If for some $i \in \{0, 1, 2\} A \subseteq C_i$, C violates the statistical relevance condition, since for the remaining $j, k \in \{0, 1, 2\}, j \neq k, i \neq j, i \neq k, P(A | C_j) = 0 = P(A | C_k)$. Similarly if B is substituted for A in the above reasoning. It follows that none of the elements of C can contain the whole event A or B . Notice also that no C_i can contain the atoms $n - 1$ and $n - 2$, but not the atom 0, as then it would not be a screener-off. This is because in such a case $P(AB | C_i) = 0$ despite the fact that $P(A | C_i) \neq 0$ and $P(B | C_i) \neq 0$. But since C is a partition of unity of the space, each of the three atoms forming $A \cup B$ has to belong to an element of C , and so each C_i contains exactly one atom from $A \cup B$. Therefore for some $j, k \in \{0, 1, 2\} P(A | C_j) > P(A | C_k)$ but $P(B | C_j) < P(B | C_k)$, which means that C violates the statistical relevance condition. All options exhausted, we conclude that the pair A, B does not have an SCCS of size 3; thus the probability space is not causally up-to-3-closed. \square

The reasoning from the ‘upward’ direction of the theorem can be extended to show that if a probability space with no 0-probability atoms has a non-uniform probability measure, it is not causally up-to- n -closed for any $n \geq 2$. The union of the two events A and B described above only contains 3 atoms; it follows that the pair cannot have an SCCS of size greater than 3, since it would have to violate the statistical relevance condition (two or more of its elements would, when conditioned upon, give probability 0 to event A or B). This, together with proposition 3 of Gyenis and Rédei (2004) justifies the following claims:

Theorem 3 *No finite probability space with a non-uniform measure and without 0-probability atoms is causally up-to- n -closed w.r.t. L_{ind} for any $n \geq 2$.*

Corollary 9 *No finite probability space with a non-uniform measure and without 0-probability atoms is causally n -closed w.r.t. L_{ind} for any $n \geq 2$.*

The proofs of lemmas 2 and 3 in Sect. 3.3 will make it clear how to generalize both theorem 3 and corollary 9 to arbitrary finite spaces (also those possessing some 0-probability atoms) with a non-uniform measure. We omit the tedious details.

3.2.1 Examples

We will now present a few examples of how our method of finding explanations for correlations works in practice, analysing a few cases of correlated logically independent events in probability spaces of various sizes (with uniform probability distribution).

Example 1 $n = 7, A = \{0, 2, 3, 5, 6\}, B = \{1, 2, 5, 6\}$.

We see that $a' = 2, b' = 1$ and $k = 3$, so we will analyse the pair $A_1 = \{0, 1, 2, 3, 4\}, B_1 = \{2, 3, 4, 5\}$. We now check whether A_1 and B_1 were independent at stage 6, and since at that stage $A_1^+ \cap B_1^+ = \emptyset$ we conclude that they were not. Therefore the pair A_1, B_1 appears from above at stage 7. Notice that $SC_7(A_1, B_1) = 1$. By construction from lemma 6 we know that an event consisting of

just a single atom from the intersection of the two events satisfies the requirements for being a common cause of the correlation. Therefore $C = \{2\}$ is a common cause of the correlation between A and B at stage 7.

Example 2 $n = 10, A = \{2, 3, 8\}, B = \{2, 8, 9\}$.

We see that $a' = 1, b' = 1$ and $k = 2$, so we will analyse the pair $A_1 = \{0, 1, 2\}, B_1 = \{1, 2, 3\}$. Since $SC_{10}(A_1, B_1) = 11$, we conclude that the lowest stage at which the pair is correlated is 5 (as remarked earlier, SC changes by k from stage to stage). A_1 and B_1 are logically independent at that stage, but not at stage 4, which means that the pair appears from above at stage 5. We employ the same method as in the previous example to come up with a 1-type common cause of the correlation at that stage—let it be the event $\{1\}$. Now the reasoning from case 2 of lemma 8 is used to ‘translate’ the explanation to stage 6, where it becomes the following 0-type SCCS: $\{\{1\}, \{0, 2, 3, 4\}, \{5\}\}$. Case 3 of the same lemma allows us to arrive at an SCCS for A_1, B_1 at stage 10: $\{\{1\}, \{0, 2, 3, 4\}, \{5, 6, 7, 8, 9\}\}$. Its structure is as follows: one element contains a single atom from the intersection of the two events, another the remainder of $A_1 \cup B_1$ as well as one atom not belonging to any of the two events, while the third element of the SCCS contains the rest of the atoms of the algebra at stage 10. We can therefore produce a 0-type SCCS for A and B at stage 10: $\{\{2\}, \{0, 3, 8, 9\}, \{1, 4, 5, 6, 7\}\}$.

Example 3 $n = 12, A = \{2, 4, 6, 8, 9, 10, 11\}, B = \{1, 3, 6, 10, 11\}$.

We see that $a' = 4, b' = 2$ and $k = 3$, so we will analyse the pair $A_1 = \{0, 1, 2, 3, 4, 5, 6\}, B_1 = \{4, 5, 6, 7, 8\}$. We also see that A_1 and B_1 were logically independent at stage 11, but were not correlated at that stage. Therefore the pair A_1, B_1 appears from below at stage 12. Notice that $z = 3$. Therefore we see that $z > b'$ and $a' > k$, which means we can use the method from case 2 of lemma 7 to construct a 0-type common cause, whose complement consists of 1 atom from $A_1 \setminus B_1$ and 1 atom from $(A_1 \cup B_1)^\perp$. Going back to A and B , we see that the role of the complement of our common cause can be fulfilled by $C^\perp = \{0, 2\}$. Therefore $C = \{1, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ is a 0-type common cause of the correlation between A and B at stage 12.¹⁰

3.3 Proofs of Lemmas 1–3

Proof (of lemma 1) If P is uniform, then $\langle S, P \rangle$ has no 0-probability atoms, which means that $S = S^+$ and $P = P^+$. Therefore P^+ is uniform, so (by theorem 1) $\langle S^+, P^+ \rangle$ (and, consequently, $\langle S, P \rangle$) is causally up-to-3-closed w.r.t. L_{ind}^+ . But in a space with no 0-probability atoms $L_{ind} = L_{ind}^+$, therefore $\langle S, P \rangle$ is also causally up-to-3-closed w.r.t. L_{ind} . \square

¹⁰ Incidentally, if we wanted to find a 1-type common cause for A and B at stage 12, we could put $C = \{2, 11\}$, in which case $P(A | C) = 1$. However, this is not always possible and there are cases in which only 0-type common causes (or only 1-type common causes) are possible. For a concrete example, take the pair $\{\{0, 1, 2, 3, 4\}, \{4, 5\}\}$, which appears from below at stage 11 and has only 0-type common causes at that stage.

The next two proofs will require “jumping” from $\langle S^+, P^+ \rangle$ to $\langle S, P \rangle$ and vice versa. We will now have to be careful about the distinction between proper and improper SCC(S)s. Some preliminary remarks are in order.

Let $A \in S$. As before, we can think of A as a set of atoms of S . Let A^+ be the set of non-zero probability atoms in A :

$$A^+ := A \setminus \{a \mid a \text{ is an atom of } S \text{ and } P(a) = 0\}.$$

Notice that

$$P(A) = \sum_{a \in A} P(a) = \sum_{a \in A^+} P(a) = P(A^+) = P^+(A^+). \quad (2)$$

Suppose $A, B, C \in S$. From (2) it follows that if A, B are correlated in $\langle S, P \rangle$, A^+, B^+ are correlated in $\langle S^+, P^+ \rangle$. Similarly, for any $D \in S$, $P(D \mid C) = P^+(D^+ \mid C^+)$. So, if C screens off the correlated events A, B in $\langle S, P \rangle$, then C^+ screens off the correlated events A^+, B^+ in $\langle S^+, P^+ \rangle$. Also, if a family $\mathbf{C} = \{C_i\}_{i \in I}$ satisfies the statistical relevance condition w.r.t. A, B in $\langle S, P \rangle$, then the family $\mathbf{C}^+ = \{C_i^+\}_{i \in I}$ satisfies the statistical relevance condition w.r.t. A^+, B^+ in $\langle S^+, P^+ \rangle$. If $\mathbf{C} = \{C_i\}_{i \in \{0, \dots, n-1\}}$ is a proper SCCS of size n for the correlation between events A, B in $\langle S, P \rangle$, then all its elements differ from both A and B by more than a measure zero event. It follows that in such a case $\mathbf{C}^+ = \{C_i^+\}_{i \in \{0, \dots, n-1\}}$ is a proper SCCS of size n for the correlation between events A^+, B^+ in $\langle S^+, P^+ \rangle$.

Proof (of lemma 2) Since P^+ is not uniform, by theorem 1 $\langle S^+, P^+ \rangle$ is not causally up-to-3-closed w.r.t. L_{ind}^+ (and, consequently, L_{ind}). Then there exist logically independent, correlated events A^+, B^+ in S^+ which do not have a proper SCCS of size at most 3 in $\langle S^+, P^+ \rangle$. The two events are also logically independent and correlated in $\langle S, P \rangle$; it is easy to show that in $\langle S, P \rangle$ the pair $\langle A^+, B^+ \rangle$ also belongs both to L_{ind}^+ and to L_{ind} . We will show that $\langle S, P \rangle$ also contains no proper SCCS of size at most 3 for these events. For suppose that for some $m \in \{2, 3\}$, $\mathbf{C} = \{C_i\}_{i \in \mathbb{N}, i < m}$ was a proper SCCS of size m for the correlation between A^+ and B^+ in $\langle S, P \rangle$. Then $\mathbf{C}^+ := \{C_i^+\}_{i \in \mathbb{N}, i < m}$ would be a proper SCCS of size m for the correlation between A^+ and B^+ in $\langle S^+, P^+ \rangle$, but by our assumption no such SCCSs exist. We infer that the correlated events A^+, B^+ have no proper SCCS of size up to 3 in $\langle S, P \rangle$, so the space $\langle S, P \rangle$ is not causally up-to-3-closed w.r.t. either L_{ind} or L_{ind}^+ . \square

Proof (of lemma 3) Since P^+ is uniform, by theorem 1 $\langle S^+, P^+ \rangle$ is causally up-to-3-closed w.r.t. L_{ind}^+ . We will first show that also $\langle S, P \rangle$ is causally up-to-3-closed w.r.t. L_{ind}^+ . Notice that if $A, B \in S$ are correlated in $\langle S, P \rangle$ and $\langle A, B \rangle \in L_{ind}^+$, then $A^+, B^+ \in S^+$ are correlated in $\langle S^+, P^+ \rangle$ and $\langle A^+, B^+ \rangle \in L_{ind}^+$. We know that in that case there exists in $\langle S^+, P^+ \rangle$ a proper SCCS of size 2 or 3 for A^+ and B^+ . If we add the 0-probability atoms of S to one of the elements of the SCCS, we arrive at a proper SCCS of size 2 or 3 for $A, B \in S$.

It remains to consider correlated events $A, B \in S$ such that $\langle A, B \rangle \in L_{ind}$ but $\langle A, B \rangle \notin L_{ind}^+$. In such a case at least one of the probabilities from definition 6 has to

be equal to 0. It is easy to show that, since we know the two events are correlated, it can only be the case that $P(A \cap B^\perp) = 0$ or $P(B \cap A^\perp) = 0$; equivalently, $A^+ \subseteq B^+$ or $B^+ \subseteq A^+$. It may happen that $A^+ = B^+$. Let us first deal with the case of a strict inclusion; suppose without loss of generality that $A^+ \subset B^+$. If $|B^+ \setminus A^+| > 1$, take an event C such that $A^+ \subset C \subset B^+$. Since both inclusions in the last formula are strict, in such a case C is a proper statistical common cause for A and B . Notice that since $\langle A, B \rangle \in L_{ind}$, from the fact that $A^+ \subset B^+$ it follows that $A \neq A^+$. Therefore, if $|B^+ \setminus A^+| = 1$, put $C = A^+$. Such a C is an improper statistical common cause of A and B .

The last case is that in which $A^+ = B^+$. From the fact that A and B are logically independent it follows that $A \setminus B^+ \neq \emptyset$ and $B \setminus A^+ \neq \emptyset$. Therefore $A \neq A^+$ and $B \neq B^+$. We can thus put $C = A^+ (=B^+)$ to arrive at an improper statistical common cause of A and B .

When $A^+ \subseteq B^+$, it is also impossible to find (even improper) SCCSs of size 3 for A and B . For suppose $\mathbf{C} = \{C_i\}_{i \in \{0,1,2\}}$ was an SCCS for A and B . If for some $j \neq l; j, l \in \{0, 1, 2\}$ it is true that $C_j \cap A^+ = C_l \cap A^+ = \emptyset$, then $P(A | C_j) = 0 = P(A | C_l)$ and so \mathbf{C} cannot be an SCCS of A and B due to the statistical relevance condition being violated. Thus at least two elements of \mathbf{C} have to have a nonempty intersection with A^+ . Every such element C_j screens off A from B . Since by our assumption $A^+ \subseteq B^+$, it follows that $P(AB | C_j) = P(AC_j)$. Therefore the screening off condition takes the form of $P(AC_j) = P(AC_j)P(BC_j)$; and so $P(BC_j) = 1$. Since we already established that C contains at least two elements which can play the role of C_j in the last reasoning, it follows that in this case the statistical relevance condition is violated too; all options exhausted, we conclude that no SCCSs of size 3 exist for A and B when $A^+ \subseteq B^+$. The argument from this paragraph can also be applied to show that if $A^+ \subseteq B^+$ and $|B^+ \setminus A^+| \leq 1$, no proper statistical common causes for the two events exist. \square

4 The “proper” / “improper” common cause distinction and the relations of logical independence

A motivating intuition for the distinction between proper and improper common causes is that a correlation between two events should be explained by a *different* event. The difference between an event A and a cause C can manifest itself on two levels: the algebraical (A and C being not identical as elements of the event space) and the probabilistic ($P(A \cap C^\perp)$ or $P(C \cap A^\perp)$ being not equal to 0). As per definition 7, in the case of improper common causes the difference between them and at least one of the correlated events (say, A) is only algebraical. For some this is intuitively enough to dismiss C as an explanation for any correlation involving A .

One could, however, have intuitions to the contrary. First, events which differ by a measure zero event can be conceptually distinct. Second, atoms with probability 0 should perhaps be irrelevant when it comes to causal features of the particular probability space, especially when the independence relation considered is defined without any reference to probability. If the space is causally up-to- n -closed w.r.t.

L_{ind} , adding 0-probability atoms should not change its status. But consider what happens when we add a single 0-probability atom to a space which is up-to-2-closed (common cause closed) w.r.t. L_{ind} by Proposition 3 from Gyenis and Rédei (2004): the space $\langle S_5, P_u \rangle$, where S_5 is the Boolean algebra with 5 atoms $\{0, 1, \dots, 4\}$ and P_u is the uniform measure on S_5 . Label the added 0-probability atom as the number 5. It is easy to check that the pair $\langle \{3, 4\}, \{4, 5\} \rangle$ belongs to L_{ind} : is correlated and has no proper common cause. The only common cause for these events, $\{4\}$, is improper. Therefore the space is not common cause closed w.r.t. L_{ind} in the sense of Gyenis and Rédei (2004) and our definition 8; this change in the space's status has been accomplished by adding a single atom with probability 0.

It should be observed that the pair of events belongs to L_{ind} , but not to L_{ind}^+ ; and that the bigger space is still common cause closed with respect to L_{ind}^+ (although not L_{ind}).

In general, suppose $\langle S, P \rangle$ is a space without any 0 probability atoms, causally up-to- n -closed w.r.t. L_{ind} , and suppose some “extra” atoms were added, so that a new space $\langle S', P' \rangle$ is obtained, where for any atom a of S' ,

$$P'(a) = \begin{cases} P(a) & \text{for } a \in S \\ 0 & \text{for } a \in S' - S \end{cases}$$

It is easy to prove, using the techniques employed in the proof of lemma 3, that all “new” correlated pairs in $\langle S', P' \rangle$ belonging to L_{ind} have (sometimes only improper) SCCSs of size up to n . This is also true in the special case of $\langle S_5, P_u \rangle$ augmented with some 0 probability atoms. Perhaps, then, we should omit the word “proper” from the requirements for a probability space to be causally up-to- n -closed (definition 8)?

This, however, is only one half of the story. Suppose the definition of causal up-to- n -closedness were relaxed in the above way, so that explaining correlations by means of improper SCC(S)s would be admissible. Consider a space $\langle S^+, P^+ \rangle$,¹¹ in which S^+ has at least 4 atoms and P^+ is not the uniform measure on S^+ . This space, as we know, is not causally up-to-3 closed in the sense of definition 8, but it is also not causally up-to-3 closed in the “relaxed” sense, since the difference between proper and improper common causes can only manifest itself in spaces with 0 probability atoms.¹² When a new 0 probability atom m is added, every hitherto unexplained correlation between some events A and B gains an SCC in the form of the event $C := A \cup \{m\}$. All such SCCs are, of course, improper.

In short, the situation is this: if proper SCC(S)s are required, this leads to somewhat unintuitive consequences regarding causal up-to- n -closedness w.r.t. L_{ind} . Omitting the requirement results, however, in unfortunate effects regarding causal up-to- n -closedness no matter whether L_{ind} or L_{ind}^+ is considered. We think the natural solution is to keep the requirement of proper SCC(S)s in the definition of causal up-to- n -closedness, but, of the two independence relations, regard L_{ind}^+ as more

¹¹ Recall that by our convention such a space has no 0 probability atoms.

¹² This is because the spaces we are dealing with are finite—so that we can be sure the Boolean algebras considered do, indeed, have atoms—and we already require an SCC for two events A and B to be distinct from both A and B , see definition 4, p. 5.

interesting. It is the rightmost column of Table 1 that contains the most important results of this paper, then; this is fortunate, since they are a “pure” implication and an equivalence, without any special disclaimers.

5 Other Independence Relations

So far, the relation of independence under consideration—determining which correlations between two events require explanation—was either the relation of logical independence or its derivative L_{ind}^+ . Let us consider using a ‘broader’ relation $R_{ind} \supset L_{ind}$, which apart from all pairs of logically independent events would also include some pairs of logically dependent events. (The spaces under consideration are still finite.) For clarity, assume the space does not have any 0-probability atoms (so that e.g. $L_{ind} = L_{ind}^+$), but make no assumptions regarding the uniformity of the measure. Will we have more correlations to explain? If so, will they have common causes?

First, observe that if A or B is $\mathbf{1}_S$, and so $P(A)$ or $P(B)$ equals 1, there is no correlation. In the sequel assume that neither A nor B equals $\mathbf{1}_S$.

Second, note that if $A \cap B = \emptyset$, then $P(AB) = 0$ and no (positive) correlation arises.

Third, if $A^\perp \cap B^\perp = \emptyset$, there is again no positive correlation. This is because in such a case $P(AB) + P(AB^\perp) + P(A^\perp B) = 1$, and since $P(A)P(B) = P(AB) [P(AB) + P(AB^\perp) + P(A^\perp B)] + P(AB^\perp)P(A^\perp B) \geq P(AB)$, the events are not correlated.

Consider the last possible configuration in which the events A, B are logically dependent: namely, that one is a subset of the other. Suppose $A \subseteq B$. Since by our assumption both $P(A)$ and $P(B)$ are strictly less than 1, the events will be correlated. It can easily be checked¹³ that when $A \subseteq B$ but $B \neq \mathbf{1}_S$, any C which screens off the correlation and has a non-empty intersection with A (and so $P(A | C) \neq 0$) has to be a subset of B (because $P(B | C) = 1$). And since it cannot be that both C and C^\perp are subsets of B , then if C is a common cause, it is necessary that $C^\perp \cap A = \emptyset$. In the other direction, it is evident that if $A \subseteq C \subseteq B$, both C and C^\perp screen off the correlation and the statistical relevance condition is satisfied. The only pitfall is that the definition of a common cause requires it be distinct from both A and B , and so none exist when $b' = 1$.

To summarise, the only correlated pairs of logically dependent events A, B are those in which one of the events is included in the other. Assume $A \subseteq B$. Then:

- if $b = 1$, there is no common cause of the correlation;
- otherwise the common causes of the correlation are precisely all the events C such that $A \subset C \subset B$.

Lastly, notice that in a space $\langle S_n, P_u \rangle$ (S_n being the Boolean algebra with n atoms and P_u being the uniform measure) we could proceed in the opposite direction and restrict rather than broaden the relation L_{ind} . If we take the independence relation

¹³ See the last paragraph of the proof of lemma 3, p. 20.

R_{ind} to be the relation of logical independence restricted to the pairs which appear from above or below at stage n , then our probability space is common cause closed w.r.t. R_{ind} .

6 A Slight Generalisation

In this section we will show that the results of this paper, which have only concerned classical probability spaces so far, are also meaningful for finite non-classical spaces. We go back to our former practice: by “common cause” we will always mean “proper common cause”; similarly with “common cause system”.

Definition 12 (*Non-classical probability space*) An ortholattice L is *orthomodular* if $\forall a, b \in L \ a \leq b \Rightarrow b = a \vee (a^\perp \wedge b)$.

Two elements a and b of L are *orthogonal* iff $a \leq b^\perp$.

An *additive state* on an orthomodular lattice (OML) L is a map P from L to $[0,1]$ such that $P(\mathbf{1}_L) = 1$ and for any $A \subseteq L$ such that A consists of mutually orthogonal elements, if $\bigvee A$ exists, then $P(\bigvee A) = \sum_{a \in A} P(a)$.¹⁴

A *non-classical probability space* is a pair $\langle L, P \rangle$, where L is a non-distributive OML and P is an additive state on L .¹⁵

A relation of compatibility needs to be introduced. Only compatible events may be correlated; and a common cause needs to be compatible with both effects. We use the word “compatibility” because it was the one used in (Hofer-Szabó et al. 2000); “commutativity” is used in its place (see e.g. Kalmbach 1983).

Definition 13 (*Compatibility, correlation, SCC(S) in non-classical spaces*) Let $\langle L, P \rangle$ be a non-classical probability space and $a, b \in L$. Event a is said to be *compatible* with b (aCb) if $a = (a \wedge b) \vee (a \wedge b^\perp)$.

Events a, b are said to be *correlated* if aCb and the events are correlated in the sense of definition 3.

The event $x \in L$ is a *proper statistical common cause* of a and b if it fulfills the requirements from definition 7, differs from both a and b by more than a measure zero event, and is compatible both with a and with b (of course, c^\perp will be compatible, too).

A partition $\{C_i\}_{i \in I}$ of $\mathbf{1}_L$ is a *proper statistical common cause system of size n* of a and b if it satisfies the requirements of definition 7, all its elements differ from both a and b by more than a measure zero event, and all its elements are compatible both with a and b .

The notion of causal up-to- n -closedness is then immediately transferred to the context of non-classical probability spaces by substituting “non-classical” for “classical” in definition 8 (p. 7).

¹⁴ Of course, in the finite case—since a lattice always contains all suprema of doubletons by virtue of being a lattice—it would suffice to say that for any two orthogonal elements a and b , $P(a \vee b) = P(a) + P(b)$.

¹⁵ Notice that if L were distributive, $\langle L, P \rangle$ would be a classical probability space.

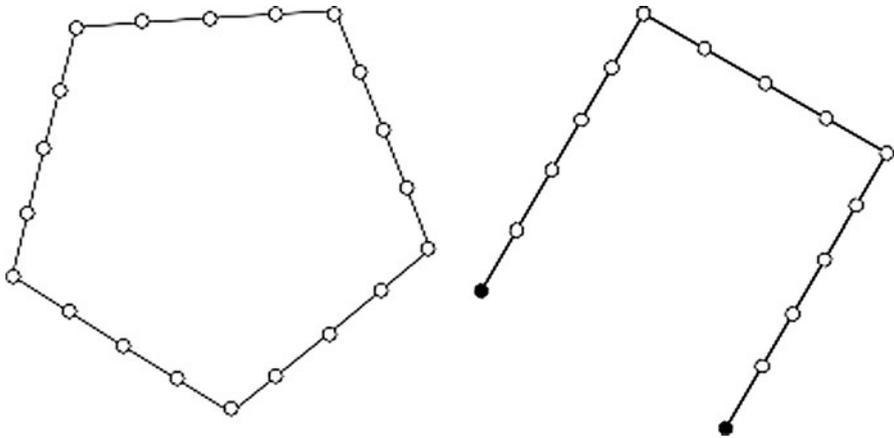


Fig. 1 Greechie diagrams of two OMLs which, if supplied with the state which assigns the number $\frac{1}{5}$ to all “white” atoms and 0 to both “black” atoms, form non-classical probability spaces which are causally up-to-2-closed (or simply “causally closed”, to use the term of Gyenis and Rédei (2004) w.r.t. L_{ind}^+



Fig. 2 In these OMLs “white” atoms have probability $\frac{1}{7}$ and the “dotted” ones $\frac{2}{7}$. The space depicted on the left is causally up-to-3-closed, but the one on the right is not

This leads us to the result of this section, which can be phrased colloquially in this way: a finite non-classical probability space is causally up-to- n closed if and only if all its blocks are causally up-to- n -closed.

Theorem 4 Suppose $\langle L, P \rangle$ is a finite non-classical probability space. Suppose all blocks of L have at least 4 atoms a such that $P(a) > 0$. Then $\langle L, P \rangle$ is causally up-to- n -closed w.r.t L_{ind} if and only if for any block B of L , the classical probability space $\langle B, P|_B \rangle$ is causally up-to- n -closed w.r.t. L_{ind} .

Proof Suppose $\langle L, P \rangle$ is causally up-to- n -closed w.r.t. L_{ind} . Let B be a block of L ; let a, b be correlated and logically independent events in $\langle B, P|_B \rangle$. Then a, b are correlated and logically independent events in $\langle L, P \rangle$, and so have an SCCS of size up to n in $\langle L, P \rangle$. But since all elements of the SCCS have to be compatible with a and b , they also have to belong to B . And so the pair has an SCCS of size up to- n in $\langle B, P|_B \rangle$.

For the other direction, suppose that for any block B of L , the space $\langle B, P|_B \rangle$ is causally up-to- n -closed w.r.t. L_{ind} . Let a, b be correlated and logically independent

events in $\langle L, P \rangle$. Being correlated entails being compatible; and so a and b belong to a block B . Since the ordering on L is induced by the orderings of the elements of B , a and b are also logically independent in B . Therefore by our assumption they have an SCCS of size up to n in $\langle B, P|_B \rangle$. This SCCS is a partition of unity of L , and so satisfies definition 13. Thus a and b have an SCCS of size up to n in $\langle L, P \rangle$. \square

6.1 Examples

We will now present a few examples of causal closedness and up-to-3-closedness of non-classical probability spaces. Figure 1 depicts two non-classical probability spaces causally closed w.r.t. L_{ind}^+ . All blocks have exactly 5 atoms of non-zero probability and each such atom receives probability $\frac{1}{5}$, and so each block is causally closed w.r.t. L_{ind}^+ . The left space is also causally closed w.r.t. L_{ind} .

The left OML in Fig. 2 has two blocks and the measure of the space is uniform on both of them, therefore the space is causally up-to-3-closed w.r.t. L_{ind} . This however is not the case with the right one: its measure is not uniform on the block with four atoms, and so there is a correlation among some two logically independent events from that block which has neither a common cause nor an SCCS of size 3. (One of these events will contain one “dotted” atom and the single “white” atom of the block; the other will contain two “dotted” atoms.) Therefore the space is not causally up-to-3-closed w.r.t. L_{ind} .

7 Conclusions and Problems

The main result of this paper is that in finite classical probability spaces with the uniform probability measure (and so no atoms with probability 0) all correlations between logically independent events have an explanation by means of a common cause or a common cause system of size 3. A few remarks are in order.

First, notice that the only SCCSs employed in our method described in Sect. 3.2 are 0-type SCCSs, and that they are required only when ‘translating’ the explanation from a smaller space to a bigger one. Sometimes (if the common cause we found in the smaller space is 0-type; see example 3 above) such a translation can succeed without invoking the notion of SCCS at all.

Second, #-type common causes, which some would view as ‘genuinely indeterministic’, are never *required* to explain a correlation – that is, a correlation can always be explained by means of a 0-type SCCS, a 0-type statistical common cause, or a 1-type statistical common cause¹⁶. Therefore one direction of the equivalence in theorem 2 can be strengthened:

¹⁶ But #-type common causes do exist: e.g. in the space with 12 atoms and the uniform measure the pair of events $\{A, B\}$, where $A = \{0, 1, 2, 3, 4, 5, 6\}$, $B = \{4, 5, 6, 7, 8\}$ (the same we dealt with in example 3, p. 18) has, apart from both 0- and 1-type common causes, a #-type common cause of shape $C = \{1, 2, 4, 5, 7, 9\}$, $C^\perp = \{0, 3, 6, 8, 10, 11\}$; $P(A | C) = \frac{2}{3}$, $P(B | C) = \frac{1}{2}$, $P(A | C^\perp) = \frac{1}{2}$, $P(B | C^\perp) = \frac{1}{3}$.

Theorem 5 Let $\langle S, P \rangle$ be a finite classical probability space. Let S^+ be the unique Boolean algebra whose set of atoms consists of all the non-zero probability atoms of S and let P^+ be the restriction of P to S^+ . Suppose S^+ has at least 4 atoms.

If P^+ is the uniform probability measure on S^+ , then any pair of positively correlated and logically independent events in $\langle S, P \rangle$ has a 1-type statistical common cause, a 0-type statistical common cause or a 0-type statistical common cause system of size 3 in $\langle S, P \rangle$.

The results of Gyenis and Rédei concerning the unique nature of the space with 5 atoms could lead one to think that in a sense it is not easy to find a space in which all correlations would be explained by Reichenbachian notions. We have shown that this is not the case—already on the level of finite spaces there are infinitely many such spaces. Moreover, recent results on causal completability show that in the case of classical probability spaces one can always extend (preserving the measure) the given (possibly infinite) space to a space which is causally closed¹⁷ and in many cases such an extension to a finite causally up-to-3-closed space is possible.¹⁸ One can think of extending a probability space while preserving the measure as of taking more factors into account when explaining some given family of correlations. We now know that it is always possible to extend the initial space so that all correlations are explained (in the Reichenbachian style) in the extension; sometimes (more often than thought before) all the explanations are there in the original space. So, we know much about explaining correlations in classical probability spaces using Reichenbachian notions: it is surprisingly easy! This strengthens the argument (which perhaps hardly needed strengthening) that a good account of causality (and causal explanation) inspired by Reichenbach should introduce something more than just bare-bones probability conditions. The account needs to be philosophically fleshed out. Another direction is investigating the fate of Reichenbach’s principle in non-classical probability spaces common in physics: in these cases decidedly less is known.¹⁹ The last option would be to move the discussion to the more general context of random variables, as opposed to events. First steps in this direction have been provided by Gyenis and Rédei (2010).

Acknowledgments The joint research was aided by The Polish Ministry for Science and Higher Education (grant no. 668/N-RNP-ESF/2010/0). L. Wroński’s research was aided by a PhD grant “Zasada wspólnej przyczyny” (“The Common Cause Principle”) of The Polish Ministry for Science and Higher Education (grant no. N N101 131936) and by the Foundation for Polish Science START Fellowship. We would like to thank the referees for their insightful comments.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

¹⁷ This was proven in Chapter 7 of Wroński (2010), but see also Marczyk and Wroński (2012) for more results on the topic and Gyenis and Rédei (2011) for a more general construction.

¹⁸ See Chapter 7 of Wroński (2010) and Marczyk and Wroński (2012).

¹⁹ See e.g. Rédei and Summers (2002).

References

- Armstrong, F. (1992). The common cause principle. *PSA: Proceedings of the Biennial meeting of the philosophy of science association* (vol. 1992). Volume Two: Symposia and Invited Papers.
- Gyenis, B., & Rédei, M. (2004). When can statistical theories be causally closed? *Foundations of Physics*, 34(9), 1284–1303.
- Gyenis, B., & Rédei, M. (2010). Causal completeness in general probability theories. In M. Suárez (Ed.), *Probabilities, Causes, and Propensities in Physics*. Synthese Library, Springer.
- Gyenis, Z., & Rédei, M. (2011). Characterizing common cause closed probability spaces. *Philosophy of Science*, 78(3), 393–409.
- Hofer-Szabó, G., & Rédei, M. (2004). Reichenbachian common cause systems. *International Journal of Theoretical Physics*, 43(7/8), 1819–1826.
- Hofer-Szabó, G., & Rédei, M. (2006). Reichenbachian common cause systems of arbitrary finite size exist. *Foundations of Physics*, 36(5), 745–756.
- Hofer-Szabó, G., Rédei, M., & Szabó, L. E. (1999). On Reichenbach's common cause principle and Reichenbach's notion of common cause. *The British Journal for the Philosophy of Science*, 50(3), 377–399.
- Hofer-Szabó, G., Rédei, M., & Szabó, L. E. (2000). Reichenbach's common cause principle: Recent results and open questions. *Reports on Philosophy*, 20, 85–107.
- Kalmbach, G. (1983). *Orthomodular lattices*. London: Academic Press.
- Marczyk, M., & Wroński, L. (2012). Completion of the causal completability problem. Accepted in the *British Journal for the Philosophy of Science*.
- Rédei, M., & Summers, S. (2002). Local primitive causality and the common cause principle in quantum field theory. *Foundations of Physics*, 32(3), 335–355.
- Reichenbach, H. (1971). *The direction of time*. University of California Press. Reprint of the 1956 edition.
- van Fraassen, B. C. (1982). The Charybdis of realism: Epistemological implications of Bell's inequality. *Synthese*, 52, 25–38. Reprinted with additions in: Cushing, J. T., & McMullin, E. (Eds.), (1989). *Philosophical consequences of quantum theory. Reflections on Bell's theorem*. University of Notre Dame Press, Indiana.
- Wroński, L. (2010). *The common cause principle. Explanation via screening off*. PhD thesis, Jagiellonian University, Kraków, archived at jagiellonian.academia.edu/LeszekWroński. http://www.academia.edu/267382/The_Common_Cause_Principle
- Wroński, L., & Marczyk, M. (2010). Only countable Reichenbachian common cause systems exist. *Foundations of Physics*, 40, 1155–1160.