

University of Groningen

## A new proposal how to handle counterexamples to Markov causation à la Cartwright, or: fixing the chemical factory

Gebharter, Alexander; Retzlaff, Nina

*Published in:*  
Synthese

*DOI:*  
[10.1007/s11229-018-02014-7](https://doi.org/10.1007/s11229-018-02014-7)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2020

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Gebharter, A., & Retzlaff, N. (2020). A new proposal how to handle counterexamples to Markov causation à la Cartwright, or: fixing the chemical factory. *Synthese*, 197(4), 1467-1486. <https://doi.org/10.1007/s11229-018-02014-7>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



# A new proposal how to handle counterexamples to Markov causation à la Cartwright, or: fixing the chemical factory

Alexander Gebharter<sup>1</sup> · Nina Retzlaff<sup>2</sup>

Received: 24 January 2018 / Accepted: 7 November 2018 / Published online: 16 November 2018  
© The Author(s) 2018

## Abstract

Cartwright (Synthese 121(1/2):3–27, 1999a; The dappled world, Cambridge University Press, Cambridge, 1999b) attacked the view that causal relations conform to the Markov condition by providing a counterexample in which a common cause does not screen off its effects: the prominent chemical factory. In this paper we suggest a new way to handle counterexamples to Markov causation such as the chemical factory. We argue that Cartwright’s as well as similar scenarios (such as decay processes, EPR/B experiments, or spontaneous macro breaking processes) feature a certain kind of non-causal dependence that kicks in once the common cause occurs. We then develop a representation of this specific kind of non-causal dependence that allows for modeling the problematic scenarios in such a way that the Markov condition is not violated anymore.

**Keywords** Causation · Markov condition · Causal modeling · Screening off · Cartwright

## 1 Introduction

From an empirical point of view, Bayes nets provide one of the most promising approaches to causation currently on the market. They can be used for formulating and testing causal hypotheses, for learning causal structure on the basis of statistical and/or experimental data, and for predicting the outcomes of possible interventions, even if only purely observational data is available (Pearl 2000; Spirtes et al. 1993). Another advantage over competing philosophical approaches to causation consists in the framework’s closeness to successful empirical theories of the sciences: Character-

---

✉ Alexander Gebharter  
alexander.gebharter@gmail.com

<sup>1</sup> Department of Theoretical Philosophy, University of Groningen, Oude Boteringestraat 52, 9712 GL Groningen, The Netherlands

<sup>2</sup> Düesseldorf Center for Logic and Philosophy of Science, University of Düesseldorf, Universitätsstraße 1, 40225 Düesseldorf, Germany

izing causation in terms of Bayes nets can be backed up by an inference to the best explanation of certain empirical phenomena and the theory as a whole can be tested on empirical grounds (Gebharder 2017b; Schurz and Gebharder 2016).

However, there seem to be problems with the Bayes net framework as a general theory of causation. One of the most prominent has been put forward by Nancy Cartwright in several works (see, e.g., Cartwright 1999a, b, 2007). Cartwright has constructed a scenario in which a chemical factory produces a substance with a certain probability. When this chemical is produced, the factory also produces a pollutant as a byproduct. Cartwright assumes that the chemical factory is the only common cause of the product and the byproduct and that neither the product is a cause of the byproduct nor the other way around. So constructed, the scenario violates the core principle of the Bayes net approach to causation, the Markov condition (MC), by implying a dependence between the chemical and the pollutant (conditional on their common cause) that is excluded by MC. Similar scenarios in which a common cause does not screen off its effects can be found in the micro as well as in the macro realm. Prominent micro examples are quantum experiments (cf. Hausman 1998; Healey 2009; Glymour 2006; Näger 2016; Retzlaff 2017). An everyday life example would be the breaking of a stone in two pieces (cf Schurz 2017).

There are two kinds of possible reactions to counterexamples to Markov causation<sup>1</sup> such as Cartwright's (1999a, b) from supporters of Bayes net methods: The first possible way to respond consists in claiming that there are no such counterexamples in the actual world (see, e.g., Glymour 1999). Markov violations in purported counterexamples arise only because the causal structure underlying these scenarios has been misrepresented in our models or variables have not been chosen correctly. Supporters of this strategy insist that to adequately represent the chemical factory (or similar scenarios), one needs to replace the original variables by more fine-grained variables or to modify the assumed causal structure, for example, by adding latent common causes or missing intermediate causes (see, e.g., Hausman and Woodward 1999; Pearl 2000; Spirtes et al. 1993). The second possible way to go consists in accepting the choice of variables and the causal structure of the purported counterexamples. In that case, the only option left for supporters of Markov causation seems to consist in modifying the framework. Such approaches have recently been put forward by Schurz (2017) and Näger (2013) who propose to weaken MC.

In this paper we explore yet another way to go when one takes counterexamples to Markov causation seriously. Instead of modifying the causal structure, changing the variables, or weakening MC, we propose to add a component to our models that has been largely ignored so far. One typical precondition for successful causal modeling is that the systems of interest do not feature variables standing in other than causal relations (cf. Woodward 2015). We argue that Cartwright's (1999a, b) chemical factory and similar scenarios violate MC because they do not meet this precondition. These scenarios involve a kind of non-causal dependence that kicks in once the common cause occurs. We argue that this non-causal dependence arises due to background assumptions which rule how quantities, properties, or parts are distributed among different objects or places if the common cause occurs. We then develop a method for

---

<sup>1</sup> Throughout the paper we loosely refer to causal relations that conform to MC as Markov causation.

representing this kind of non-causal dependence in such a way that MC is not violated anymore and highlight several possible advantages of our approach.

The paper is structured as follows: in Sect. 2 we introduce the basics of the Bayes net framework and its causal interpretation. In Sect. 3 we present Cartwright's (1999a, b) counterexample to Markov causation as well as three other structurally similar scenarios: a decay process, an EPR/B experiment, and an eroding stone. These four scenarios shall stand proxy for all kinds of such scenarios from the macro or micro realm. In Sect. 4 we very briefly discuss standard reactions to the problematic scenarios and Schurz' (2017) recent approach. In Sect. 5 we then develop our own approach to handle counterexamples to Markov causation. We conclude in Sect. 6.

## 2 Bayes nets and their causal interpretation

Bayes nets were originally developed to graphically store independence information and to simplify reasoning under uncertainty (Pearl 1988). A Bayes net is a triple  $\langle V, E, P \rangle$  that satisfies the Markov condition (MC), where  $V$  is a set of random variables,  $E$  is a set of edges connecting pairs of variables in  $V$ , and  $P$  is a probability distribution over  $V$ . Throughout the paper we will use the following (global) version of MC (cf. Schurz and Gebharter 2016, p. 1084):<sup>2</sup>

**Definition 2.1** (*Markov condition*) A graph  $G = \langle V, E \rangle$  and a probability distribution  $P$  satisfy the (global) Markov condition iff it holds for all  $X, Y \in V$  and  $Z \subseteq V \setminus \{X, Y\}$  that  $X$  and  $Y$  are  $d$ -connected given  $Z$  in  $G$  if  $X$  and  $Y$  are probabilistically dependent conditional on  $Z$  in  $P$ .

**Definition 2.2** ( *$d$ -connection/ $d$ -separation*) Variables  $X \in V$  and  $Y \in V$  are  $d$ -connected given a set  $Z \subseteq V \setminus \{X, Y\}$  in a graph  $G = \langle V, E \rangle$  iff  $X$  and  $Y$  are connected by a path  $\pi$  such that

- (i) no non-collider  $C \in V$  on  $\pi$  is in  $Z$ , and
- (ii) every collider  $C \in V$  on  $\pi$  is in  $Z$  or there is a  $C' \in Z$  such that  $C \longrightarrow \dots \longrightarrow C'$  is part of  $G$ .

$X \in V$  and  $Y \in V$  are  $d$ -separated by  $Z \subseteq V \setminus \{X, Y\}$  in  $G$  iff  $X$  and  $Y$  are not  $d$ -connected given  $Z$  in  $G$ .

A path  $\pi$  between two variables  $X$  and  $Y$  is a chain of edges connecting  $X$  and  $Y$ . No variable  $Z$  (different from  $X$  and  $Y$ ) is allowed to appear more often than once on  $\pi$ . An edge  $X \longrightarrow Y$  is called a directed edge from  $X$  to  $Y$ , and a path of the form  $X \longrightarrow \dots \longrightarrow Y$  is called a directed path from  $X$  to  $Y$ . Finally, a collider  $X$  on a path  $\pi$  is a variable on  $\pi$  such that the edges connecting  $X$  with its neighbors on  $\pi$  both feature an arrowhead pointing at  $X$ .

Note that a Bayes net's edges might lack any realistic interpretation and just represent probabilistic independence patterns. They might, however, also represent all kinds of relations that conform to MC. Examples for such relations are causation, superven-

<sup>2</sup> For the more established (but also more limited) local version of MC see, for example, (Spirtes et al. 1993, p. 33).

nience and constitution (see, e.g., Gebharder 2017a,c), but also the dependence of different pieces of evidence on a corresponding hypothesis (see, e.g., Sprenger and Hartmann in press). The causal interpretation of Bayes nets will be especially relevant for the present paper. It was developed by Clark Glymour and his students around 1990 (Spirtes et al. 1993) and later by Pearl (2000), but the idea of connecting causal structure to probabilistic independence patterns is already present in Reichenbach's (1956) work. Glymour, Spirtes, and Scheines (1991, p. 151), for example, put Reichenbach's insights this way:

**Screening Off:** If A causes C only through the mediation of a set of variables B, then A and C are statistically independent conditional on B.

**Common Cause:** If A does not cause B and B does not cause A, and A and B are statistically dependent, then there exists a common cause of A and B.

The causal interpretation of Bayes nets basically covers and advances Reichenbach's (1956) insights. But before we will go into details, let us briefly introduce a few more terminological conventions. In a causally interpreted Bayes net  $\langle V, E, P \rangle$ , the variables in  $V$  describe properties or events, and the directed edges in  $E$  typically represent direct causal connections between variables in  $V$ , meaning that  $X \rightarrow Y$  is interpreted as  $X$  being a direct cause of  $Y$  (w.r.t.  $V$ ). If  $\pi$  is a directed path from  $X$  to  $Y$ , then  $X$  is assumed to be a (direct or indirect) cause of  $Y$ . A variable  $Z$  (different from  $X$  and  $Y$ ) lying on such a directed path  $\pi$  is called an intermediate cause. Finally, a path of the form  $X \leftarrow \dots \leftarrow Z \rightarrow \dots \rightarrow Y$  is called a common cause path with  $Z$  being a common cause of  $X$  and  $Y$ .

If a Bayes net is causally interpreted, then the directed edges of its graph  $G$  are intended to provide information about the causal structure of the system of interest, and the probability distribution  $P$  to provide information about the strengths of the influences propagated over this structure. If all common causes of variables in  $V$  are included in  $V$  and the variables in  $V$  do not stand in other than causal relations, then it is assumed that every dependence among variables in  $V$  can be explained by some causal path  $d$ -connecting these variables. This is often referred to as the causal Markov assumption in the literature. If it holds for a variable set  $V$ , then the corresponding Bayes net can account for and explain every probabilistic dependence among variables in  $V$  in purely causal terms.

Assuming that for every system there is a (maybe larger) structure satisfying MC reflects a deep-seated metaphysical principle which will play a major role later on: Dependencies do not occur randomly. They are produced by their underlying structure. If we have reasons to exclude non-causal relations among variables in  $V$ , then every dependence between variables in  $V$  must be due to some causal path. And, vice versa, there is no dependence without structure, meaning that whenever we block all paths between two variables  $X$  and  $Y$  (by conditionalizing on a set  $Z \subseteq V \setminus \{X, Y\}$   $d$ -separating  $X$  and  $Y$ ), then any dependence between  $X$  and  $Y$  will vanish. As a consequence, explaining a dependence between  $X$  and  $Y$  in case of a common cause structure  $X \leftarrow \dots \leftarrow Z \rightarrow \dots \rightarrow Y$  (without other causal or non-causal connections between  $X$  and  $Y$  around) amounts to the fact that conditionalizing on  $Z$  screens  $X$  and  $Y$  off each other (since all paths causally connecting  $X$  and  $Y$  are blocked).

### 3 Cartwright's chemical factory and other counterexamples

In this section we will introduce Cartwright's (1999a, b) chemical factory as well as the three other counterexamples to Markov causation mentioned in Sect. 1: a decay process, an EPR/B experiment, and an eroding stone. They shall stand proxy for all counterexamples structurally similar to Cartwright's from the macro or micro realm. Though they are to some extent controversial, we hope that almost everyone will find at least one of them convincing or at least alarming.

#### 3.1 Chemical factory

Assume that there is a chemical factory that produces a certain substance. Unfortunately, the chemical process involved in producing the target substance is not perfect: The factory does not always succeed in producing this substance. On average, it succeeds only in 8 of 10 attempts. Whenever it succeeds, however, it also produces a nasty pollutant. The target chemical and the pollutant can only be produced together. They are output in two different storage tanks.

Here comes our reconstruction of the chemical factory scenario in terms of Bayes nets. We represent the chemical factory by the binary variable *Chem*, where *Chem* = 1 means that the chemical factory is active and *Chem* = 0 means that it is inactive. We represent the target substance by the binary variable *Sub* and the pollutant by the binary variable *Poll*. *Sub* = 1 means that the target substance occurs in its storage tank, *Sub* = 0 means that it does not. *Poll* = 1 means that the pollutant occurs in its storage tank, while *Poll* = 0 means that it does not. The causal structure underlying the system is  $Sub \leftarrow Chem \rightarrow Poll$ . Due to the assumptions made above, the system's associated probability distribution features the following conditional probabilities:

$$P(Sub = 1|Chem = 1) = P(Poll = 1|Chem = 1) = 0.8$$

$$P(Sub = 1|Chem = 1, Poll = 1) = P(Poll = 1|Chem = 1, Sub = 1) \approx 1$$

MC together with  $Sub \leftarrow Chem \rightarrow Poll$  implies that *Sub* and *Poll* are independent conditional on *Chem*. But this implies that  $P(Sub = 1|Poll = 1, Chem = 1)$  has to equal  $P(Sub = 1|Chem = 1)$ . From the probabilities specified above, however, it follows that these two conditional probabilities cannot be equal. As a consequence, the chemical factory clearly violates MC.<sup>3</sup>

<sup>3</sup> Glymour (1999) asks the reader to "consider where this factory is" (p. 72). He answers this rhetorical question with "nowhere" (ibid.) and argues that presumably nothing can be inferred from fictional examples such as the chemical factory. However, we think that the chemical factory can still be seen as a rather innocuous attempt to illustrate in a simple and easily accessible way that macro-level common causes not screening off their effects are a serious possibility a general theory of causation should be able to handle. No scientific evidence so far excludes the existence of such common causes. We hope that readers nevertheless tempted by Glymour's argumentation might find one of the other macro-level counterexamples (such as the eroding stone discussed below or the rich uncle scenario described in Sect. 5) more appealing.

### 3.2 Decay process

Assume that a particle—for example, a heavy isotope—is present in a certain region of space. After a while it might spontaneously decay in two subparticles of the same kind in such a way that the two subparticles move away from each other in opposite directions. This would clearly be an indeterministic process with a common cause structure. There seems to be no other causal connection between the two subparticles' behaviors. Now assume that we are interested in whether a subparticle occurs in one of two specific regions, one is at the left hand side of the decaying particle, the other one is at the right hand side. When the particle decays, there is a certain chance that its subparticles will move away from its original position in such a way that they end up in our two regions of interest. However, if the particle has decayed and there is a subparticle in one of the two regions, then (with a probability of almost 1) there is also a subparticle of the same kind present in the other region.

The common cause involved in this spontaneous process is modeled by the binary variable *Particle*.  $Particle = 1$  means that a particle of a certain kind is present in a certain region of space  $r_0$ , and  $Particle = 0$  that no such particle is present in that region. The two effects are modeled by the binary variables *SubP1* and *SubP2*.  $SubP1 = 1$  means that a subparticle occurs in another region of space  $r_1$ ,  $SubP1 = 0$  that no such particle occurs in that region.  $SubP2 = 1$  means that a subparticle occurs in yet another region of space  $r_2$ ,  $SubP2 = 0$  that no such particle occurs in that region. Note that the two regions  $r_1$  and  $r_2$  must be chosen in such a way that they lie in opposite directions of  $r_0$ . MC applied to the causal structure  $SubP1 \leftarrow Particle \rightarrow SubP2$  then implies that *Particle* screens *SubP1* and *SubP2* off each other. However, this is not the case: From the fact that a particle of a certain kind was present in a certain space region nothing follows about whether it decayed, nor in which directions the subparticles have moved if it decayed. However, the subparticles must have moved in opposite directions if the particle decayed. Thus, if the particle decayed and there is a subparticle in one region—say in  $r_1$ —the probability that there is, at the same time, a subparticle of the same kind present in  $r_2$  will be almost 1. But this just means, contrary to what is implied by MC, that *SubP1* and *SubP2* are dependent given  $Particle = 1$ .

### 3.3 EPR/B experiment

Assume that two photons are emitted in an entangled quantum state in opposite directions from a source. On the left hand side as well as on the right hand side of the source are polarizers and behind each polarizer is a detector with a light bulb. We assume that the measurement settings of the two polarizers can be chosen at will (by the experimenter). If a photon passes one of the polarizers, it is polarized according to the setting of that polarizer. The photon then moves on to the detector and the light bulb on that detector goes on. If the photon, on the other hand, is absorbed by the polarizer, then no photon is detected and, accordingly, the bulb on the corresponding detector does not go on. If we know nothing about the settings of the two polarizers, then we cannot infer anything about the behavior of one of the detectors by observing how the other detector behaves, even if we know the quantum state. However, if we know the

quantum state and that the settings of both polarizers are identical, then we can infer whether the light bulb on one of the detectors is on or off by observing whether the bulb on the other detector is on or off. If the quantum state is, for example, the Bell state  $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B)$ , then the light on one of the detectors will be on if and only if the light on the other detector will be on as well.

We represent the quantum state by a variable  $QS$ . The settings of the two polarizers are represented by variables  $Pol1$  and  $Pol2$ , and the detectors' behaviors are modeled by the binary variables  $Det1$  and  $Det2$  (with the possible values *on* and *off* standing for whether the light bulb on the respective detector is on or off). Though it is still controversial which causal structure is the true one underlying EPR/B experiments (cf. Näger 2016; Wood and Spekkens 2015), the intuitively most plausible structure seems to be  $Pol1 \rightarrow Det1 \leftarrow QS \rightarrow Det2 \leftarrow Pol2$  (cf. Glymour 2006). Whether the bulb on a detector is on or off causally depends on the quantum state and the setting of its respective polarizer. There seem to be no other causal influences around. Assuming that the causal structure  $Pol1 \rightarrow Det1 \leftarrow QS \rightarrow Det2 \leftarrow Pol2$  is the correct one, the EPR/B experiment clearly violates MC, which implies that  $Det1$  is independent of  $Det2$  given  $\{QS, Pol1, Pol2\}$ . As we have seen above, however, there is at least one setting in the EPR/B experiment in which  $Det1$  and  $Det2$  depend on each other conditional on  $\{QS, Pol1, Pol2\}$ .

### 3.4 Breaking stone

The breaking stone counterexample is inspired by (Schurz 2017).<sup>4</sup> Assume that there is a stone (with a certain mass  $m > 0$ ) on the top of a mountain or not. If there is such a stone on the mountain's top, we assume that it slowly erodes and, after many years, it might spontaneously break in two parts. If it breaks, then one part falls down on the one side of the mountain, while the other one falls down on the opposite side. The presence of the eroding stone with mass  $m$  at the mountain's top would be the common cause of one part with a certain mass  $m_1$  falling down on the one side of the mountain and the other part with a certain mass  $m_2$  falling down on the opposite side. There would be no other causal connection among these events.

We model the common cause in such a way that it involves a spontaneous breaking event. We do so by means of the variable  $Stone$ .  $Stone = 0$  means that no stone is present at the mountain's top.  $Stone = m$  means that there is a stone with mass  $m$  (with  $m > 0$ ) present at the top of the mountain. The variable  $Part1$  describes what is going on on one side of the mountain.  $Part1 = 0$  means that no stone falls down on this side of the mountain.  $Part1 = m_1$  (with  $m_1 > 0$ ) means that a stone falls down on this side and has the mass  $m_1$ . The variable  $Part2$  describes what is going on on the other side of the mountain.  $Part2 = 0$  means that no stone falls down on the other side of the mountain.  $Part2 = m_2$  (with  $m_2 > 0$ ) means that a stone falls down on that side and has the mass  $m_2$ . The causal structure underlying the breaking stone scenario is  $Part1 \leftarrow Stone \rightarrow Part2$ . Which values  $Part1$  and  $Part2$  take if  $Stone = m$  is not fully determined by  $Stone = m$ . According to MC, the causal

<sup>4</sup> Examples violating MC in a similar way can be constructed for all kinds of macro objects that can spontaneously break apart.



structure  $Part1 \leftarrow Stone \longrightarrow Part2$  implies that  $Part1$  and  $Part2$  are screened off by  $Stone$ . But if  $Stone = m$ , then  $Part1$  and  $Part2$  are clearly dependent. The reason for this is simple: If a stone with mass  $m$  sat on the top of the mountain, then a stone with mass  $m_1$  falling down on one side of the mountain increases the probability that this stone falling down the mountain is a part of the stone that sat on the mountain's top a moment before. This, together with the fact that the masses  $m_1$  and  $m_2$  of the parts must sum up to the mass  $m$  of the stone before it broke, would increase the probability for a stone with mass  $m_2 = m - m_1$  falling down on the other side of the mountain.

## 4 Strategies to save Markov causation

In this section we present and review responses to counterexamples to Markov causation à la Cartwright (1999a, b). The first four responses claim that violations of MC arise only because the causal structure is misrepresented or variables are not correctly chosen in these scenarios. All of these strategies to save MC are to some extent controversial. Since they have already been discussed in detail elsewhere, our presentation of these responses will be quite brief. As a fifth and more recent possible response we discuss (in more detail) a proposal how to avoid violations of MC put forward by Schurz (2017). Though this approach is technically elegant and promising, we think that it leaves important questions open. In Sect. 5 we will then provide our own story about what might go on in the problematic scenarios and propose a new way to handle them that does not share the problems other strategies might have to face.

### 4.1 Too coarse-grained common cause variables

The first strategy to avoid Markov violations consists in claiming that the common cause variables in the problematic scenarios might be too coarse-grained. To adequately represent the processes going on in nature, one would need to replace the original common cause variables of our models by more fine-grained variables. If one would do that, then conditionalizing on the more fine-grained common cause variables would screen these common causes' effects off each other.<sup>5</sup> But since the chemical factory scenario is fictional anyway, Cartwright could simply insist that there is no way to represent the chemical process by a more fine-grained variable in her counterexample to MC (see, e.g., Cartwright 2002). The breaking of a stone due to erosion might be perfectly spontaneous as well, and in the decay process scenario it is unclear how or whether the common cause could be modeled more precisely in such a way that it would screen off its effects. Note that screening off could be restored in the EPR/B case by fine-graining the common cause variable if one is ready to subscribe to Bohmian mechanics (see, e.g., Egg and Esfeld 2014). But since Bohmian mechanics is a minority view among physicists that comes with its own problems, we stick with the standard interpretation in this paper and, hence, consider also the EPR/B example as not easily fixable.

<sup>5</sup> For an argumentation going in this direction see, for example, (Hausman and Woodward 1999).

## 4.2 Latent common causes

The second strategy to avoid Markov violations is to assume a latent common cause  $C'$  that is not represented in our causal models built in Sect. 3. Such an additional common cause  $C'$  would explain why the effect variables  $E_1$  and  $E_2$  in the problematic scenarios are not screened off by the common cause  $C$  represented in these models. But in the case of the chemical factory latent common causes are excluded by assumption. In the breaking stone example there is no obvious candidate for such an additional common cause, and in the case of the EPR/B experiment, latent common causes are excluded due to Bell inequalities. An argument why one should in general doubt the hidden common cause strategy has been put forward by Schurz (2017). In a nutshell, Schurz remarks that the effect variables  $E_1$  and  $E_2$  in the problematic scenarios are independent whenever the common cause is absent, i.e., when  $C = 0$ . In our examples this is the case if the chemical factory is inactive, there is no stone sitting at the top of the mountain, there is no particle of a certain kind present in the space region described by *Particle*, and the source does not emit any entangled pair of photons in the EPR/B experiment. But if there were another common cause  $C'$  of  $E_1$  and  $E_2$  in addition to the common cause  $C$  represented in our models, then the effects  $E_1$  and  $E_2$  could be expected to be correlated over the causal path  $E_1 \leftarrow C' \rightarrow E_2$ , even if  $C = 0$ . Note that postulating such an additional common cause  $C'$  does not strictly exclude an independence between  $E_1$  and  $E_2$  if  $C = 0$ . It is—though highly unlikely—possible to fine-tune the model's parameters in such a way that  $C = 0$  actually does screen  $E_1$  and  $E_2$  off each other. However, we think that Schurz' argument still succeeds in causing suspicion about the strategy to avoid Markov violations by postulating hidden common causes.

## 4.3 Missing intermediate causes

Another way to go consists in claiming that the common cause  $C$  used in the models we built in Sect. 3 is in fact not a common cause of the effect variables  $E_1$  and  $E_2$ , but rather a cause of such a common cause  $C'$  which is missing in our models (cf. Spirtes et al. 1993, pp. 61ff). An objection to Cartwright's (1999a, b) chemical factory counterexample in this spirit has, for example, been launched by Hausman and Woodward (1999). Hausman and Woodward suggest that there might be an intermediate common cause, viz. the firing of the chemical process. The active chemical factory would cause the firing of the process (with probability 0.8). But once the process fires, it determines the occurrence of the target substance and the pollutant in their corresponding tanks. Cartwright (2002) replied that she basically never assumed such an intermediate cause. As an alternative, one could also assume that the firing of the process is itself indeterministic. Also the breaking of a stone might be a perfectly spontaneous process, and the decay scenario and the EPR/B example (at least according to the predominant standard interpretation of quantum mechanics) are the prime examples for indeterministic processes per se. In addition, there seems to be no plausible can-

didate for an intermediate common cause  $C'$  available in the decay and the EPR/B scenarios.<sup>6</sup>

#### 4.4 Direct causal connections between effects

Yet another possible solution consists in assuming that one of the effects  $E_1$  or  $E_2$  in examples like the chemical factory causes the other one. Such a move seems implausible for several reasons. First of all, causation seems to need some kind of physical realizer. But we do not know of any physical processes or forces connecting the effects  $E_1$  and  $E_2$  in our examples. In addition, the effects  $E_1$  and  $E_2$  occur simultaneously, but causation is typically assumed to be forward directed in time (cf. Cartwright 1979; Reichenbach 1956; Suppes 1970). Finally, causal relations are typically assumed to transport influences due to interventions (cf. Pearl 2000; Woodward 2003) from the cause to the effect, but not the other way round. In case of the problematic scenarios we have discussed in Sect. 3 it seems, however, to be the case that we cannot influence one of the effects  $E_1$  or  $E_2$  by intervening on the other effect.<sup>7</sup> Putting the target chemical in one of the storage tanks in the chemical factory scenario, for example, does clearly not increase the probability that the pollutant can be found in the other tank (and vice versa).

#### 4.5 Interactive common causes

Another way to handle counterexamples to Markov causation like the chemical factory has recently been put forward by Schurz (2017). (See Näger 2013 for a similar proposal.) Instead of fiddling with the variables involved in the problematic scenarios or with the causal structure of the models representing these scenarios, Schurz proposes to modify MC. Schurz starts—inspired by Salmon (1984)—by distinguishing between two different kinds of common causes. A common cause  $C$  in a structure  $E_1 \leftarrow C \rightarrow E_2$  might either be a conjunctive or an interactive common cause.<sup>8</sup> Conjunctive common causes work like ordinary common causes: If two variables are causally connected only via a conjunctive common cause, then fixing this common cause's value will screen its effects off each other. Interactive common causes, on the other hand, work quite differently: Variables representing interactive common causes can be on or off. Conditionalizing on an interactive common cause's off value will screen its effects off each other (provided no other causal connections are around),

<sup>6</sup> For an in-depth discussion of possible causal structures in the EPR/B experiment, see, for example, (Näger 2013, 2016; Wood and Spekkens 2015). For an argument why such an intermediate common cause  $C'$  is excluded for theoretical reasons in the decay example, see (Schurz 2017).

<sup>7</sup> For an argumentation for the view that it is not even possible to independently intervene on the effects in the chemical factory example and in EPR/B experiments, see, for example, (Hausman and Woodward 1999).

<sup>8</sup> Schurz (2017) marks interactive common causes by an additional arc between the causal arrows. When looking at the details of the definitions he provides, however, it becomes clear that interactive common causes are not new theoretical elements of a model's causal structure  $\langle V, E \rangle$ . Schurz (2017, pp. 475f) defines an interactive common cause  $C$  as part of an ordinary common cause structure  $E_1 \leftarrow C \rightarrow E_2$  that satisfies several additional constraints.

while conditionalizing on one of its on values will still allow for a probabilistic dependence of the common cause's effects on each other. To make room for interactive common causes within the Bayes net framework, Schurz then proposes to weaken MC in such a way that interactive common causes are not required to screen off their effects when taking one of their on values anymore. (The technical details are not relevant here; for technical details how the revision of MC works, see Schurz 2017, p. 476.) Schurz' proposal to handle the problematic scenarios introduced in Sect. 3 would then amount to the claim that all the common causes in these scenarios are interactive common causes that can be handled by the weakened version of MC.

Schurz' (2017) approach to handle Markov violations can avoid the problems the other strategies outlined in section 4 have to face in a technically elegant way. We think, however, that it also leaves important questions open. It seems, for example, to be unclear whether Schurz' weaker version of MC can capture the main metaphysical assumption lying at the very heart of the project of approaching causation in terms of Bayes nets already mentioned in Sect. 2. This basic assumption is clearly reflected by the original Markov condition (Definition 2.1). It says that dependencies do not occur randomly, but are produced by structure. But this is only one side of the coin. That dependencies can only be due to structure means, on the other hand, that there can be no dependence when all pathways connecting two variables are blocked. Recall that one consequence of this basic assumption is that explaining a dependence between two effects  $E_1$  and  $E_2$  of a common cause  $C$  (with no other causal or non-causal relations around) amounts to citing  $E_1$  and  $E_2$ 's common cause  $C$ . The dependence is only fully explained when  $C$  screens  $E_1$  and  $E_2$  off each other. If it does not, then—since dependence is only due to structure—there must be other causal or non-causal relations involved which have been overlooked in the explanation. This should also hold for the systems in the four examples discussed. In our examples, however, there is still a dependence between the effect variables  $E_1$  and  $E_2$  conditional on  $C$ . But how is this possible? How should a common cause path  $E_1 \leftarrow C \rightarrow E_2$  (may  $C$  be conjunctive or interactive) be able to transport any probabilistic dependence between the effects  $E_1$  and  $E_2$  if the common cause variable  $C$ 's value is not allowed to vary? Or in other words: How is it possible that citing the common cause  $C$  fails to fully explain the dependence between its effects  $E_1$  and  $E_2$ ?

There are basically two metaphysical possibilities: either (i)  $E_1$  and  $E_2$  are only causally dependent (over the causal path  $E_1 \leftarrow C \rightarrow E_2$ ). If this is the case, then we have to take it as a brute fact (i.e., without any explanation) that some common causes do not screen off their effects. This would mean that there are, metaphysically speaking, true interactive common causes out there in the world and, hence, that there are dependencies that cannot be fully explained. Thus, it would go against the basic assumption that dependence is always due to structure. The other metaphysical possibility is that (ii) there is also some kind of non-causal dependence involved between  $E_1$  and  $E_2$ . In that case, it seems that there would be no true interactive common causes. What would actually happen in the problematic scenarios is that the purely causal dependence propagated over the causal path  $E_1 \leftarrow C \rightarrow E_2$  vanishes after conditionalizing on  $C$ . However, there would still be some kind of non-causal dependence between  $E_1$  and  $E_2$  given  $C$  that cannot be explained by citing the common cause  $C$ . This would mean that the structure  $\langle V, E \rangle$  underlying

the problematic scenarios is in truth richer. We have to deal with a mixed (causal and non-causal) dependence structure. Conditionalizing on  $C$  does not block all paths between  $E_1$  and  $E_2$  in the full mixed (causal and non-causal) structure.

Summarizing, one can tell two quite different metaphysical stories about why  $C$  does not screen off  $E_1$  and  $E_2$  from each other. Schurz' (2017) proposal to model the problematic scenarios is so general that it can capture both possibilities. This can clearly be seen as a merit of the approach. The downside of the approach's generality is, on the other hand, that it does not distinguish between these different metaphysical possibilities: Do interactive common causes really exist or are additional non-causal dependencies responsible for the Markov violation in the problematic scenarios? In the next section we will try to fill this gap.

## 5 Distribution conditions and common cause triggered non-causal dependencies

In this section we will explore yet another way to handle counterexamples to MC such as the chemical factory. We think that the scenarios discussed in Sect. 3 might violate MC because they involve non-causal dependencies of a certain kind. We proceed in two steps: First, we argue that some kind of non-causal dependence must be involved in scenarios like the chemical factory. Second, we provide a general but metaphysically cautious characterization of this kind of non-causal dependence.<sup>9</sup> We argue that it kicks in once the common cause occurs because of background assumptions that rule how quantities, properties, or parts are distributed among different objects or places. We finally propose a way to model these non-causal dependencies in such a way that the original version of MC and the metaphysical principle that dependence is due to structure can be preserved.

We start with the first step. Let us take it for granted that dependencies are always produced by structure. In all of our exemplary scenarios introduced in Sect. 3 the effects  $E_1$  and  $E_2$  are causally connected only via common cause structures  $E_1 \leftarrow C \rightarrow E_2$ . But this means that we should expect that all the probabilistic information we can get for one of the effects by observing the other effect has to be mediated over the common cause  $C$ . In constructing the counterexamples to Markov causation, however, we also clearly needed to say something about how the effects  $E_1$  and  $E_2$  depend on each other once  $C$ 's value is fixed. Let us briefly illustrate this by means of the chemical factory. To get the Markov violation, it does not suffice to just specify the conditional probabilities  $P(\text{Sub}|\text{Chem})$  and  $P(\text{Poll}|\text{Chem})$ . We also had to say that observing the value of  $\text{Poll}$  makes a difference for the probability of  $\text{Sub}$  when  $\text{Chem} = 1$ , and, vice versa, that observing the value of  $\text{Sub}$  has a probabilistic influence on  $\text{Poll}$  when  $\text{Chem} = 1$ . If we (i) take this dependence of  $\text{Sub}$  and  $\text{Poll}$  on each other when  $\text{Chem} = 1$  seriously, we (ii) take it for granted that the causal structure underlying the scenario is  $\text{Sub} \leftarrow \text{Chem} \rightarrow \text{Poll}$ , and we (iii) also take the metaphysical assumption that dependence can only be due to structure seriously, then there must

<sup>9</sup> We prefer lightweight metaphysical approaches, meaning that we will, instead of providing a deep metaphysical analysis, content ourselves with capturing only some characteristic marks and relevant formal properties of this kind of non-causal dependence.

be some kind of non-causal relationship between the variables *Sub* and *Poll* that can account for their dependence when  $Chem = 1$ . We see no other way to think about the situation once one accepts (i)–(iii). The same considerations apply to the other scenarios discussed.

Let us now come to the second step. First, note that for every one of the four scenarios we have discussed in Sect. 3 there is a background story to be told in addition to the causal and probabilistic information captured by the models we built for these scenarios. Sometimes the commitments that come with these background stories were more explicit, and sometimes they were more implicit. However, what all the scenarios we have discussed have in common is that there is some background story featuring laws of nature or assumptions which rule that if the common cause  $C$  occurs (i.e., if  $C$  takes one of its on values), then quantities, properties, or parts of a whole are distributed among objects or places modeled by the effect variables  $E_1$  and  $E_2$  in a specific way and with a non-extreme probability. In the following, we will refer to such laws or background assumptions as distribution conditions.<sup>10</sup>

Regardless of whether the distribution conditions are themselves of causal nature or not, our argumentation in step 1 seems to make it clear that they lead to a non-causal dependence between  $E_1$  and  $E_2$  once  $C$  has taken one of its on values. Let us illustrate this by the following example in which an amount of money is distributed among different persons.<sup>11</sup> Assume a rich uncle has ten million dollars as well as a poor nephew and a poor niece. According to the rich uncle's last will, his nephew should inherit four million dollars and his niece six million dollars after his death. Now assume that (for whatever reason) a random process (e.g., a coin toss) is used to decide whether the rich uncle's last will is enforced after his death. These are the distribution conditions. Next, we build the model. The binary variable *Death* models whether the rich uncle dies or not, *Nephew* the amount of money his nephew possesses, and *Niece* the amount of money his niece possesses. The causal structure of the model is  $Nephew \leftarrow Death \rightarrow Niece$ . If  $Death = 1$  and the niece possesses an amount of money close to six million dollars, then this drastically increases the probability for the nephew to possess an amount of money close to four million dollars, and vice versa. The cause  $Death = 1$  triggers, due to the distribution conditions, an additional dependence between *Nephew* and *Niece* that is not screened off by *Death*. This dependence must, according to assumptions (i)–(iii) in step 1, be of non-causal nature.

In the following, we will discuss the four counterexamples introduced in Sect. 3 again. It will turn out that all of these scenarios feature non-causal dependencies that kick in because an indeterministic common cause occurs and certain distribution conditions are in place.

<sup>10</sup> Note that distribution conditions—as will become clearer below—often come in the form of conservation laws. That conservation laws might play a central role in scenarios involving common causes that do not screen off their effects is not a new insight. It can, for example, already be found in (Salmon 1984). For further discussion of how conservation laws might contribute to Markov violations, see, for example, Cartwright (2007), Hausman and Woodward (1999), or Schurz (2017).

<sup>11</sup> The example also illustrates how distribution conditions together with an indeterministic common cause structure provide the ingredients for a general recipe for constructing counterexamples to MC like the chemical factory.

## 5.1 Chemical factory

In the case of the chemical factory, a certain chemical compound is split up in a product and a byproduct. How else should the chemical factory produce a target chemical in such a way that the target substance is always accompanied by the pollutant when the procedure succeeds? That a successful splitting process, as it is carried out in the chemical factory, always leads to the target substance and the pollutant seems to be due to the very nature of the chemical compound we split up. There seem to be chemical laws involved which rule that this specific kind of chemical compound can only be used to produce the target substance when it is split in such a way that the pollutant always occurs as a byproduct. These laws are the distribution conditions for the chemical factory scenario. They rule that there is a certain probability that specific parts of a whole are distributed among certain places (i.e., the storage tanks) if the chemical factory is active. If one of the parts can be found at the one place, this increases the probability that the other one will be at the other place. Note that this dependence between *Sub* and *Poll* given  $Chem = 1$  is not causal. There is no causal connection between these variables in addition to  $Sub \leftarrow Chem \rightarrow Poll$ . There is also no physical process or force that might realize such an additional causal connection.

## 5.2 Decay process

In this scenario a particle decays in two subparticles of the same kind which move away in opposite directions. This specific behavior lies in the nature of particles and is regulated by the laws of nature. In particular, the dependence between *SubP1* and *SubP2* given  $Particle = 1$  is due to the laws of conservation of momentum and energy. These laws are the distribution conditions in the decay process scenario. They rule how the subparticles are distributed in space if the original particle decays. Also the dependence of *SubP1* and *SubP2* on each other triggered by  $Particle = 1$  is something that comes in addition to the causal structure underlying the scenario. There seems to be no story to be told about how *SubP1* and *SubP2* are causally connected in addition to the one that cites the path  $SubP1 \leftarrow Particle \rightarrow SubP2$ .

## 5.3 EPR/B experiment

Also EPR/B experiments seem to involve distribution conditions. Pairs of entangled photons can, for example, be produced by splitting a photon beam with a non-linear crystal obeying energy and momentum conservation. So there seem to be laws of nature (i.e., the principles of energy and momentum conservation) ruling how quantities are distributed that are responsible for the fact that *Det1* and *Det2* are dependent conditional on  $\{QS, Pol1, Pol2\}$ . Also this dependence between *Det1* and *Det2* can—so we think—not be causally interpreted. Again, there are no physical processes or forces that might realize an additional causal relation between *Det1* and *Det2*. More importantly, causal influences are assumed to not spread faster than light. But the dependencies between the effects in EPR/B experiments triggered by their common causes would clearly violate locality.

## 5.4 Breaking stone

Also in the breaking stone scenario an object is split in two parts with a non-extreme probability. Here *Part1* and *Part2* seem to become dependent when  $Stone = m$  because of the very nature of macro physical objects described by the principle of mass conservation. So the law of mass conservation is the distribution condition in this scenario. It rules that when a macro object breaks, its mass  $m$  must be distributed among the parts resulting from the breaking event in such a way that the masses of all these parts sum up to  $m$ . The dependence between *Part1* and *Part2* that arises when  $Stone = m$  is, again, clearly non-causal. No time lag and also no physical process or force which might realize such an additional causal connection between *Part1* and *Part2* is involved here.

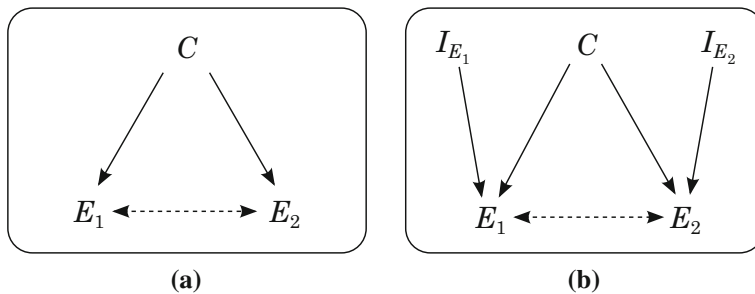
Summarizing, it seems that the effects  $E_1$  and  $E_2$  in all of the problematic scenarios introduced in Sect. 3 are not only dependent because of the common cause structure  $E_1 \leftarrow C \rightarrow E_2$ . There is also an additional (non-causal) dependence between  $E_1$  and  $E_2$  that is triggered when the common cause variable is on and that arises due to model-external distribution conditions. These distribution conditions rule how quantities, properties, or parts are distributed among objects or space regions once the common cause is in place.

Now there are two possible ways to proceed for supporters of the Bayes net approach to causation. The first one consists in accepting that the framework is not applicable to scenarios in which common causes do not screen off their effects because a basic assumption of that framework is not met in these scenarios: For successful causal modeling it is typically assumed that there are no other than causal dependencies around. The alternative way to go would consist in trying to represent the kind of non-causal dependence involved in the problematic scenarios within one's models. In this paper we follow the second possible way to go. This seems to be more challenging because, as Woodward (2015) remarks, it is still unclear how non-causal dependencies should be represented in causal models.

In the remainder of this paper we argue that non-causal dependencies due to distribution conditions that are triggered by indeterministic common causes have the same formal properties as latent common causes (which are typically represented by double-headed arrows).<sup>12</sup> To distinguish this kind of non-causal dependence from dependence due to latent common causes ( $\longleftrightarrow$ ), we represent it by dashed double-headed arrows ( $\leftarrow\text{---}\rightarrow$ ). Treating these non-causal dependencies similarly to latent common causes will allow us to keep the original version of MC since, as is well-known from the

<sup>12</sup> We would like to thank an anonymous referee for pointing at a possible tension between our claim that the kind of non-causal dependence we are interested in formally behaves like dependence due to latent common causes and Schurz' (2017) argument (presented in Sect. 4) against latent common causes. Schurz' argument is plausible because  $C = 0$  would most likely not screen off  $E_1$  and  $E_2$  in the presence of an additional common cause  $C'$ , which would require a very specific parameter fine-tuning which—according to what we know about causal relations—can be expected to almost never occur. We think, however, that this objection does not apply to the kind of non-causal dependence we have in mind. This kind of dependence only kicks in once some quantity, property, etc. is distributed. Without the common cause in place, no such distribution occurs and, thus, no dependence between  $E_1$  and  $E_2$  should be expected if  $C = 0$ .





**Fig. 1** Mixed (causal and non-causal) model not violating MC; (a) without intervention variables and (b) with intervention variables

causal modeling literature, models featuring single-headed as well as double-headed arrows representing latent common causes both comply with it.

Now why should the non-causal dependencies involved in the chemical factory scenario and the like formally behave like dependencies due to latent common causes? Latent common causes have basically two main features: They (i) typically produce a dependence between the effects  $E_1$  and  $E_2$  that cannot be accounted for by other causal connections between  $E_1$  and  $E_2$ , and they (ii) do not propagate probabilistic influences due to interventions on one of the effect variables to the other one. Both features are, in accordance with MC, captured by drawing a double-headed arrow  $E_1 \longleftrightarrow E_2$ . The kind of non-causal dependence we are interested in in this paper seems to share these key features: It produces an additional dependence between  $E_1$  and  $E_2$  that cannot be explained by other paths such as  $E_1 \leftarrow C \rightarrow E_2$ , and it does not support the mediation of probabilistic influences induced by interventions on  $E_1$  or  $E_2$ . In the case of the chemical factory scenario, for example, intervening on *Sub* would amount to putting the target chemical in its storage tank. This will not have any effect on whether also the pollutant is present in its tank (and vice versa). Hence, we can represent this kind of non-causal dependence by drawing a double-headed arrow  $E_1 \longleftrightarrow E_2$  and end up with the model in Fig. 1a which does not violate MC anymore.

Let us further explain why adding the double-headed arrow  $E_1 \longleftrightarrow E_2$  provides the right intervention properties. One standard way to represent an intervention on a variable  $X$  consists in adding an intervention variable  $I_X$  to one's model.  $I_X$  is assumed to be exogenous and a direct cause only of its target variable  $X$ . Variables  $Y$  that are  $d$ -connected to (and might, hence, be dependent on) such an intervention variable  $I_X$  can—at least in principle—be manipulated by  $I_X$ . Adding intervention variables for the effects of our three problematic scenarios would lead to the structure depicted in Fig. 1b. In this structure the intervention variable  $I_{E_1}$  for  $E_1$  is  $d$ -separated from  $E_2$ . Hence, intervening on  $E_1$  cannot have any effect on  $E_2$ . And, vice versa, any intervention variable  $I_{E_2}$  for  $E_2$  we add to our model is  $d$ -separated from  $E_1$ . Hence, no intervention on  $E_2$  can lead to a change in  $E_1$ .

Let us finally highlight some possible advantages of our approach to handle purported counterexamples to Markov causation à la Cartwright (1999a, b). One advantage over the first four strategies discussed in Sect. 4 is that we can accept scenarios such as the chemical factory as they are presented, i.e., without the need to deny that their causal structures have been misrepresented or that variables have not been chosen cor-

rectly. We can account for the non-causal aspects of these scenarios by allowing for richer structures  $\langle V, E \rangle$  featuring causal as well as non-causal components. We can avoid Markov violations by formally treating the specific kind of non-causal dependence involved in the problematic scenarios as if it were a causal dependence due to latent common causes.

Our approach also seems to have some advantages over approaches such as Schurz' (2017) and Näger's (2013) which modify the Bayes net framework by weakening MC. First, we can clearly preserve the original version of MC (Definition 2.1) together with the metaphysical core assumption underlying the causal interpretation of MC that dependence does not occur randomly, but is always produced by structure. In the case of the problematic scenarios introduced in Sect. 3, this structure is of mixed nature: It contains causal as well as non-causal elements. Blocking the common cause path  $E_1 \leftarrow C \rightarrow E_2$  by conditionalizing on  $C$  does not lead to an independence between the effect variables  $E_1$  and  $E_2$  because there is also the non-causal path  $E_1 \leftarrow \dots \rightarrow E_2$ . We cannot fully explain the dependence between  $E_1$  and  $E_2$  by citing the common cause  $C$  because there is an additional non-causal dependence between the two effect variables. Second, no sophisticated modification of the Bayes net approach is required to handle the problematic scenarios. How models featuring single-headed as well as double-headed arrows are connected to probabilities is already well-known (see, e.g., Richardson and Spirtes 2002): These models conform to the original version of the (global) Markov condition. Because of this, search algorithms capable of handling latent common causes such as FCI (Spirtes et al. 1993, p. 188) should also be applicable to sets of variables featuring the kind of non-causal dependence we were interested in in this paper.<sup>13</sup>

Another possible advantage of the mixed (causal and non-causal) model strategy over interactive common cause approaches is that it allows to draw a more specific metaphysical picture of what is going on in the problematic scenarios. As we have seen in Sect. 4, Schurz' (2017) more general approach leaves it open whether the dependence between the effects  $E_1$  and  $E_2$  when  $C = 1$  is only due to the common cause  $C$  or due to  $C$  and an additional non-causal dependence between  $E_1$  and  $E_2$ . In Sect. 4 we argued for the view that the common causes in the problematic scenarios do not screen off their effects because a certain kind of non-causal dependence between  $E_1$  and  $E_2$  is involved in these scenarios. If this is correct, then it seems justified to explicitly represent this non-causal dependence as a new structural element in one's models.

Here is the last possible merit of our approach we would like to mention: The approach might be used as a basis for developing or investigating the possibility of an account of mixed (causal and non-causal) explanation for certain phenomena.<sup>14</sup> Let us illustrate this by means of the chemical factory scenario. Assume that we observe

<sup>13</sup> We are indebted to an anonymous referee for pointing this out to us.

<sup>14</sup> Note that non-causal (and mixed causal and non-causal) explanations are currently experiencing a revival of some sort. Lange's (2016) work (but see also Reutlinger and Saatsi in press), for example, suggests that there is a recent interest in the possibility and merits of non-causal explanation in general. And then there are, for example, also mechanistic approaches to explanation which seem to flourish in philosophy of biology, medicine, and neuroscience (see, e.g., Bechtel and Abrahamsen 2005; Craver 2007, but also Illari et al. 2011). They heavily rely on causal as well as on non-causal constitutional relationships and, thus, on models that can combine both kinds of relations.

that the chemical factory is active, i.e., we learn that  $Chem = 1$ . Assume that we also observe that the target substance and the pollutant are present in their corresponding storage tanks, i.e., we learn that  $Sub = 1$  and that  $Poll = 1$ . Now we can ask why  $Poll = 1$  happened, i.e., why the pollutant is present in the right storage tank. When giving a purely causal explanation for a certain event, one has to refer to this event's causes (and only to its causes). If there is an additional non-causal dependence between  $Sub$  and  $Poll$ , however, it might make sense to cite  $Poll$ 's cause  $Chem$  as well as  $Sub$  for explaining  $Poll = 1$ . The purely causal explanation of  $Poll = 1$  could go as follows: The pollutant is present in the right storage tank ( $Poll = 1$ ) because the chemical factory was active ( $Chem = 1$ ). The probability of the explanandum event ( $Poll = 1$ ) the explanans  $Chem = 1$  of the purely causal explanation provides is 0.8. If we were also allowed to use the observation that the target substance is present in the left storage tank ( $Sub = 1$ ) for the explanation, we would, of course, get a higher probability of almost 1 for the explanandum event  $Poll = 1$ . Learning that the target substance is present in the left storage tank increases the probability that the chemical process was successful, which means, because of the distribution conditions (i.e., specific chemical laws) involved, that the pollutant must have been produced together with the target substance. At the same time it decreases the probability that the chemical process failed and that the pollutant is in the right tank for other reasons. (Maybe the tank has not been emptied the day before or someone put the pollutant there by hand.) Whether an account of explanation based on mixed (causal and non-causal) models might be fruitful and how exactly such an account might look like is something we will hopefully be able to investigate in future work.

## 6 Conclusion

In this paper we argued that purported counterexamples to Markov causation such as the chemical factory scenario do not meet an assumption typically made in causal modeling. These scenarios seem to violate MC because they involve non-causal dependencies. In particular, we have argued that distribution conditions together with the presence of an indeterministic common cause trigger a certain kind of non-causal dependence between the effects that formally behaves like dependence due to latent common causes. Once the common cause occurs, the distribution conditions allow the inference from observing the value of one of the effects to how quantities, properties, or parts of the object modeled by the common cause variable might have been distributed among the other effects. They do, however, not allow the propagation of influences due to interventions on the effect variables. We have then proposed a strategy to model the specific kind of non-causal dependence involved in the problematic scenarios. We proposed to represent these non-causal dependencies just like latent common causes by adding a new structural element to one's models: double-headed arrows.

We finally highlighted some possible advantages of our approach to handle counterexamples to Markov causation such as Cartwright's (1999a, b). First, we have a new way to deal with scenarios for which we find no plausible alternative causal structure or variables which would allow the cause to screen off its effects from each other. If we find no obvious errors in the building procedure of the models, it might be plausible to accept the causal and probabilistic features of these purported counterexamples just

the way they are presented and rather search for an additional non-causal dependence among the effects. Second, we do not have to modify the original version of MC and, thus, abandon the core idea underlying the project of approaching causation in terms of Bayes nets that dependencies are produced by structure, which seems to be an essential motivation for accepting Markov causation in the first place. The third advantage of our approach is that it allows to draw a metaphysically more specific picture of what is going on in the problematic scenarios than interactive common cause accounts: In truth, there are no interactive common causes; there are just ordinary (conjunctive) common causes that sometimes come hand in hand with additional non-causal dependencies between their effects. We also suggested a strategy for explicitly representing the non-causal dependencies involved in the problematic scenarios as parts of our models' structures. The resulting mixed (causal and non-causal) models might be used in future work as a basis for investigating the possibility of mixed (causal and non-causal) explanation.

There is a couple of open questions and problems. Here are some of them. This paper was about metaphysics and representation. But how does search work for systems involving causal as well as non-causal dependencies of the kind described? Since the kind of non-causal dependence we were interested in in this paper formally behaves like dependence due to latent common causes, search algorithms such as FCI should be applicable to such systems as well. But note that in the scenarios discussed we just assumed that we already know which parts of the structures underlying our models are of causal and which are of non-causal nature. So the problem how to distinguish non-causal dependencies from latent common causes remains to be solved. Another important question is whether there are common causes not involving distribution conditions that do not screen off their effects. If there are such scenarios, they might produce non-causal dependencies with different formal properties and our modeling strategy might not be applicable. Can the approach put forward in this paper be used to shed new light on open philosophical questions concerning the quantum realm? Can it, for example, contribute to the discussion of causal Markov violations versus faithfulness violations in quantum experiments (cf. Näger 2016; Retzlaff 2017; Wood and Spekkens 2015)? And, finally, can the approach be used as a basis for developing an account of mixed (causal and non-causal) explanation? Questions like these have to await exploration in future work.

**Acknowledgements** We thank Florian Boge, Christian Feldbacher-Escamilla, Slawa Kabanovic, Gry Oftedal, Gerhard Schurz, Annika Schuster, Anders Strand, Naftali Weinberger, and two anonymous referees for important discussions and helpful comments.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 421–441.

- Cartwright, N. (1979). Causal laws and effective strategies. *Noûs*, 13(4), 419–437.
- Cartwright, N. (1999a). Causal diversity and the Markov condition. *Synthese*, 121(1/2), 3–27.
- Cartwright, N. (1999b). *The dappled world*. Cambridge: Cambridge University Press.
- Cartwright, N. (2002). Against modularity, the causal Markov condition, and any link between the two: Comments on Hausman and Woodward. *British Journal for the Philosophy of Science*, 53(3), 411–453.
- Cartwright, N. (2007). *Hunting causes and using them*. Cambridge: Cambridge University Press.
- Craver, C. (2007). *Explaining the brain*. Oxford: Clarendon Press.
- Egg, M., & Esfeld, M. (2014). Non-local common cause explanations for EPR. *European Journal for Philosophy of Science*, 4(2), 181–196.
- Gebharter, A. (2017a). Causal exclusion and causal Bayes nets. *Philosophy and Phenomenological Research*, 95(2), 353–375.
- Gebharter, A. (2017b). *Causal nets, interventionism, and mechanisms*. Cham: Springer.
- Gebharter, A. (2017c). Uncovering constitutive relevance relations in mechanisms. *Philosophical Studies*, 174(11), 2645–2666.
- Glymour, C. (1999). Rabbit hunting. *Synthese*, 121(1/2), 55–78.
- Glymour, C. (2006). Markov properties and quantum experiments. In W. Demopoulos & I. Pitowsky (Eds.), *Physical theory and its interpretation* (pp. 117–126). Netherlands: Springer.
- Glymour, C., Spirtes, P., & Scheines, R. (1991). Causal inference. *Erkenntnis*, 35(1/3), 151–189.
- Hausman, D. (1998). *Causal asymmetries*. Cambridge: Cambridge University Press.
- Hausman, D., & Woodward, J. (1999). Independence, invariance and the causal Markov condition. *British Journal for the Philosophy of Science*, 50(4), 521–583.
- Healey, R. (2009). Causation in quantum mechanics. In H. Beebe, C. Hitchcock, & P. Menzies (Eds.), *The Oxford handbook of causation*. Oxford: Oxford University Press.
- Illari, P. M., Russo, F., & Williamson, J. (Eds.). (2011). *Causality in the sciences*. Oxford: Oxford University Press.
- Lange, M. (2016). *Because without cause: Non-causal explanations in science and mathematics*. Oxford: Oxford University Press.
- Näger, P. M. (2013). Causal graphs for EPR experiments. Retrieved from <http://philsci-archive.pitt.edu/id/eprint/9915>.
- Näger, P. M. (2016). The causal problem of entanglement. *Synthese*, 193(4), 1127–1155.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo, CA: Morgan Kaufmann.
- Pearl, J. (2000). *Causality* (1st ed.). Cambridge: Cambridge University Press.
- Reichenbach, H. (1956). *The direction of time*. Berkeley: University of California Press.
- Retzlaff, N. (2017). Another counterexample to Markov causation from quantum mechanics: Single photon experiments and the Mach–Zehnder interferometer. *Kriterion*, 31(2), 17–42.
- Reutlinger, A., & Saatsi, J. (in press). *Explanation beyond causation: Philosophical perspectives on non-causal explanations*. Oxford University Press.
- Richardson, T., & Spirtes, P. (2002). Ancestral graph Markov models. *Annals of Statistics*, 30(4), 962–1030.
- Salmon, W. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.
- Schurz, G. (2017). Interactive causes: Revising the Markov condition. *Philosophy of Science*, 84(3), 456–479.
- Schurz, G., & Gebharter, A. (2016). Causality as a theoretical concept: Explanatory warrant and empirical content of the theory of causal nets. *Synthese*, 193(4), 1073–1103.
- Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search* (1st ed.). Dordrecht: Springer.
- Sprenger, J., & Hartmann, S. (in press). *Bayesian philosophy of science*. Oxford: Oxford University Press.
- Suppes, P. (1970). *A probabilistic theory of causality*. Amsterdam: North-Holland.
- Wood, C. J., & Spekkens, R. W. (2015). The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning. *New Journal of Physics*, 17, 1–29.
- Woodward, J. (2003). *Making things happen*. Oxford: Oxford University Press.
- Woodward, J. (2015). Interventionism and causal exclusion. *Philosophy and Phenomenological Research*, 91(2), 303–347.