Synthese
https://doi.org/10.1007/s11229-017-1650-2

Generalised Reichenbachian common cause systems

Claudio Mazzola[1]

Received: 11 May 2017 / Accepted: 28 November 2017

Abstract The principle of the common cause claims that if an improbable coincidence has occurred, there must exist a common cause. This is generally taken to mean that positive correlations between non-causally related events should disappear when conditioning on the action of some underlying common cause. The extended interpretation of the principle, by contrast, urges that common causes should be called for in order to explain positive deviations between the estimated correlation of two events and the expected value of their correlation. The aim of this paper is to provide the extended reading of the principle with a general probabilistic model, capturing the simultaneous action of a system of multiple common causes. To this end, two distinct models are elaborated, and the necessary and sufficient conditions for their existence are determined.

Keywords Common cause · Probabilistic causality · Correlation · Common cause system

1 Introduction

Chancy coincidences happen everyday, but sometimes coincidences are just too striking, or too improbable, not to reveal the presence of some coordinating process. To wit, if all the electrical appliances in a building were to shut down at exactly the same time, it would not be unreasonable to search for a breakdown in their common power supply. Similarly, if the price of petrol were to simultaneously rise in all oil

✉ Claudio Mazzola
c.mazzola@uq.edu.au

1 UQ Critical Thinking Project, School of Historical and Philosophical Inquiry, The University of Queensland, E345, Forgan Smith Building (1), Brisbane, St. Lucia, QLD 4072, Australia

Published online: 11 December 2017

importing countries, it would be a fair bet that exporters had concertedly decided to reduce extraction. The principle of the common cause is the inferential rule governing instances of this kind: informally stated, it asserts that improbable coincidences are to be put down to the action of a common cause.

Reichenbach (1956) was the first to provide the principle of the common cause with a mathematical characterisation. His treatment relied on three major ingredients. First, he represented improbable coincidences as positive probabilistic correlations between random events. Second, he demanded that common causes should increase the probability of their effects. Third, he further required that conditioning on the presence, or on the absence, of a common cause should make its effects probabilistically independent from one another.

In Mazzola (2013), however, I argued that Reichenbach's treatment is overly restrictive, as it rests on a too narrow conception of improbable coincidences, and on a correspondingly narrow understanding of the explanatory function of common causes. I accordingly proposed an improved interpretation of the principle, along with a suitably revised probabilistic model for common causes, which generalises Reichenbach's original model in two respects. On the one hand, it represents improbable coincidences not as positive correlations, but rather as positive differences between the correlation actually exhibited by a specified pair of events, and the correlation that they should be expected to exhibit according to historical data, background beliefs, or established theory. On the other hand, and correspondingly, it demands that conditioning on the presence or on the absence of a common cause should restore the expected correlation between its effects.

Reichenbach's understanding of the principle is demonstrably a special case of this interpretation, applying when the expected correlation between the events of interest is null. Nevertheless, there is one respect in which the probabilistic model proposed in Mazzola (2013) is still not general enough. Like Reichenbach's original account, in fact, it depicts the action of a single common cause, and it is accordingly inadequate to capture instances whereby two coordinated effects are brought about by a system of distinct common causes. The aim of this paper is precisely to further expand the model in this direction. To this end, two avenues for the generalisation of the model will be explored, each based on a different probabilistic characterisation for systems of common causes.

The article will be structured in four main sections. Firstly, in Sect. 2 the extended interpretation of the principle elaborated in Mazzola (2013) will be briefly outlined, and given formal treatment. Next, in Sect. 3 said interpretation will be incorporated into Hofer-Szabó and Rédei's (2004) Reichenbachian common cause systems model. Third, in Sect. 4 the extended version of the principle will be integrated with my own revisitation of Reichenbachian common cause systems (Mazzola 2012). The necessary and sufficient conditions for the existence of the resulting models in classical probability spaces will also be investigated.[1] Finally, Sect. 5 will address two major objections.

[1] The existence of common causes in general or quantum probability spaces has been investigated by Gyenis and Rédei (2014, 2016), Kitajima (2008), and Kitajima and Rédei (2015). Whether, and under what conditions, the generalised models here proposed exist in such spaces will be left as an open question.

2 Generalised conjunctive common causes

Reichenbach originally applied the principle of the common cause to pairs of positively correlated, albeit causally unrelated, events. Before introducing his probabilistic model for common causes, a definition of probabilistic correlation is thus needed:

Definition 1 Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B, C \in \Omega$ such that $p(C) \neq 0$, we define:

$$Corr(A, B|C) := p(A \wedge B|C) - p(A|C)\, p(B|C). \tag{1}$$

Moreover,

$$Corr(A, B) := p(A \wedge B) - p(A)\, p(B). \tag{2}$$

The expression $Corr(A, B|C)$ denotes the *correlation* of events A and B *conditional* on event C. The expression $Corr(A, B)$, instead, denotes the *absolute correlation* or *unconditional correlation* of events A and B. Two events are said to be *positively* (*negatively*) *correlated* (conditional on another event) if their correlation (conditional on said event) is greater (smaller) than zero; by the same token, they are said to be *uncorrelated* or *probabilistically independent* (conditional on another event) if their correlation (conditional on that event) is equal to zero.

The existence of a positive correlation between two events is often an indication that one of them is a cause of the other. However, this is not invariably the case: as is well known, correlation does not imply causation. Reichenbach's interpretation of the common cause principle could indeed be seen as an attempt to preserve a one-one correspondence between probabilistic correlation and causal dependence (Hofer-Szabó et al. 2013): in his account, acquiring full information about the occurrence of common causes should dissolve, as it were, any positive correlation between causally unrelated events. Reichenbach gave formal shape to this intuition by demanding that conditioning on the presence of a common cause, or on its absence, should make its effects probabilistically independent. The result was a probabilistic model for common causes known as *conjunctive fork*. With only a slight terminological modification and few minor notational variants, we can introduce his model as follows:

Definition 2 Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any three distinct $A, B, C \in \Omega$, the event C is a conjunctive common cause for $Corr(A, B)$ if and only if:

$$p(C) \neq 0 \tag{3}$$
$$p\left(\overline{C}\right) \neq 0 \tag{4}$$
$$Corr(A, B|C) = 0 \tag{5}$$
$$Corr\left(A, B\left|\overline{C}\right.\right) = 0 \tag{6}$$

$$p(A|C) - p(A|\overline{C}) > 0 \tag{7}$$
$$p(B|C) - p(B|\overline{C}) > 0. \tag{8}$$

Conjunctive common causes, as just defined, are intended to explain the occurrence of non-causal positive correlations in two ways. On the one hand they increase the joint probability of their effects, consequently favouring their correlation, as established by the following proposition:

Proposition 1 *Let* (Ω, p) *be a classical probability space with* σ*-algebra of random events* Ω *and probability measure* p*. For any three distinct* $A, B, C \in \Omega$*, if* C *is a conjunctive common cause for* $Corr(A, B)$*, then:*

$$Corr(A, B) > 0. \tag{9}$$

On the other hand, conditions (5)–(6) demand that the correlation between the effects of a conjunctive common cause should disappear when conditioning on the occurrence, or on the absence, of said cause: in jargon, we say that the common cause *screens-off* the two effects from one another. This is meant to indicate that the positive correlation between the two effects is purely epiphenomenal, being a mere by-product of the underlying action of the common cause.

Reichebach's conjunctive common cause model has exerted considerable influence in both probabilistic causal modelling and the philosophy of science. To mention but few of its contributions to the latter field, it anticipated the probabilistic causality program (Good 1961; Suppes 1970; Cartwright 1979; Skyrms 1980; Eells 1991), fostered the development of probabilistic accounts of scientific explanation (Salmon 1971; Suppes and Zaniotti 1981), and inspired causal interpretations of Bell's no-go theorem in quantum physics (van Fraassen 1991; Graßhoff et al. 2006). Simultaneously, the Bayesian Networks movement in probabilistic causal modelling incorporated and generalised the screening-off constraints (5)–(6) in the guise of the so-called Causal Markov Condition, according to which any two variables that are not related as cause and effect must be probabilistically independent conditional on the set of their direct causes (Pearl 1988, 2000; Spirtes et al. 2001). Nonetheless, the conjunctive common cause model relies on a demonstrably restrictive understanding of the principle of the common cause, and on a correspondingly narrow conception of the explanatory function performed by common causes.

To fully appreciate this, it will be instructive to start by taking a deeper look at the very thing the principle of the common cause is intended to apply to: improbable coincidences. Reichenbach, as we saw, understood improbable coincidences as positive correlations between causally unrelated events. Positively correlated events tend to be coinstantiated, so it is clear why positive correlations can be used to give coincidences a probabilistic representation. The problem is: in what sense, then, can coincidences between causally unrelated events be deemed *improbable*? The underlying presupposition is that *in general* causally unrelated events tend to be uncorrelated, so *in general* positive correlations between such events are not to be expected. Reichenbach, in other words, applied the principle of the common cause to pairs of events that happen to be positively correlated, even though we would expect them to be not. To be even more explicit: he applied the principle to cases where the observed value of the correlation between two events is strictly higher than its expected value, *which is zero*.

Once the principle is presented in this way, however, it becomes apparent that there is no reason not to demand that it should equally apply to *all* cases in which two events are more strongly correlated than expected, whatever the value of their expected correlation. The *extended principle of the common cause* is specifically tailored to meet this demand. Compressed in one sentence, it claims that the role of common causes is to explain statistically significant deviations between the estimated value of a correlation and its expected value, by conditionally restoring the latter. *If two events are more strongly correlated than they should, there must exist a common cause.*

To illustrate, let us consider an economic example. Let us imagine that an econometric analysis revealed a strong positive correlation between holding a postgraduate degree and earning higher-than-average income. This positive correlation, in and of itself, would not be surprising, as it would be consistent with both common sense and microeconomic theory: people who study more are likely to earn higher wages, owing to the comparatively scarce supply and higher productivity of skilled labour. But suppose that, in the case at hand, the estimated correlation were remarkably strong: strong enough to be significantly dissimilar from the average correlation reported by other similar studies. Then, excluding any mistakes in the analysis, it would be natural for one to wonder if there were anything about the selected sample, which could bring about said discrepancy.

The extended principle of common cause urges that the explanation should be sought in the presence of some unacknowledged common cause. To wit, we may imagine that the econometric analysis in our example were conducted in a relatively wealthy subpopulation. People coming from wealthy families are more likely to undergo additional years of study, since they can more easily afford the opportunity costs this involves. Moreover, they are more likely to earn their degrees from renowned but expensive academic institutions, whose graduates have a higher chance to be hired in high-earning appointments. By simultaneously increasing the probability of holding a postgraduate degree and the probability of earning higher-than-average income, family wealth would consequently increase their joint probability, and explain their stronger-than-usual correlation.

Remarkably, in this case it would be unreasonable to require that the correlation between holding a postgraduate degree and of earning higher-than-average income should disappear conditional on family wealth: after all, as we already noticed, some positive correlation between wage and qualification is to be expected. Rather, conditioning on the common cause should restore the expected correlation between the two events, consequently eliminating the apparent disagreement between the econometric analysis and the preceding studies.

To provide the extended principle of the common cause with some formal bite, let us first define:

Definition 3 Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$, the *deviation* of $Corr\,(A, B)$ is the quantity

$$\delta\,(A, B) := Corr\,(A, B) - Corr_e\,(A, B)\,, \tag{10}$$

where $Corr_e(A, B)$ denotes the *expected correlation* between A and B.

Notice that the notions of *deviation* and *expectation*, as they are understood here, are not necessarily restricted to the corresponding statistical concepts: in particular, the expected correlation between two values may be determined by non-statistical means, e.g. on the basis of logical or mathematical rules, or simply on the grounds of entrenched prior beliefs.

On this basis, we can now define:

Definition 4 Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any three distinct $A, B, C \in \Omega$, the event C is a *generalised common cause* for $\delta(A, B)$ if and only if:

$$p(C) \neq 0 \tag{11}$$

$$p\left(\overline{C}\right) \neq 0 \tag{12}$$

$$Corr(A, B\,|C) = Corr_e(A, B) \tag{13}$$

$$Corr\left(A, B\,\middle|\overline{C}\right) = Corr_e(A, B) \tag{14}$$

$$p(A\,|C) - p\left(A\,\middle|\overline{C}\right) > 0 \tag{15}$$

$$p(B\,|C) - p\left(B\,\middle|\overline{C}\right) > 0. \tag{16}$$

Just like conjunctive common causes do for positive correlations, generalised common causes explain positive deviations in two ways. On the one hand, (13)–(14) demand that conditioning on the presence of a common cause, or on its absence, should restore the expected correlation between its effects. On the other hand, generalised common causes increase the unconditional correlation between their effects, consequently generating the observed discrepancy between the estimated value of said correlation and its expected value:

Proposition 2 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any three distinct $A, B, C \in \Omega$, if C is a generalised common cause for $\delta(A, B)$ then*

$$\delta(A, B) > 0. \tag{17}$$

Quite evidently, conjunctive common causes can be thought of as generalised common causes whose effects are expected to be uncorrelated. Nonetheless, the generalised common cause model is demonstrably immune from some of the most common objections to the standard interpretation of the common cause principle.

For one thing, it has been objected that the screening-off conditions (5)–(6) are too restrictive, either because they are only satisfied by deterministic common causes (van Fraassen 1980; Cartwright 1999), or because they exclusively apply when the effects of a common cause are independently produced (Salmon 1984; Cartwright 1988). This objection is easily met by the generalised common cause model, which drops (5)–(6) in favour of the more general constraints (13)–(14). For another thing, it has been contended that non-causal positive correlations that result from logical, mathematical, semantic, or nomic relations do not generally admit of a conjunctive common cause

(Arntzenius 1992; Williamson 2005). The existence of similar correlations is clearly detrimental to the common understanding of the principle of the common cause, but it is perfectly consistent with its extended version. The reason is that, according to the extended principle, similar correlations simply do not *call* for a common cause explanation. By hypothesis, they are determined by logical, mathematical, semantic, or physical laws, so they must be expected. They accordingly fall outside the scope of the extended common cause principle, and as such they can be no counterexample to it.

One may retort, at this point, that generalised common causes are exposed to an objection that is often made to other models, such as the *interactive fork* model (Salmon 1984), that are not strictly committed to the screening-off conditions (5)–(6). The objection has it that the role of the common cause principle is to ensure that every positive correlation between causally independent events should be given a fully causal explanation, in adherence with the metaphysical thesis that there can be no correlation without causation. The screening-off condition is instrumental in this respect, as it guarantees that any correlation between causally unrelated events should be *entirely* explained by the action of the common cause. By contrast, renouncing the screening-off condition would leave some positive correlations at least partially unexplained. The problem with this objection is that the thesis that every correlation must have a causal basis is, quite simply, false. Correlations, as we have just noticed, may arise from all sorts of non-causal relations. The extended principle of the common cause explicitly takes this fact into account. The modified conditions (13)–(14), in particular, encapsulate the idea that common causes should explain a correlation *only as long as that correlation is not determined either by a direct causal influence between the correlated pair, or by other non-causal means*, i.e. as long as that correlation is not to be expected. The generalised common cause model is clearly an improvement on conjunctive common causes in this respect, as it is not similarly tied to a questionable metaphysical assumption.

The following sections will be dedicated to further enrich the generalised common cause model, so as to cover systems of multiple common causes.

3 Generalised HR-Reichenbachian common cause systems

The first attempt to extend the conjunctive common cause model to comprise systems of multiple common causes was made by Hofer-Szabó and Rédei (2004), who proposed, to this end, the following definition:

Definition 5 Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$, a *HR-Reichenbachian Common Cause System (HR-RCCS)* of size $n \geq 2$ for $Corr\,(A, B)$ is a partition $\{C_i\}_{i=1}^{n}$ of Ω such that:

$$p\,(C_i) \neq 0 \quad (i = 1, \ldots, n) \tag{18}$$

$$Corr\,(A, B\,|C_i\,) = 0 \quad (i = 1, \ldots, n) \tag{19}$$

$$\left[p\,(A\,|C_i\,) - p\left(A\,\middle|C_j\,\right)\right]\left[p\,(B\,|C_i\,) - p\left(B\,\middle|C_j\,\right)\right] > 0 \quad (1, \ldots, n{=}i{\neq}j{=}1, \ldots, n). \tag{20}$$

Hofer-Szabó and Rédei refer to Reichenbachian Common Cause Systems using the acronym RCCS. The acronym HR-RCCS is here employed to distinguish their model from the one utilized in the next section.

The notion of a HR-RCCS is meant to generalise the notion of a conjunctive common cause in two respects. On the one hand, Hofer-Szabó and Rédei demonstrate that only positively correlated pairs admit of a HR-RCCS, thus replicating the result of Proposition 1. On the other hand, conditions (18), (19) and (20) are intended to generalise, respectively, conditions (3)–(12), (5)–(6), and (7)–(8) from Definition 2. Specifically, (19) demands that each element of a HR-RCCS should screen-off its common effects from one another. This means that HR-RCCSs increase the correlation between otherwise uncorrelated pairs, emulating as a consequence the explanatory function of conjunctive common causes.

Modifying the above definition in accordance with the extended interpretation of the common cause principle only requires replacing the screening-off condition (19) with a suitably generalised variant of (13)–(14). Let us accordingly define:

Definition 6 Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$, a *Generalised HR-Reichenbachian Common Cause System (GHR-RCCS)* of size $n \geq 2$ for $\delta(A, B)$ is a partition $\{C_i\}_{i=1}^n$ of Ω such that:

$$p(C_i) \neq 0 \quad (i = 1, \ldots, n) \tag{21}$$

$$Corr(A, B \,|C_i) = Corr_e(A, B) \quad (i = 1, \ldots, n) \tag{22}$$

$$\left[p(A\,|C_i) - p\left(A\,\middle|C_j\right)\right]\left[p(B\,|C_i) - p\left(B\,\middle|C_j\right)\right] > 0 \quad (1, \ldots, n{=}i{\neq}j{=}1, \ldots, n). \tag{23}$$

This definition generalises at once Definitions 4 and 5: it extends the former by admitting systems of *any number* of common causes;[2] it extends the latter by requiring that every common cause in a system should restore the expected correlation between its two effects *whatever its value*.

Not surprisingly, every GHR-RCCS increases the correlation between its effects, consequently emulating the explanatory function of generalised common causes.

Proposition 3 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ and any $\{C_i\}_{i=1}^n \subseteq \Omega$, if $\{C_i\}_{i=1}^n$ is a GHR-RCCS of size $n \geq 2$ for $\delta(A, B)$, then (17) obtains.*

To show this, let us first prove the following lemma:

Lemma 1 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. Let $A, B \in \Omega$ and let $\{C_i\}_{i=1}^n$ be a partition of*

[2] Or, more exactly, any *finite or countable* number: in fact, no uncountable set of disjoint and non-zero probability events can exist given a countably additive probability measure (Wroński and Marczyk 2010). Thanks to an anonymous reviewer for pointing this out.

Ω satisfying conditions (21)–(22). *Then:*

$$\delta(A,B)=\frac{1}{2}\sum_{i,j=1}^{n}p(C_i)\,p\left(C_j\right)[p(A\,|C_i)-p\left(A\,\middle|C_j\right)][p(B\,|C_i)-p\left(B\,\middle|C_j\right)]. \tag{24}$$

Proof Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. Let $A, B \in \Omega$ and let $\{C_i\}_{i=1}^{n}$ be a partition of Ω satisfying (21). From the theorem of total probability it follows that:

$$\begin{aligned}Corr(A,B)&=\frac{1}{2}\sum_{i,j=1}^{n}p(C_i)\,p\left(C_j\right)\left[p(A\,|C_i)-p\left(A\,\middle|C_j\right)\right]\left[p(B\,|C_i)-p\left(B\,\middle|C_j\right)\right]\\&+\frac{1}{2}\left[\sum_{i=1}^{n}p(C_i)\,Corr(A,B\,|C_i)+\sum_{j=1}^{n}p\left(C_j\right)Corr\left(A,B\,\middle|C_j\right)\right].\end{aligned} \tag{25}$$

On the other hand, by hypothesis $\{C_i\}_{i\in I}$ is a partition of the given probability space, which implies that:

$$\sum_{i=1}^{n}p(C_i)=1. \tag{26}$$

Further assuming (22) will therefore produce the following equality:

$$\begin{aligned}&Corr(A,B)-Corr_e(A,B)\\&=\frac{1}{2}\sum_{i,j=1}^{n}p(C_i)\,p\left(C_j\right)[p(A\,|C_i)-p\left(A\,\middle|C_j\right)][p(B\,|C_i)-p\left(B\,\middle|C_j\right)],\end{aligned} \tag{27}$$

which in the light of (10) is but a different formulation of (24). □

Demonstrating Proposition 5 on this basis would be straightforward, so we are omitting the details of the proof. One interesting thing to notice about this demonstration, however, is that setting $Corr_e(A,B) = 0$ in (27) would reduce it to the equation employed by Hofer-Szabó and Rédei to demonstrate that HR-RCCSs produce positive correlations. This fact, in itself, is further confirmation of the adequacy of GHR-RCCSs as a generalisation of HR-RCCSs.

3.1 Existence of GHR-RCCSs

Hofer-Szabó and Rédei (2006) argue that a HR-RCCS of arbitrary finite size exists for every positively correlated pair of events, in some suitable extension of the original probability space.[3] The discussion to follow will be dedicated to establish a similar

[3] For the existence of HR-RCCS of countably infinite size, see Marczyk and Wroński (2015).

result for GHR-RCCSs. Remarkably, it will turn out that not all positive deviations admit of a GHR-RCCS.

Hofer-Szabó and Rédei's proof proceeds by noticing that, in general, a set $\{C_i\}_{i=1}^n$ is a HR-RCCS of size $n \geq 2$ for $Corr\,(A, B)$ in probability space (Ω, p) if and only if the values of $p\,(A\,|C_1\,)$, …, $p\,(A\,|C_n\,)$, $p\,(B\,|C_1\,)$, …, $p\,(B\,|C_n\,)$, and $p\,(C_1)$, …, $p\,(C_n)$ satisfy some specified constraints. They call any set $\{a_i, b_i, c_i\}_{i=1}^n$ of $3n$ numbers satisfying said constraints *admissible* for $Corr\,(A, B)$ and demonstrate that, for any two positively correlated events A and B and any $n \geq 2$, a set of n admissible numbers for $Corr\,(A, B)$ can be found. On this basis, they finally show how an extension of the given probability space can always be constructed, in which some partition $\{C_i\}_{i=1}^n$ exists such that $n \geq 2$ and the values of $p\,(A\,|C_1\,)$, …, $p\,(A\,|C_n\,)$, $p\,(B\,|C_1\,)$, …, $p\,(B\,|C_n\,)$, and $p\,(C_1)$, …, $p\,(C_n)$ are admissible for $Corr\,(A, B)$, thereby establishing the existence of a HR-RCCS of size n for $Corr\,(A, B)$ in that space.

The following proof will follow the broad logical structure of Hofer-Szabó and Rédei's argumentation. Our first step will consist in identifying the necessary and sufficient conditions that must be satisfied by the values of $p\,(A\,|C_1\,)$, …, $p\,(A\,|C_n\,)$, $p\,(B\,|C_1\,)$, …, $p\,(B\,|C_n\,)$, $p\,(A \wedge B\,|C_1\,)$, …, $p\,(A \wedge B\,|C_n\,)$, and $p\,(C_1)$, …, $p\,(C_n)$ to make $\{C_i\}_{i=1}^n$ a GHR-RCCS of size $n \geq 2$ for $\delta\,(A, B)$. This, however, will be done in two stages, as some of the conditions that we are going to single out will be shared by the model to be developed in Sect. 4. Let us begin by isolating these.

Definition 7 Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ satisfying (17) and any $n \geq 2$, the set

$$\{a_i, b_i, c_i, d_i\}_{i=1}^n$$

of real numbers is called *quasi-admissible for* $\delta\,(A, B)$ if and only if the following conditions hold:

$$\sum_{i=1}^n a_i c_i = p\,(A) \tag{28}$$

$$\sum_{i=1}^n b_i c_i = p\,(B) \tag{29}$$

$$\sum_{i=1}^n c_i = 1 \tag{30}$$

$$d_i - a_i b_i = Corr_e\,(A, B) \qquad (i = 1, \ldots, n) \tag{31}$$

$$0 \leq a_i, b_i, d_i \leq 1 \qquad (i = 1, \ldots, n) \tag{32}$$

$$0 < c_i < 1 \qquad (i = 1, \ldots, n). \tag{33}$$

The attentive reader will have noticed that, for each $n \geq 2$, quasi-admissible sets include $4n$ numbers, whereas admissible sets, as defined by Hofer-Szabó and Rédei, include only $3n$ numbers. Moreover, while (28)–(30) and (32)–(33) are either identical to or straightforward generalizations of some of Hofer-Szabó and Rédei's original conditions for admissible numbers, constraint (31) is not. Similar changes are needed to avoid a logical mistake in their original proof, along the lines illustrated in more detail by Mazzola and Evans (2017).

To complete this part of the proof, we need to supplement quasi-admissible sets of numbers with one more condition:

Definition 8 Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ and any $n \geq 2$, a set $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ of real numbers is called *HR-admissible* for $\delta(A, B)$ if and only if it is quasi-admissible for $\delta(A, B)$ and it further satisfies

$$[a_i - a_j][b_i - b_j] > 0 \qquad (1, \ldots, n = i \neq j = 1, \ldots, n). \tag{34}$$

The adequacy of the above definition is testified by the following lemma, whose proof is straightforward and which can consequently be omitted:

Lemma 2 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ and any $\{C_i\}_{i=1}^n \subseteq \Omega$ where $n \geq 2$, the set $\{C_i\}_{i=1}^n$ is a GHR-RCCS of size n for $\delta(A, B)$ if and only if there exists a set $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ of HR-admissible numbers for $\delta(A, B)$ such that*

$$p(C_i) = c_i \qquad (i = 1, \ldots, n) \tag{35}$$
$$p(A|C_i) = a_i \qquad (i = 1, \ldots, n) \tag{36}$$
$$p(B|C_i) = b_i \qquad (i = 1, \ldots, n) \tag{37}$$
$$p(A \wedge B|C_i) = d_i \qquad (i = 1, \ldots, n). \tag{38}$$

The next step in our proof will be to establish the necessary and sufficient conditions for the existence of HR-admissible numbers for $\delta(A, B)$. To this purpose, however, we shall need the following lemma:

Lemma 3 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. Moreover, let $A, B \in \Omega$ and let $\{C_i\}_{i=1}^n \subseteq \Omega$ with $n \geq 2$. Then, any set $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ of real numbers satisfying identities (35)–(38) is quasi-admissible for $\delta(A, B)$ if and only if it further satisfies (32)–(33) as well as*

$$a_n = \frac{a - \sum_{k=1}^{n-1} c_k a_k}{1 - \sum_{k=1}^{n-1} c_k} \tag{39}$$

$$b_n = \frac{b - \sum_{k=1}^{n-1} c_k b_k}{1 - \sum_{k=1}^{n-1} c_k} \tag{40}$$

$$c_n = 1 - \sum_{k=1}^{n-1} c_k \tag{41}$$

$$d_n = \varepsilon + \frac{\left[a - \sum_{k=1}^{n-1} a_k c_k\right]\left[b - \sum_{k=1}^{n-1} b_k c_k\right]}{\left[1 - \sum_{k=1}^{n-1} c_k\right]^2} \tag{42}$$

$$d_k = \varepsilon + a_k b_k \qquad (k = 1, \ldots, n-1), \tag{43}$$

where

$$a = p(A) \tag{44}$$

$$b = p(B) \tag{45}$$

$$\varepsilon = Corr_e(A, B). \tag{46}$$

Proof Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. Moreover, let $A, B \in \Omega$ satisfy (17) and let the set $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ of $n \geq 2$ real numbers satisfy conditions (35)–(38) and (32)–(33). Finally, let (44)–(46) be in place.

Given the aforesaid hypothesis, (28)–(30) can be directly obtained from (39)–(41) thanks to the theorem of total probability, and vice-versa. Therefore, we only need to show that (31) obtains if and only if (42)–(43) do. To this purpose, let us first observe that, as a further consequence of the theorem of total probability, the following equality holds:

$$d_n = [d_n - a_n b_n] + \frac{\left[a - \sum_{k=i}^{n-1} a_k c_k\right]\left[b - \sum_{k=1}^{n-1} b_k c_k\right]}{\left[1 - \sum_{k=1}^{n-1} c_k\right]^2}. \tag{47}$$

Thanks to (47) it is then immediate to verify that (42)–(43) are simultaneously satisfied if (31) is. Conversely, let us suppose that (42)–(43) are the case. Then (42) and (47) will jointly imply that

$$d_n - a_n b_n = \varepsilon, \tag{48}$$

which, together with (43), straightforwardly implies (31), as required. □

Endowed with the above result, we are now in a position to determine the necessary conditions so that, in general, HR-admissible numbers $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ could exist for $\delta(A, B)$ and $n \geq 2$. Quite interestingly:

Lemma 4 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$, no HR-admissible numbers exist for $\delta(A, B)$ if*

$$Corr_e(A, B) + p(A)\, p(B) \leq 0. \tag{49}$$

Proof Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p, and let $A, B \in \Omega$ be arbitrarily chosen. Lemma 4 will be established by contraposition, so let us assume that, for some $n \geq 2$, a set

$\{a_i, b_i, c_i, d_i\}_{i=1}^n$ of HR-admissible numbers does exist for $\delta(A, B)$. Moreover, let us assume identities (44)–(46).

To prove our lemma, two preliminary steps will be required. First, we shall prove that some $a_j, b_k \in \{a_i, b_i, c_i, d_i\}_{i=1}^n$ exist such that

$$a - a_j > 0 \tag{50}$$

$$b - b_k > 0. \tag{51}$$

Next, on that basis, we shall demonstrate that at least some such $a_j, a_k \in \{a_i, b_i, c_i, d_i\}_{i=1}^n$ exist, for which $j = k$.

To establish the first claim, let us begin by noticing that, as a plain consequence of (28)–(30):

$$0 = a - a = \sum_{i=1}^n ac_i - \sum_{i=1}^n a_i c_i = \sum_{i=1}^n c_i(a - a_i) \tag{52}$$

$$0 = b - b = \sum_{i=1}^n bc_i - \sum_{i=1}^n b_i c_i = \sum_{i=1}^n c_i(b - b_i). \tag{53}$$

On the other hand, (33) demands that $c_i > 0$ for all $i = 1, \ldots, n$, while (34) implies that $a_j = a$ and $b_k = b$ can be satisfied by at most one term a_j and one term b_k for $j, k = 1, \ldots, n \geq 2$. The above equalities therefore imply that $a - a_i$ should be positive for some values of i and negative for others, while similarly $b - b_i$ should be positive for some values of i and negative for others. This is enough to prove (50) and (51), as desired.

To prove our second auxiliary result, let us first relabel all numbers in $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ so that

$$a_1 < \cdots < a_k < a \leq a_{k+1} < \cdots < a_n. \tag{54}$$

This in turn implies that

$$a_i - a_j < 0 \qquad i = 1, \ldots, k; \qquad j = k+1, \ldots, n. \tag{55}$$

Now, let us proceed by reductio, and let us assume that

$$b_i - b > 0 \qquad i = 1, \ldots, k. \tag{56}$$

Then, according to the result previously established, some $b_j \in \{b_i\}_{i=k+1}^n \subset \{a_i, b_i, c_i, d_i\}_{i=1}^n$ should exist such that

$$b - b_j > 0. \tag{57}$$

However, in that case

$$b_i - b_j > 0 \qquad i = 1, \ldots, k \tag{58}$$

would ensue. Together with (55), this would imply

$$[a_i - a_j][b_i - b_j] < 0 \qquad i = 1, \ldots, k \tag{59}$$

consequently contradicting (34). By reductio, this shows that (50)–(51) must be satisfied for some $a_j, b_k \in \{a_i, b_i, c_i, d_i\}_{i=1}^n$ where $j = k$.

Let us now come to the main part of our proof. Thanks to the results so established, we can now safely claim that, for any set $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ of HR-admissible numbers, some $a_i, b_i \in \{a_i, b_i, c_i, d_i\}_{i=1}^n$ are always to be found such that

$$ab - a_i b_i > 0. \tag{60}$$

Together with (31) and (32), this implies that

$$ab + \varepsilon > a_i b_i + \varepsilon = d_i \geq 0, \tag{61}$$

contradicting (49). By contraposition, this means that whenever (49) is satisfied, no set $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ of real numbers can satisfy (31) given the other conditions for a HR-admissible set for $\delta(A, B)$. Hence, no HR-admissible set can exist for $\delta(A, B)$. $\square$

Let us now move to the sufficient condition for the existence of HR-admissible numbers $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ for $\delta(A, B)$ and $n \geq 2$. Remarkably, this turns out to be the same as the necessary condition:

Lemma 5 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ satisfying (17), a set $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ of HR-admissible numbers for $\delta(A, B)$ exists for each $n \geq 2$ if*

$$Corr_e(A, B) + p(A)\, p(B) > 0. \tag{62}$$

The lengthy demonstration of this lemma can be found in the "Appendix". Before getting to the end of our existential proof, we need one more definition:

Definition 9 Let (Ω, p) and (Ω', p') be classical probability spaces with σ-algebras of random events Ω and Ω' and with probability measures p and p', respectively. Then (Ω', p') is called an *extension* of (Ω, p) if and only if there exists an injective lattice homomorphism $h : \Omega \to \Omega'$, preserving complementation, such that

$$p'(h(X)) = p(X) \qquad \text{for all } X \in \Omega. \tag{63}$$

The results of our demonstration can thus be crystallized into the following proposition:

Proposition 4 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ satisfying (17) and any $n \geq 2$, an extension (Ω', p') of (Ω, p) including a GHR-RCCS of size n for $\delta(A, B)$ exists if and only if A and B satisfy (62).*

Proof The only-if clause immediately follows from Lemmas 2 and 4. Proof of the if-clause, instead, is structurally similar to Step 2 of Hofer-Szabó and Rédei's proof for the existence of HR-RCCSs of arbitrary finite size in (2006), although Hofer-Szabó and Rédei's conditions (60)–(63) will have to be replaced by:

$$r_i^1 = \frac{c_i d_i}{p(A \wedge B)} \tag{64}$$

$$r_i^2 = \frac{c_i a_i - c_i d_i}{p\left(A \wedge \overline{B}\right)} \tag{65}$$

$$r_i^3 = \frac{c_i b_i - c_i d_i}{p\left(\overline{A} \wedge B\right)} \tag{66}$$

$$r_i^4 = \frac{c_i - c_i a_i - c_i b_i + c_i d_i}{p\left(\overline{A} \wedge \overline{B}\right)}, \tag{67}$$

which, owing to (31), actually reduce to the aforesaid conditions for $\varepsilon = 0$. □

4 Generalised Reichenbachian common cause systems revisited

There are two aspects in which HR-RCCSs may not be considered fully satisfactory generalisations of conjunctive common causes. The first aspect is that they can admit of elements that are probabilistically independent of one or both events from the corresponding correlated pair. This is at odds with the intuition that positive causes should *ceteris paribus* increase the probability of their effects, and that negative causes should *ceteris paribus* decrease their probability. The second aspect is that they rule out the possibility that two distinct causes could equally alter the probability of one, or both, of their effects. On the face of it, there is simply no reason why a systems of common causes should be so constrained. To overcome these limitations, in Mazzola (2012) I proposed a revisitation of HR-RCCSs, along the following lines:

Definition 10 Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$, a *M-Reichenbachian Common Cause System (M-RCCS)* of size $n \geq 2$ for $Corr\,(A, B)$ is a partition $\{C_i\}_{i=1}^n$ of Ω such that:

$$p(C_i) \neq 0 \quad (i = 1, \ldots, n) \tag{68}$$

$$Corr\,(A, B\,|C_i) = 0 \quad (i = 1, \ldots, n) \tag{69}$$

$$[p(A\,|C_i) - p(A)]\,[p(B\,|C_i) - p(B)] > 0 \quad (i = 1, \ldots, n). \tag{70}$$

Whether M-RCCSs are really to be preferred to HR-RCCSs is open to dispute (Stergiou 2015). However, this is no place to settle that issue. Rather, in this section we shall limit ourselves to offer an alternative extension of the generalised common cause model, by taking M-RCCSs as a basis. Let us accordingly define:

Definition 11 Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$, a *Generalised*

M-Reichenbachian Common Cause System (GM-RCCS) of size $n \geq 2$ for $Corr\,(A, B)$ is a partition $\{C_i\}_{i=1}^{n}$ of Ω such that:

$$p\,(C_i) \neq 0 \quad (i = 1, \ldots, n) \tag{71}$$

$$Corr\,(A, B\,|C_i\,) = Corr_e\,(A, B) \quad (i = 1, \ldots, n) \tag{72}$$

$$[p\,(A\,|C_i\,) - p\,(A)]\,[p\,(B\,|C_i\,) - p\,(B)] > 0 \quad (i = 1, \ldots, n). \tag{73}$$

Just as with GHR-RCCSs, it can be shown that GM-RCCSs invariably produce a positive deviation between the observed correlation of their effects and their expected correlation. To this end, let us first introduce the following lemma:

Lemma 6 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. Let $A, B \in \Omega$ and let $\{C_i\}_{i=1}^{n}$ be a partition of Ω satisfying conditions* (71)–(72). *Then:*

$$\delta\,(A, B) = \sum_{i=1}^{n} p\,(C_i)\,[p\,(A\,|C_i\,) - p\,(A)][p\,(B\,|C_i\,) - p\,(B)]. \tag{74}$$

Proof Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. Let $A, B \in \Omega$ and let $\{C_i\}_{i=1}^{n}$ be a partition of Ω for which (71) holds. The theorem of total probability thereby implies that

$$\begin{aligned} Corr\,(A, B) = &\sum_{i=1}^{n} p\,(C_i)\,[p\,(A\,|C_i\,) - p\,(A)][p\,(B\,|C_i\,) - p\,(A)] \\ &+ \sum_{i=1}^{n} p\,(C_i)\,Corr\,(A, B\,|C_i\,). \end{aligned} \tag{75}$$

Let us now suppose that (72) is satisfied, too. Then, owing to the fact that

$$\sum_{i=1}^{n} p\,(C_i) = 1, \tag{76}$$

few elementary calculations would transform the above equality into:

$$\begin{aligned} &Corr\,(A, B) - Corr_e\,(A, B) \\ &\quad = \sum_{i=1}^{n} p\,(C_i)\,[p\,(A\,|C_i\,) - p\,(A)][p\,(B\,|C_i\,) - p\,(A)], \end{aligned} \tag{77}$$

which according to (10) is just a restatement of (74). □

Based on the above lemma, it would then be easy to demonstrate the following proposition:

Proposition 5 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ and any $\{C_i\}_{i \in I} \subseteq \Omega$, if $\{C_i\}_{i \in I}$ is a GM-RCCS of size $n \geq 2$ for $\delta(A, B)$, then (17) obtains.*

GM-RCCSs accordingly perform a similar explanatory function as GHR-RCCSs. Quite interestingly, moreover, for every two events A and B in a classical probability space such that $\delta(A, B) > 0$ and every $n \geq 2$, a GM-RCCSs of size n for $\delta(A, B)$ exists in some extension of the given probability space if and only if a GHR-RCCS does. Demonstrating this will be our next objective.

4.1 Existence of GM-RCCSs

The existential proof we shall elaborate in this section will follow the broad lines of the one developed in Sect. 3.1. Just as with GHR-RCCSs, we shall first determine the necessary and sufficient conditions the values of $p(A|C_1), \ldots, p(A|C_n), p(B|C_1), \ldots, p(B|C_n), p(A \wedge B|C_1), \ldots, p(A \wedge B|C_n)$, and $p(C_1), \ldots, p(C_n)$ ought to satisfy so that the set $\{C_i\}_{i=1}^n$ is a GM-RCCS of size $n \geq 2$ for $\delta(A, B)$. Subsequently, we shall determine the necessary and sufficient conditions for the existence of such numbers, and on that basis we shall finally establish the necessary and sufficient conditions for the existence of an extension of the given probability space, where a GM-RCCS of size n for $\delta(A, B)$ could be found.

Quite evidently, GM-RCCSs differ from GHR-RCCSs only in that they substitute condition (23) with (73). Consequently, in order to complete the first step of our proof, we only need to replace Definition 8 with the following one:

Definition 12 Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ and any $n \geq 2$, a set $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ of real numbers is called *M-admissible* for $\delta(A, B)$ if and only if it is quasi-admissible for $\delta(A, B)$ and it further satisfies

$$[a_i - p(A)][b_i - p(B)] > 0 \qquad (i = 1, \ldots, n). \tag{78}$$

Just as before, the adequacy of the above definition is easily established:

Lemma 7 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ and any $\{C_i\}_{i=1}^n \subseteq \Omega$ where $n \geq 2$, the set $\{C_i\}_{i=1}^n$ is a GM-RCCS of size n for $\delta(A, B)$ if and only if there exists a set $\{a_i, b_i, c_i, d_i\}_{i=1}^n$ of M-admissible numbers for $\delta(A, B)$ for which identities (35)–(38) are true.*

M-admissible numbers are quasi-admissible by definition. This fact allows us to build on our previous discussion, to easily prove the following result:

Lemma 8 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$, no M-admissible numbers exist for $\delta(A, B)$ if*

$$Corr_e(A, B) + p(A)\, p(B) \leq 0. \tag{49}$$

Proof Lemma 8 is demonstrated in a similar way as Lemma 4. Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p, let $A, B \in \Omega$ be arbitrarily chosen so as to satisfy (17), and let us further assume that for some $n \geq 2$, a set $\{a_i, b_i, c_i, d_i\}_{i=1}^{n}$ of M-admissible numbers exists for $\delta(A, B)$. Moreover, let identities (35)–(38) and (44)–(46) be in place.

Showing that some $a_i, b_i \in \{a_i, b_i, c_i, d_i\}_{i=1}^{n}$ exist satisfying

$$a - a_i > 0 \tag{79}$$

$$b - b_i > 0 \tag{80}$$

in this case would only require some elementary calculations, as the above inequalities directly follow from (78) along with (28) and (29). The remainder of the proof would then proceed in exactly the same way as the analogous proof for Lemma 4. □

Condition (62) is thus necessary for the existence of M-admissible numbers for $\delta(A, B)$, for any two events A and B satisfying (17) and any $n \geq 2$. Moreover, as with HR-admissible numbers, it is also sufficient:

Lemma 9 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ satisfying (17), a set $\{a_i, b_i, c_i, d_i\}_{i=1}^{n}$ of M-admissible numbers for $\delta(A, B)$ exists for each $n \geq 2$ if*

$$Corr_e(A, B) + p(A)\, p(B) > 0. \tag{62}$$

Proof Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. Moreover, let $A, B \in \Omega$ satisfy (17) and (62). Proof will proceed by induction on n.

Let $n = 2$ be our inductive basis. Because for $n = 2$ conditions (34) and (78) become equivalent, this case was already covered in the inductive proof for Lemma 5. Next, as our inductive hypothesis, let $n = m$ and let $\{a_i, b_i, c_i, d_i\}_{i=1}^{m}$ be M-admissible for $\delta(A, B)$. On this basis, let us now proceed to the last step of our inductive proof, and let $n = m + (r - 1)$, where $r \geq 2$.

Let us first choose some $c_k \in \{a_i, b_i, c_i, d_i\}_{i=1}^{m}$. Then, (30) and (33) ensure that it is possible to find a set $\left\{c_k^j\right\}_{j=1}^{r}$ of $r \geq 2$ identical real numbers, lying inside the interval $(0, 1)$, such that

$$\sum_{j=1}^{r} c_k^j = \sum_{j=1}^{r} \frac{c_k}{r} = c_k. \tag{81}$$

Furthermore, it is trivially possible to find three sets $\left\{a_k^j\right\}_{j=1}^{r}$, $\left\{b_k^j\right\}_{j=1}^{r}$ and $\left\{d_k^j\right\}_{j=1}^{r}$ of $r \geq 2$ identical real numbers satisfying

$$a_k^j = a_k \qquad j = 1, \ldots, r \tag{82}$$

$$b_k^j = b_k \qquad j = 1, \ldots, r \tag{83}$$

$$d_k^j = d_k \qquad j = 1, \ldots, r. \tag{84}$$

Given (44)–(45), our inductive hypothesis then implies:

$$a = \sum_{i=1}^{m} a_i c_i = \sum_{k \neq i=1}^{m} a_i c_i + a_k c_k = \sum_{k \neq i=1}^{m} a_i c_i + a_k \frac{c_k}{r} r = \sum_{k \neq i=1}^{m} a_i c_i + \sum_{j=1}^{r} a_k^j c_k^j = \sum_{k \neq i=1}^{m+r} a_i c_i \tag{85}$$

$$b = \sum_{i=1}^{m} b_i c_i = \sum_{k \neq i=1}^{m} b_i c_i + b_k c_k = \sum_{k \neq i=1}^{m} b_i c_i + b_k \frac{c_k}{r} r = \sum_{k \neq i=1}^{m} b_i c_i + \sum_{j=1}^{r} b_k^j c_k^j = \sum_{k \neq i=1}^{m+r} b_i c_i \tag{86}$$

$$1 = \sum_{i=1}^{m} c_i = \sum_{k \neq i=1}^{m} c_i + c_k = \sum_{k \neq i=1}^{m} c_i + \sum_{j=1}^{r} c_k^j = \sum_{k \neq i=1}^{m+r} c_i \tag{87}$$

$$\varepsilon = d_k - a_k b_k = d_k^j - a_k^j b_k^j \qquad j = 1, \ldots, r. \tag{88}$$

This guarantees that the set $\{a_i, b_i, c_i, d_i\}_{k \neq i=1}^{m} \cup \{a_j, b_j, c_j, d_j\}_{j=1}^{r}$ of $4(m+r-1)$ numbers so obtained satisfies (28)–(31). Furthermore, (32)–(33) and (78) are clearly satisfied owing to our inductive hypothesis and to the way numbers $\{a_j, b_j, c_j, d_j\}_{j=1}^{r}$ were chosen. This is enough to prove that $\{a_i, b_i, c_i, d_i\}_{k \neq i=1}^{m} \cup \{a_j, b_j, c_j, d_j\}_{j=1}^{r}$ is M-admissible for $\delta(A, B)$, therefore concluding our inductive proof. □

Our existential proof is now virtually complete. Let us just add one final touch:

Proposition 6 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ satisfying (17) and any $n \geq 2$, an extension (Ω', p') of (Ω, p) including a GM-RCCS of size n for $\delta(A, B)$ exists if and only if A and B satisfy (62).*

Proof Proof is in all similar to the proof for Proposition 4, mutatis mutandis. □

Two final remarks may be added at this point. First, Propositions 4 and 6 both rectify the results announced in Mazzola (2013), where it was implicitly assumed that expected correlations should be greater than or equal to zero. This led to the erroneous claim that a generalised common cause should exist, in some extension of the initial probability space, for every positive deviation. Second, GHR-RCCSs and GM-RCCSs for a given deviation may not coexist in the same probability space. Nevertheless, Propositions 4 and 6 jointly guarantee the following result:

Proposition 7 *Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. For any $A, B \in \Omega$ satisfying (17), an extension (Ω', p') of (Ω, p) including a GHR-RCCS of size $n \geq 2$ for $\delta(A, B)$ exists if and only if there is some extension (Ω'', p'') of (Ω, p) including a GM-RCCS of size $m \geq 2$ for $\delta(A, B)$.*

This means that, irrespective of the different probabilistic properties of GHR-RCCSs and GM-RCCSs, neither model can explain more or different deviations than the other. The two models are thus to be assessed based not on what they can explain, but how. The question as to whether M-RCCSs should be preferred to HR-RCCSs, therefore, is carried over to their generalised counterparts.

5 Discussion

The time has come to consider a few objections. These were raised by two anonymous referees, so credit for this section will be shared with them.

The first objection focuses on the notion of expected correlation. Back in Sect. 2, we formulated the generalised principle of the common cause as the proposition that the role of common causes is to explain statistically significant deviations between the estimated value of a correlation and its *expected* value. One may wonder whether this could make the principle sound subjective, in contrast with the metaphysical status that is usually conferred on it. Relatedly, as suggested by one of the referees, one may wish to rephrase the extended principle in purely objectivist terms: conditioning on the common cause should restore not the expected correlation between two events, but rather the correlation due the (possibly null) direct causal dependence between them.

Restating the principle along these lines, however, would be neither necessary nor effective. The notion of expectation here employed, in fact, is deliberately polysemic, allowing for objective (e.g. statistical) as well as for subjective interpretations. This polysemy does not prevent a metaphysical interpretation of the principle, but at the same time it makes room for an epistemological interpretation: in this sense, the principle may be seen as a rule to preserve subjective beliefs about probabilistic dependences in the face of recalcitrant statistical evidence. Epistemological interpretations of this kind have been proposed, for instance, by Fraassen (1982) and Sober (1984).

Notice, moreover, that reformulating the extended principle in purely causal language in the way suggested by the referee would immediately expose it to two major problems. To start with, the principle could not make room for probabilistic correlations resulting from logical or mathematical relations. To wit, certainly there is a correlation between the number of sides of a regular plane figure and the number of its internal angles, but it is equally evident that said correlation is the result of neither a direct causal connection between the two numbers, nor of any common cause. Introducing the notion of expected correlation easily avoids this problem, as argued in Sect. 2. Secondarily, if reformulated in the way suggested, the extended principle would become moot unless it were supplemented with some criterion to tell how much of the correlation between two events is due to their causal interaction, and how much of it needs to be explained by some common cause instead. On the face of it, there seems to be no other way to do this than looking at the population mean value of the correlation between events of the same type. Pushed out of the door, the notion of expected correlation would then get back through the window.

The second objection that we shall address questions the very significance of our investigation. Such misgivings are based on the following considerations. Conditions such as (20) and (70) from HR-RCCS and M-RCCS models are intended to ensure that the effects of a common cause should be positively correlated. Nevertheless, the objection goes, they are not strictly needed to explain the positive correlation between the effects of a common cause: *given* that these are positively correlated, it is immaterial whether (20) and (70) actually hold. Rather, all is required for the explanation is that their correlation should disappear when conditioning on the common cause, i.e. the screening-off condition (19), (69). On the other hand the existence of a partition satisfying (19), (69) is a trivial matter, for some such partition can always be constructed. The question whether a common cause system of any finite or countable size exists for a given correlated pair is thus, in one of the referee's words, a 'pseudo-question'. By extension, so too must be the question whether a generalised common cause system exists for each positive deviation.

Luckily, there are at least two reasons why the above misgivings are unwarranted. The first reason is that, in reality, conditions (20) and (70) do perform an important explanatory function. They make it possible to *infer* the positive correlation between two events from the existence of their common cause. Hence, they ensure that positive correlations get explained in accordance with the classical covering-law model (Hempel 1965). Of course the general applicability of the covering-law model has long been put into question, but in and of itself this is no reason to deny the explanatory significance of (20) and (70). Similar considerations also apply to conditions (22) and (72).

The second reason is that, in any event, the existence of a generalised common cause system is no trivial matter. More exactly, it can be demonstrated that a partition satisfying (22), (72) may not exist for a deviation $\delta(A, B)$ if $Corr_e(A, B) + p(A)\,p(B) < 0$. The details of this proof, however, are beyond the scope of this paper.

6 Conclusion

The principle of the common cause decrees that improbable coincidences ought to be put down to the action of some common cause. The standard interpretation of the principle takes this as a requirement that positive correlations between causally unrelated events should be removed by conditioning on some conjunctive common cause. The interpretation here promoted, and encapsulated in the extended principle of the common cause, urges by contrast that common causes should be called for in order to explain positive deviations between the estimated correlation of two events and their expected correlation. This paper has outlined two distinct probabilistic models for systems of common causes that incorporate the extended interpretation of the principle. GHR-RCCSs have been elaborated by combining the generalised common cause model with HR-RCCSs. GM-RCCSs, instead, have been obtained by integrating generalised common causes with M-RCCSs. The necessary and sufficient conditions for the existence of finite systems of either kind in classical probability spaces have been determined.

Our demonstration led to the unexpected result that some extension of the given classical probability space can be found including a GHR-RCCS of arbitrary finite size for some specified positive deviation, if and only if a similar extension can be found including a GM-RCCS of finite size for the same deviation. Even more interestingly, in either case the existence of such space is guaranteed if and only if the sum of the expected correlation of the pair of events under consideration and the product of their probabilities is greater than zero. The mathematical reason for this limitation is clear: only under said constraint, in fact, can HR-admissible numbers and M-admissible numbers for a positive deviation exist. The philosophical interpretation of this result, instead, is an open question.

Acknowledgements I am grateful to two anonymous referees for their challenging but constructive feedback.

Appendix: Proof of Lemma 5

Let (Ω, p) be a classical probability space with σ-algebra of random events Ω and probability measure p. Moreover, let $A, B \in \Omega$ satisfy (17) and (62).

To start with, let us observe that (43) is in fact a system of $n-1$ equations, namely one for each value of $i = 1, \ldots, n-1$. Therefore, (39)–(43) jointly comprise a system of $4+(n-1) = n+3$ equations in $4n$ variables. This means that each HR-admissible set for $\delta(A, B)$ is determined by a set of $4n - (n+3) = 3n - 3$ parameters, for every $n \geq 2$. To establish the existence of such set, we accordingly need to prove that such parameters exist. To this purpose, let numbers a, b and ε be understood as per (44)–(46). Proof will proceed by induction on n.

Let us begin by assuming $n = 2$ as our inductive basis. This has the effect of transforming (39)–(43), respectively, into:

$$a_2 = \frac{a - c_1 a_1}{1 - c_1} \tag{89}$$

$$b_2 = \frac{b - c_1 b_1}{1 - c_1} \tag{90}$$

$$c_2 = 1 - c_1 \tag{91}$$

$$d_2 = \varepsilon + \frac{[a - a_1 c_1][b - b_1 c_1]}{[1 - c_1]^2} \tag{92}$$

$$d_1 - a_1 b_1 = \varepsilon. \tag{93}$$

Since a, b and ε are known by hypothesis, choosing numbers c_1, a_1 and b_1 will therefore suffice to fix the values of all $4n = 8$ variables in the system. In particular, (89)–(92) immediately produce:

$$lim_{c_1 \to 0}\, a_2 = a \tag{94}$$

$$lim_{c_1 \to 0}\, b_2 = b \tag{95}$$

$$lim_{c_1 \to 0}\, c_2 = 1 \tag{96}$$

$$lim_{c_1 \to 0}\, d_2 = \varepsilon + ab, \tag{97}$$

while on the other hand (17) directly requires that

$$1 > a, b > 0, \tag{98}$$

as it would be easy to verify. Taken together this ensures that, as c_1 is taken sufficiently close to zero:

$$1 \geq a_2, b_2 \geq 0 \tag{99}$$
$$1 \geq c_1, c_2 \geq 0, \tag{100}$$

while (62) and (17) imply that

$$1 \geq d_2 \geq 0. \tag{101}$$

To determine the remaining numbers, we further need to set a_1 and b_1. In this case, our choice will depend on the value of ε, as follows:

$$\varepsilon \geq 0 \qquad \begin{cases} a > a_1 \geq 0 \\ b > b_1 \geq 0 \end{cases} \tag{102}$$
$$\varepsilon < 0 \qquad \begin{cases} 1 \geq a_1 > a \\ 1 \geq b_1 > b \end{cases} \tag{103}$$

Either option is allowed by (98), and either will ensure that

$$1 \geq d_i \geq 0, \tag{104}$$

as it would be straightforward to check with the aid of (43), (17) and (62). Thanks to Lemma 3, this is enough to establish that some set $\{a_i, b_i, c_i, d_i\}_{i=1}^{n}$ of quasi-admissible numbers exist for $\delta(A, B)$ if $n = 2$. To further show that such set is HR-admissible for $\delta(A, B)$, we only need to observe that (34) can be obtained from both (102) and (103), owing to (28)–(30).

Let us now assume, as our inductive hypothesis, that some set $\{a_i, b_i, c_i, d_i\}_{i=1}^{m}$ is HR-admissible for $\delta(A, B)$, where $n = m > 2$. To prove that a HR-admissible set for $\delta(A, B)$ also exists if $n = m + 1$, let us consider the set

$$\left\{a_j, bj, c_j, a_{m-1}, b_{m-1}\right\}_{j=1}^{m-2} \subset \{a_i, b_i, c_i, d_i\}_{i=1}^{m},$$

and let us choose numbers $a'_m, b'_m, c'_{m-1}, c'_m$ such that:

$$a_j > a'_m > 0 \qquad (j = 1, \ldots, m-1) \tag{105}$$
$$b_j > b'_m > 0 \qquad (j = 1, \ldots, m-1) \tag{106}$$
$$c_m, c_{m-1} > c'_{m-1} > 0 \tag{107}$$
$$1 > c'_m > 0. \tag{108}$$

Given (39)–(43), the set

$$\left\{a_j, bj, c_j, a_{m-1}, b_{m-1}\right\}_{j=1}^{m-2} \cup \left\{a'_m, b'_m, c'_{m-1}, c'_m\right\}$$

of $3n-3$ parameters will then suffice to determine $4(m+1)$ numbers:

$$\left\{a_j, b_j, c_j, d_j, a_{m-1}, b_{m-1}, c'_{m-1}, d_{m-1}, a'_m, b'_m, c'_m, d'_m, a_{m+1}, b_{m+1}, c_{m+1}, d_{m+1}\right\}_{j=1}^{m-2}.$$

Because numbers $\left\{a_j, b_j, c_j, d_j, a_{m-1}, b_{m-1}, c'_{m-1}, d_{m-1}, a'_m, b'_m, c'_m,\right\}_{j=1}^{m-1}$ satisfy conditions (31)–(33) and (34) by hypothesis, all we need to show is that said constraints are also satisfied by the remaining numbers $\left\{d'_m, a_{m+1}, b_{m+1}, c_{m+1}, d_{m+1}\right\}$. To this purpose, let us first notice that (32) must be true of d'_m by virtue of (43) and (105)–(106). Next, thanks to (39)–(42), it will be sufficient to suppose that

$$a'_m \rightarrow 0 \tag{109}$$
$$b'_m \rightarrow 0 \tag{110}$$
$$c'_m \rightarrow c_{m-1} - c'_{m-1} \tag{111}$$

to obtain

$$lim_{a'_m \rightarrow 0,\ c'_m \rightarrow c_{m-1}-c'_{m-1}} a_{m+1} = a_m \tag{112}$$
$$lim_{b'_m \rightarrow 0,\ c'_m \rightarrow c_{m-1}-c'_{m-1}} b_{m+1} = b_m \tag{113}$$
$$lim_{c'_m \rightarrow c_{m-1}-c'_{m-1}} c_{m+1} = c_m \tag{114}$$
$$lim_{a'_m \rightarrow 0,\ b'_m \rightarrow 0,\ c'_m \rightarrow c_{m-1}-c'_{m-1}} d_{m+1} = d_m \tag{115}$$

which we already know, by our inductive hypothesis, to satisfy (31)–(33). Moreover, (105) and (106), along with the inductive assumption whereby

$$[a_m - a_i][b_m - b_i] > 0 \qquad (i = 1, \ldots, m-1), \tag{116}$$

ensures that

$$[a_{m+1} - a_i][b_{m+1} - b_i] > 0 \qquad (i = 1, \ldots, m), \tag{117}$$

which together with our inductive hypothesis suffices to establish (34). Due to Lemma 3 and Definition 8, the set of $4(m+1)$ numbers so determined is therefore HR-admissible for $\delta(A, B)$. Lemma 5 is thus demonstrated by induction.

References

Arntzenius, F. (1992). The common cause principle. In *PSA: Proceedings of the biennial meeting of the philosophy of science association. Volume Two: Symposia and invited papers* (pp. 227–237).

Cartwright, N. (1979). Causal laws and effective strategies. *Noûs*, *13*, 419–437.

Cartwright, N. (1988). How to tell a common cause: Generalizations of the common cause criterion. In J. H. Fetzer (Ed.), *Probability and causality* (pp. 181–188). Dordrecht: Reidel.

Cartwright, N. (1999). Causal diversity and the Markov condition. *Synthese*, *121*, 3–27.

Eells, E. (1991). *Probabilistic causality*. Cambridge: Cambridge University Press.

Good, I. J. (1961). A causal calculus I. *British Journal for the Philosophy of Science*, *11*, 305–318.

Graßhoff, G., Portmann, S., & Wuthrich, A. (2006). Minimal assumption derivation of a Bell-type inequality. *The British Journal for the Philosophy of Science*, *56*, 663–680.

Gyenis, Z., & Rédei, M. (2014). Atomicity and causal completeness. *Erkenntnis*, *79*, 437–451.

Gyenis, Z., & Rédei, M. (2016). Common cause completability of non-classical probability spaces. *Belgrade Philosophical Annual*, *29*, 15–32.

Hempel, C. (1965). *Aspects of scientific explanation and other essays in philosophy of science*. New York: Free Press.

Hofer-Szabó, G., & Rédei, M. (2004). Reichenbachian common cause systems. *International Journal of Theoretical Physics*, *43*, 1819–1826.

Hofer-Szabó, G., & Rédei, M. (2006). Reichenbachian common cause systems of arbitrary finite size exist. *Foundations of Physics*, *36*, 745–756.

Hofer-Szabó, G., Rédei, M., & Szabó, L. E. (2013). *The principle of the common cause*. Cambridge: Cambridge University Press.

Kitajima, Y. (2008). Reichenbach's common cause in an atomless and complete orthomodular lattice. *International Journal of Theoretical Physics*, *47*, 511–519.

Kitajima, Y., & Rédei, M. (2015). Characterizing common cause closedness of quantum probability theories. *Studies in the History and Philosophy of Modern Physics*, *52*(B), 234–241.

Marczyk, M., & Wroński, L. (2015). Completion of the causal completability problem. *British Journal for the Philosophy of Science*, *66*, 307–326.

Mazzola, C. (2012). Reichenbachian common cause systems revisited. *Foundations of Physics*, *42*, 512–523.

Mazzola, C. (2013). Correlations, deviations and expectations: The extended principle of the common cause. *Synthese*, *190*, 2853–2866.

Mazzola, C., & Evans, P. (2017). Do Reichenbachian common cause systems of arbitrary finite size exist? *Foundations of Physics*, *47*, 1543–1558.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Mateo: Morgan Kauffman.

Pearl, J. (2000). *Causality*. Cambridge: Cambridge University Press.

Reichenbach, H. (1956). *The direction of time*. Berkeley: University of California Press.

Salmon, W. C. (1971). *Statistical explanation and statistical relevance*. Pittsburgh: Pittsburgh University Press.

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.

Skyrms, B. (1980). *Causal necessity*. New Haven: Yale University Press.

Sober, E. (1984). Common cause explanation. *Philosophy of Science*, *5*, 212–241.

Spirtes, P., Glymour, C., & Scheines, R. (2001). *Causation, prediction, and search* (2nd ed.). Cambridge: MIT Press.

Stergiou, C. (2015). Explaining correlations by partitions. *Foundations of Physics*, *45*, 1599–1612.

Suppes, P. (1970). *A probabilistic theory of causality*. Amsterdam: North-Holland Publishing Company.

Suppes, P., & Zaniotti, M. (1981). When are probabilistic explanations possible? *Synthese*, *48*, 191–199.

van Fraassen, B. (1980). *The scientific image*. Oxford: Clarendon Press.

van Fraassen, B. (1982). Rational belief and the common cause principle. In R. McLaughlin (Ed.), *What? Where? When? Why?* (pp. 193–209). Dordrecht: Reidel.

van Fraassen, B. (1991). *Quantum mechanics: An empiricist view*. Oxford: Clarendon Press.

Williamson, J. (2005). *Bayesian nets and causality*. Oxford: Oxford University Press.

Wroński, L., & Marczyk, M. (2010). Only countable Reichenbachian common cause systems exist. *Foundations of Physics*, *40*, 1155–1160.