Bell's Local Causality is a d-Separation Criterion

Gábor Hofer-Szabó

Abstract This paper aims to motivate Bell's notion of local causality by means of Bayesian networks. In a locally causal theory any superluminal correlation should be screened off by atomic events localized in any so-called *shielder-off region* in the past of one of the correlating events. In a Bayesian network any correlation between non-descendant random variables are screened off by any so-called *d-separating set* of variables. We will argue that the shielder-off regions in the definition of local causality conform in a well defined sense to the d-separating sets in Bayesian networks.

Keywords Local causality · Shielder-off region · Bayesian network · d-separation

1 Introduction

John Bell's notion of local causality is one of the central notions in the foundations of relativistic quantum physics. Bell himself has returned to the notion of local causality from time to time providing a more and more refined formulation for it. The final formulation stems from Bell's posthumously published paper "La nouvelle cuisine." It reads as follows[1]:

> A theory will be said to be locally causal if the probabilities attached to values of local beables in a space-time region V_A are unaltered by specification of values of local beables in a space-like separated region V_B, when what happens in the backward light cone of V_A is already sufficiently specified, for example by a full specification of local beables in a space-time region V_C [1, 2, p. 239–240].

[1] For the sake of uniformity we slightly changed Bell's notation and figure.

G. Hofer-Szabó (✉)
Institute of Philosophy, Research Center for the Humanities, Országház u. 30, Budapest 1014, Hungary
e-mail: szabo.gabor@btk.mta.hu

M. Ozawa et al. (eds.), *Reality and Measurement in Algebraic Quantum Theory*,
Springer Proceedings in Mathematics & Statistics 261,
https://doi.org/10.1007/978-981-13-2487-1_2

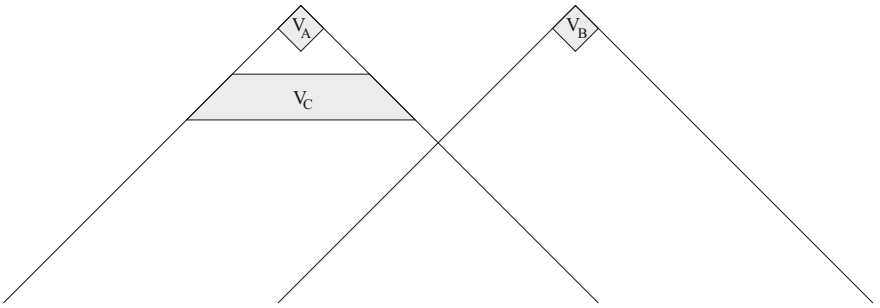

Fig. 1 Full specification of what happens in V_C makes events in V_B irrelevant for predictions about V_A in a locally causal theory

The figure Bell is attaching to his formulation of local causality is reproduced in Fig. 1 with Bell's original caption. In a rough translation, a theory is locally causal if any superluminal correlation can be screened-off by a "full specification of local beables in a space-time region" in the past of one of the correlating events.

The terms in quotation marks, however, need clarification. What are "local beables"? What is "full specification" and why is it important? Which are those regions in spacetime which, if fully specified, render superluminally correlating events probabilistically independent? The first two questions have attracted much interest among philosophers of science. As Bell puts it, "*be*ables of the theory are those entities in it which are, at least tentatively, to be taken seriously, as corresponding to something real" [1, 2, p. 234]. Furthermore, "it is important that events in V_C be specified completely. Otherwise the traces in region V_B of causes of events in V_A could well supplement whatever else was being used for calculating probabilities about V_A" [1, 2, p. 240].

The third question, however, concerning the localization of the screener-off regions has gained much less attention in the literature. How to characterize the regions which region V_C in Fig. 1 is an example of? Bell's answer is instructive but brief: "It is important that region V_C completely shields off from V_A the overlap of the backward light cones of V_A and V_B [1, 2, p. 240]." But why to shield off the common past of the correlating events? Why the region V_C cannot be in the remote past of V_A as for example in Fig. 2? Well, intuition dictates that in this latter case some event might occur *above* the shielder-off region but still *within* the common past establishing a correlation between events in V_A and V_B. This intuition is correct. The aim of this paper, however, is to provide a more precise explanation for the localization of the shielder-off regions in spacetime. This explanation will consists in drawing a parallel between local physical theories and Bayesian networks. It will turn out that *the*

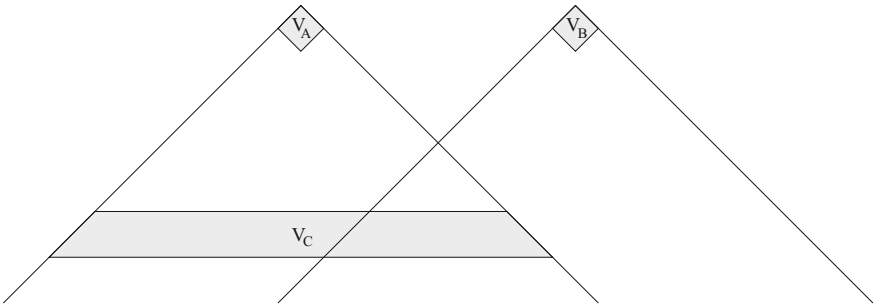

Fig. 2 A *not* completely shielding-off region V_C

shielder-off regions in the definition of local causality play an analogous role to the so-called d-separating sets of random variables in Bayesian networks.

There is a renewed interest in Bell's notion of local causality [22–24], its relation to separability [12]; the role of full specification in local causality [13, 32]; its role in relativistic causality [3, 4, 27]; its status as a local causality principle [10, 11, 28]. A similar closely related topic, the Common Cause Principle is also given much attention [15, 17, 26, 29]. On the other hand, there is also an intensive discussion on the applicability of the Causal Markov Condition in the EPR scenario [5, 21, 33, 34], Hausman and Woodward 1999. Despite the rich and growing literature on the topic I am unaware of any work relating Bayesian networks and especially d-separation directly to local causality. This paper intends to fill this gap. For a precursor of this paper investigating Causal Markov Condition in a specific local physical theory see [14]. For a comprehensive formally rigorous investigation of the relation of Bell's local causality to the Common Cause Principle and other relativistic locality concepts see [19]; for a more philosopher-friendly version see [20].

In the paper we will proceed as follows. In Sect. 2 we introduce the basics of the theory of Bayesian networks and the notion of d-separation and m-separation. In Sect. 3 we define the notion of a local physical theory and formulate Bell's notion of local causality within this framework. We prove our main claim in Sect. 4 and conclude in Sect. 5.

2 Bayesian Networks and d-Separation

A *Bayesian network* [6, 25] is a pair $(\mathscr{G}, \mathscr{V})$ where $\mathscr{G}$ is a directed acyclic graph and $\mathscr{V}$ is a set of random variables on a classical probability space (X, Σ, p) such

that the elements $A, B \ldots$ of $\mathscr{V}$ are represented by the vertices of $\mathscr{G}$ and the arrows (directed edges) $A \to B$ on the graph represent that A is *causally relevant* for B. Two vertices are called *adjacent* if they are connected by an arrow. For a given $A \in \mathscr{V}$, the set of vertices that have directed edges in A is called the *parents* of A, denoted by $Par(A)$; the set of vertices from which a directed paths is leading to A is called the *ancestors* of A, denoted by $Anc(A)$; and finally the set of vertices that are endpoints of a directed paths from A is called the *descendants* of A, denoted by $Des(A)$. For a set $\mathscr{C}$ of vertices $Par(\mathscr{C})$, $Anc(\mathscr{C})$ and $Des(\mathscr{C})$ are defined similarly.

The set $\mathscr{V}$ is said to satisfy the *Causal Markov Condition* relative to the graph $\mathscr{G}$ if for any $A \in \mathscr{V}$ and any $B \notin Des(A)$ the following is true:

$$p(A \mid Par(A) \wedge B) = p(A \mid Par(A)) \tag{1}$$

or equivalently

$$p(A \wedge B \mid Par(A)) = p(A \mid Par(A))\, p(B \mid Par(A)) \tag{2}$$

That is conditioning on its parents any random variable will be probabilistically independent from any of its non-descendant. Non-descendants can be of two types: either ancestors or *collaterals* (non-descendants and non-ancestors). As we will see, being independent of collaterals is what relates the Causal Markov Condition to Bell's local causality.

Causal Markov Condition establishes a special conditional independence relation between some random variables of $\mathscr{V}$. But there are many other conditional independences. In a faithful Bayesian network these other conditional independences are all implied by the Causal Markov Condition by means of the so-called *d-separation* criterion. Let $\mathscr{P}$ be a *path* in $\mathscr{G}$, that is a sequence of adjacent vertices. A variable E on $\mathscr{P}$ is a *collider* if there are arrows to E from both its neighbors on $\mathscr{P}$ ($D \to E \leftarrow F$). Now, let $\mathscr{C}$ be a set of vertices and let A and B two different vertices not in $\mathscr{C}$. The vertices A and B are said to be *d-connected* by $\mathscr{C}$ in $\mathscr{G}$ iff there exists a path $\mathscr{P}$ between A and B such that every non-collider on $\mathscr{P}$ is not in $\mathscr{C}$ and every collider is in $Anc(\mathscr{C}) \vee \mathscr{C}$. A and B are said to be *d-separated* by $\mathscr{C}$ in $\mathscr{G}$, iff they are not d-connected by $\mathscr{C}$ in $\mathscr{G}$.

The intuition behind d-separation is the following. A vertex E on a path (not at the endpoints) can be either a *collider* ($D \to E \leftarrow F$), an *intermediary cause* ($D \to E \to F$) or a *common cause* ($D \leftarrow E \to F$). The idea here is that only intermediary and common causes (together called *non-colliders*) can transmit causal dependence and hence establish probabilistic dependence. This dependence can be *blocked* by conditioning on the non-collider. Colliders behave just the opposite way. They represent two events causing a common effect. These two causes are causally and probabilistically independent, but become dependent upon conditioning on their common effect. Moreover, they also become dependent upon conditioning on any of the descendants of the effect. Putting these together, the causal dependence on a path $\mathscr{P}$ connecting two vertices is blocked by a set $\mathscr{C}$ if either there is at least one non-collider on $\mathscr{P}$ which is in $\mathscr{C}$ or there is at least one collider E on $\mathscr{P}$ such that

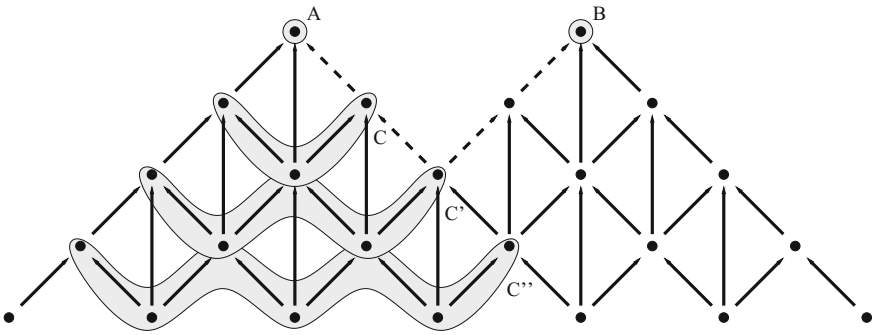

Fig. 3 A and B are d-separated by $\mathscr{C}$ and $\mathscr{C}'$ but d-connected by $\mathscr{C}''$

either E or a descendant of E is not in $\mathscr{C}$. The two vertices are d-separated by $\mathscr{C}$ if causal dependence is blocked on every path connecting them.

As an example for d-connection and d-separation consider the causal graph in Fig. 3. (The arrows are directed to up, left up and right up.) Let A be the left "peak" and B the right "peak" in the graph and let $\mathscr{C}$, $\mathscr{C}'$ and $\mathscr{C}''$ be the sets shown in the figure containing 3, 5 and 7 vertices, respectively. Then A and B are d-separated by $\mathscr{C}$ since the parents are always d-separating due to the Causal Markov Condition. A and B are d-separated also by $\mathscr{C}'$ since for every path connecting the peaks there is a non-collider in $\mathscr{C}'$. However, A and B are d-connected by $\mathscr{C}''$ since there is a path (denoted by a broken line in Fig. 3) connecting the peaks which contains only non-colliders outside $\mathscr{C}''$. Consequently, the following probabilistic relations hold:

$$p(A \wedge B \,|\, \mathscr{C}) = p(A \,|\, \mathscr{C})\, p(B \,|\, \mathscr{C}) \tag{3}$$

$$p(A \wedge B \,|\, \mathscr{C}') = p(A \,|\, \mathscr{C}')\, p(B \,|\, \mathscr{C}') \tag{4}$$

$$p(A \wedge B \,|\, \mathscr{C}'') \neq p(A \,|\, \mathscr{C}'')\, p(B \,|\, \mathscr{C}'') \tag{5}$$

Looking at in Fig. 3, what stands out immediately is that a set which is too far in the causal past of A cannot d-separate A from a collateral event since there might be paths connecting them "above" the set. As we will see, a similar moral will be valid in case of local causality: regions with are too far in the causal past of an event cannot screen it off from a spacelike separated event since there might be events "above" the region which can establish correlation between them.

In analyzing local causality sometimes we need to go beyond directed acyclic graphs. A graph which may contain both directed ($A \to B$) and bi-directed ($A \leftrightarrow B$) edges is called *mixed*. The d-separation criterion extended to mixed acyclic graphs

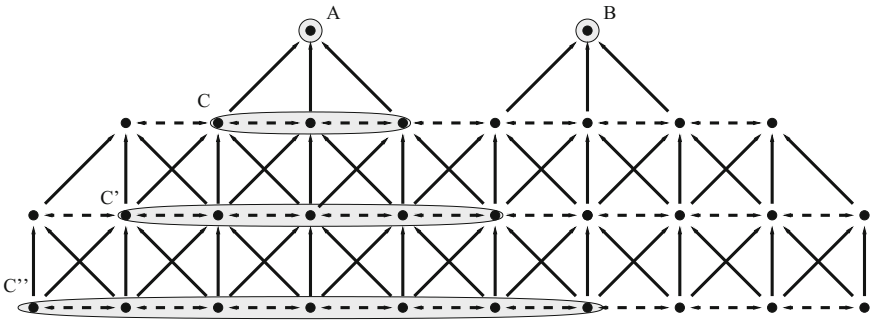

Fig. 4 A and B are m-separated by $\mathscr{C}$ but m-connected by both $\mathscr{C}'$ and $\mathscr{C}''$

is called *m-separation* [30, 31]. Two vertices A and B are said to be *m-connected* by $\mathscr{C}$ in a mixed acyclic graph $\mathscr{G}$ iff there exists a path $\mathscr{P}$ between A and B such that every non-collider on $\mathscr{P}$ is not in $\mathscr{C}$ and every collider is in $Anc(\mathscr{C}) \vee \mathscr{C}$. A and B are said to be *m-separated* by $\mathscr{C}$ in $\mathscr{G}$, iff they are not m-connected by $\mathscr{C}$ in $\mathscr{G}$. In a directed acyclic graph m-separation reduces to d-separation.

An example for a mixed acyclic graph is depicted in Fig. 4. Here the bi-directed edges are represented by dotted lines. Again, let A be the left "peak" and B the right "peak" in the graph and let $\mathscr{C}$, $\mathscr{C}'$ and $\mathscr{C}''$ be the sets shown in the figure containing 3, 5 and 7 vertices, respectively. Then A and B are m-separated by $\mathscr{C}$ but m-connected by both $\mathscr{C}'$ and $\mathscr{C}''$. The connecting path is the shortest path connecting A and B.

Now, let us connect the terminology of Bayesian networks to that of standard physics. Before doing that note that probability is commonly interpreted in Bayesianism subjectively as partial belief and in physics objectively as long-run relative frequency. This interpretative difference, however, does not undermine the analogy between local causality and d-separation, since Bayesian networks are well open to statistical interpretation and, conversely, there is a growing tendency to understand quantum physics in a subjectivist way.

Let us start with random variables. A random variable is a real-valued Borel-measurable function on X. Each random variable $A \in \mathscr{V}$ generates a sub-σ-algebra of Σ by the inverse image of the Borel sets:

$$\sigma(A) := \left\{A^{-1}(b) \mid b \in \mathscr{B}(\mathbb{R})\right\} \tag{6}$$

Similarly, each set $\mathscr{C}$ of n random variables generates a sub-σ-algebra of Σ by the inverse image of the n-dimensional Borel sets:

$$\sigma(\mathscr{C}) := \left\{(C_1, C_2 \ldots C_n)^{-1}(b) \mid C_i \in \mathscr{C},\ b \in \mathscr{B}(\mathbb{R}^n)\right\} \tag{7}$$

From this perspective d-separation tells us which sub-σ-algebras are probabilistically independent conditioned on which other sub-σ-algebras of Σ.

Now, instead of using σ-algebras it is more instructive to use a richer structure in physics, namely *von Neumann algebras*. Consider the characteristic functions on X projecting on the elements of Σ, called *events*. The set $\{\chi_S \mid S \in \Sigma\}$ of characteristic functions generates an abelian von Neumann algebra, namely $\mathscr{L}^\infty(X, \Sigma, p)$, the space of essentially bounded complex-valued functions on X. Starting from the characteristic functions of the sub-σ-algebra $\sigma(A)$, one arrives at a subalgebra of $\mathscr{L}^\infty(X, \Sigma, p)$. Denote this abelian von Neumann algebra determined by the random variable A by $\mathscr{N}_A$. Similarly, denote by $\mathscr{N}_\mathscr{C}$ the von Neumann algebra determined by a set $\mathscr{C}$ of random variables.

Instead of using a probability measure on Σ or on a sub-σ-algebra $\sigma(A)$, one can also use a state on the corresponding von Neumann algebra $\mathscr{N}_A$. A *state* ϕ is a positive linear functional of norm 1 on a von Neumann algebra. States on $\mathscr{N}_A$ and probability measures on $\sigma(A)$ mutually determine one another: a state restricted to the characteristic functions in $\mathscr{N}_A$ is a probability measure on $\sigma(A)$; and vice versa, integrating elements of $\mathscr{N}_A$ according to a probability measure on $\sigma(A)$ yields a state on $\mathscr{N}_A$.

Therefore, a conditional independence between random variables A and B given the set $\mathscr{C}$

$$p(A \wedge B \mid \mathscr{C}) = p(A \mid \mathscr{C})\, p(B \mid \mathscr{C}) \tag{8}$$

can be rewritten as follows: for any projection $A \in \mathscr{N}_A$, $B \in \mathscr{N}_B$ and $C \in \mathscr{N}_\mathscr{C}$:

$$\frac{\phi(A \wedge B \wedge C)}{\phi(C)} = \frac{\phi(A \wedge C)}{\phi(C)} \frac{\phi(B \wedge C)}{\phi(C)} \tag{9}$$

Although in this paper we stay at the classical level, the theory of von Neumann algebras is wide enough to incorporate also quantum physics. In this case the von Neumann algebras are nonabelian. The events, just like in the classical case, are represented by projections of the von Neumann algebras. In the quantum case conditional independence between the projection $A \in \mathscr{N}_A$ and $B \in \mathscr{N}_B$ given $C \in \mathscr{N}_\mathscr{C}$ reads as follows:

$$\frac{\phi(CABC)}{\phi(C)} = \frac{\phi(CAC)}{\phi(C)} \frac{\phi(CBC)}{\phi(C)} \tag{10}$$

which in the classical case reduces to (9).

The last point in converting the formalism of Bayesian networks into physics, is to swap the causal graph for spacetime. We can then replace the causal relations embodied in the causal graph by spatiotemporal relations of a given spacetime. Instead of saying that a random variable is the *ancestor* of another variable we will

then say that an event is in the *past* of the other. But to do so first we need to *localize* events in spacetime that is we need to have an association of algebras of events to spacetime regions. Such a principled association is offered by the formalism of algebraic quantum field theory. Hence, in the next section we will introduce some elements of algebraic quantum field theory which is indispensable for our purpose which is to come up with a mathematically precise definition of Bell's notion of local causality.

3 Bell's Local Causality in a Local Physical Theory

Let $\mathscr{M}$ be a globally hyperbolic spacetime and let $\mathscr{K}$ be a covering collection of bounded, globally hyperbolic subspacetime regions of $\mathscr{M}$ such that $(\mathscr{K}, \subseteq)$ is a directed poset under inclusion $\subseteq$. A *local physical theory* is a net $\{\mathscr{A}(V), V \in \mathscr{K}\}$ associating algebras of events to spacetime regions which satisfies *isotony* and *microcausality* defined as follows [7, 8, 19, 20]:
Isotony. The net of local observables is given by the isotone map $\mathscr{K} \ni V \mapsto \mathscr{A}(V)$ to unital C^*-algebras, that is $V_1 \subseteq V_2$ implies that $\mathscr{A}(V_1)$ is a unital C^*-subalgebra of $\mathscr{A}(V_2)$. The *quasilocal algebra* $\mathscr{A}$ is defined to be the inductive limit C^*-algebra of the net $\{\mathscr{A}(V), V \in \mathscr{K}\}$ of local C^*-algebras.
Microcausality: $\mathscr{A}(V')' \cap \mathscr{A} \supseteq \mathscr{A}(V)$, $V \in \mathscr{K}$, where primes denote spacelike complement and algebra commutant, respectively.

If the quasilocal algebra $\mathscr{A}$ of the local physical theory is commutative, we speak about a *local classical theory*; if $\mathscr{A}$ is noncommutative, we speak about a *local quantum theory*. For local classical theories microcausality fulfills trivially.

Given a state ϕ on the quasilocal algebra $\mathscr{A}$, the corresponding GNS representation $\pi_\phi \colon \mathscr{A} \to \mathscr{B}(\mathscr{H}_\phi)$ converts the net of C^*-algebras into a net of C^*-subalgebras of $B(\mathscr{H}_\phi)$. Closing these subalgebras in the weak topology one arrives at a net of local von Neumann observable algebras: $\mathscr{N}(V) := \pi_\phi(\mathscr{A}(V))''$, $V \in \mathscr{K}$. The *net* $\{\mathscr{N}(V), V \in \mathscr{K}\}$ *of local von Neumann algebras* also obeys isotony and microcausality, hence we can also refer to it as a local physical theory.

Given a local physical theory, we can turn now to the definition of Bell's notion of local causality. Recall that according to Bell a theory is locally causal if any superluminal correlation is screened-off by a "full specification of local beables in a space-time region V_C" as shown in Fig. 1. As indicated in the Introduction we need to address three questions. What are "local beables"? What is "full specification"? Which are the shielder-off regions? The brief answer to the first two questions is the following. In a local physical theory a "local beable" in a region V is an *element* of the local von Neumann algebra $\mathscr{N}(V)$. A "full specification" of local beables in region V is an *atomic element* of the local von Neumann algebra $\mathscr{N}(V)$. In this paper we do not comment on these two answers. For a more thoroughgoing discussion on why we think this to be the correct translation of Bell's intuition into our framework see [19, 20].

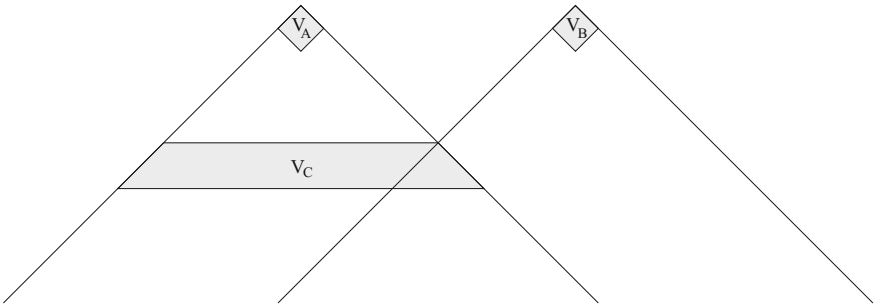

Fig. 5 A completely shielding-off region V_C intersecting with the common past of V_A and V_B

To the third question, which is the topic of our paper, the answer is this: a shielder-off region V_C is a region in the causal past of V_A which can block any causal influence on V_A arriving from the common past of V_A and V_B. But there is an ambiguity in this answer. Bell's Fig. 1 suggests that a shielder-off region should not intersect with the common past. Whereas the requirement of simply blocking causal influences from the past allows for also regions depicted in Fig. 5 intersecting with the common past. This means that one can define a *shielder-off region of* V_A *relative to* V_B either as a region V_C satisfying:

$\mathbf{L}_1 : V_C \subset J_-(V_A)$ (V_C is in the causal past of V_A),
$\mathbf{L}_2 : V_A \subset V_C''$ (V_C is wide enough such that its causal shadow contains V_A),
$\mathbf{L}_3^Q : V_C \subset V_B'$ (V_C is spacelike separated from V_B)

in tune with Bell's Fig. 1; or one can replace L_3^Q by the weaker requirement

$\mathbf{L}_3^C : J_-(V_C) \supset J_-(V_A) \cap J_-(V_B)$ (The causal past of V_C contains the common past of V_A and V_B)

allowing for regions such as in Fig. 2. It turns out that (with respect to the Bell inequalities, see [16, 18]) it is more appropriate to demand L_3^Q in case of a local *quantum* theory and L_3^C in case of a local *classical* theory (hence the superscripts). But note that as the covering regions become infinitely thin shrinking down to a Cauchy surface, requirement L_3^C coincides with requirement L_3^Q.

With all these considerations in mind Bell's notion of local causality in the framework of a local physical theory will be the following:

Definition 1 A local physical theory represented by a net $\{\mathcal{N}(V), V \in \mathscr{K}\}$ of von Neumann algebras is called *locally causal* (in Bell's sense), if

1. for any pair $A \in \mathscr{N}(V_A)$ and $B \in \mathscr{N}(V_B)$ of events represented by projections in spacelike separated regions $V_A, V_B \in \mathscr{K}$;
2. for every locally normal and faithful state ϕ establishing a correlation $\phi(AB) \neq \phi(A)\phi(B)$ between A and B;
3. for any spacetime shielder-off region V_C defined by requirements L_1, L_2 and L_3^Q/L_3^C;
4. for any event C in the set $\mathscr{C}$ of atomic events in $\mathscr{A}(V_C)$ the following screening-off condition holds:

$$\frac{\phi(CABC)}{\phi(C)} = \frac{\phi(CAC)}{\phi(C)}\frac{\phi(CBC)}{\phi(C)} \tag{11}$$

which for a local *classical* theory is equivalent to

$$p(A \wedge B \,|\, \mathscr{C}) = p(A \,|\, \mathscr{C})\, p(B \,|\, \mathscr{C}) \tag{12}$$

In short, a local physical theory is locally causal in Bell's sense if every superluminal correlation is screened off by all atomic events in all shielder-off region. (For many delicate questions such as what if the algebras are non-atomic, how this definition of local causality relates to the Common Cause Principle and the Bell inequalities see again [19, 20].)

The question left is, however: *why* shielder-off regions are characterized by requirements L_1, L_2 and L_3^Q/L_3^C? To this we turn in the next section.

4 Shielder-Off Regions are d-Separating

The point we are going to make in this Section is that shielder-off regions in the definition of local causality conform to d-separating sets in directed acyclic graphs and to m-separating sets in mixed acyclic graphs.

First we show how a local physical theory gives rise to a causal graph. Consider a local *classical* theory $\{\mathscr{N}(V), V \in \mathscr{K}\}$ where the covering collection is induced by a partition $\mathscr{T}$ of a spacetime $\mathscr{M}$. By *partition* we mean a countable set of disjoint, bounded spacetime regions such that their union is $\mathscr{M}$. The local classical theory $\{\mathscr{N}(V), V \in \mathscr{K}\}$ gives rise to a causal graph $\mathscr{G}$ as follows: Let the *vertices* of the $\mathscr{G}$ be the regions in the partition, $\{V \in \mathscr{T}\}$. For two vertices V_A and V_B, let there be an *edge* pointing from V_A and V_B, $V_A \to V_B$, iff there is a future directed causal curve from V_A to V_B such that the curve does not enter any region, except for V_A and V_B. It will turn out that the type of the graph we obtain is crucially depending on the partition $\mathscr{T}$ of the spacetime. Let us see some different cases.

If $\mathscr{M}$ is the 1+1 dimensional Minkowski spacetime, then it can be covered by double cones of equal size. (See Fig. 6.) The causal graph corresponding to this covering emerges simply by connecting those adjacent double cones which lie in the

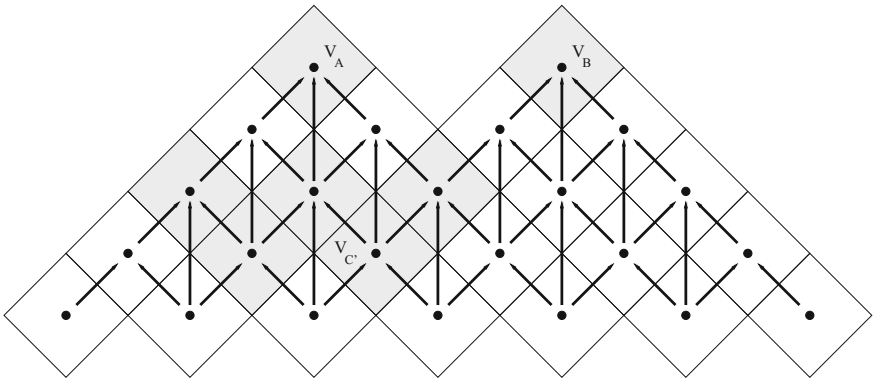

Fig. 6 The directed acyclic graph generated by double cones of equal size covering the 1+1 dimensional Minkowski spacetime

causal past of one another. What we get is just the directed acyclic graph depicted in Fig. 3 in Sect. 2.

Figure 6 is a kind of "superposition" of a spacetime diagram and a Bayesian network. Consider for example region $V_{\mathscr{C}'}$. Reading Fig. 6 as a spacetime diagram, one sees that $V_{\mathscr{C}'}$ is a *shielder-off region*. Reading Fig. 6 as a causal graph, one observes that the set $\mathscr{C}'$ corresponding to $V_{\mathscr{C}'}$ (depicted in Fig. 3) is a *d-separating set*. Similarly, one can check that the region associated to the d-separating set $\mathscr{C}$ in Fig. 3 is a shielder-off region and the region associated to the d-connecting set $\mathscr{C}''$ is not a shielder-off region.

A general spacetime $\mathscr{M}$ cannot be partitioned to globally hyperbolic regions, let alone to double cones. Still one can construct the causal graph corresponding to a partition $\mathscr{T}$. In Fig. 7 we illustrate such a construction where a 1+1 dimensional Minkowski spacetime is covered by boxes of equals size. (This example, in contrast to the previous one, can be generalized for a $3+1$-dimensional Minkowski spacetime covered by $3+1$-dimensional boxes of equals size.) The causal graph emerging from this construction is not a directed acyclic graph since it contains bi-directed edges: spacelike neighboring boxes will be *spouses*. What we get is a mixed acyclic graph depicted in Fig. 4. Again, confronting Figs. 4 and 7 one can see that the set $\mathscr{C}'$ is *not* an m-separating set and at the same time the corresponding region $V_{\mathscr{C}'}$ is *not* a shielder-off region of V_A relative to V_B.

The exact characterization of the graphs emerging from a different coverings of a given spacetime is a subtle question which we do not go into here. Instead we turn now to the construction of random variables. Let $\mathscr{N}(V)$ be the local von

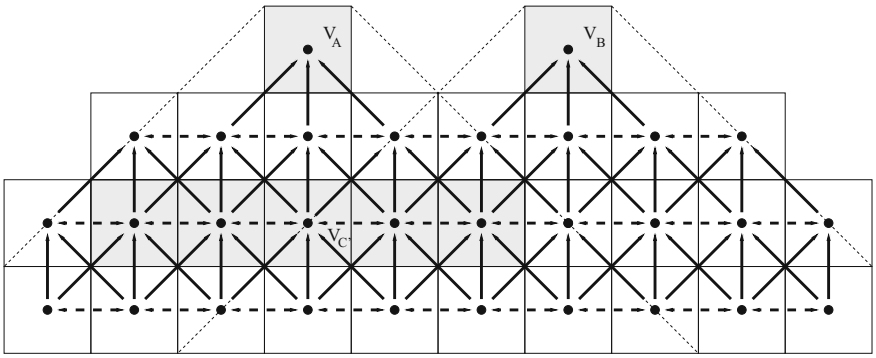

Fig. 7 The mixed acyclic graph generated by boxes of equals size covering of the 1+1 dimensional Minkowski spacetime

Neumann algebra associated to the spacetime region $V \in \mathscr{T}$. Denote by $\sigma(V)$ the sigma-algebra of the projections of $\mathscr{N}(V)$. Let the random variable associated to V be any Borel-measurable function from $\sigma(V)$ to $\mathscr{B}(\mathbb{R})$. Any state ϕ will then define a probability measure p on $\sigma(V)$ for any $V \in \mathscr{T}$ and, due to isotony of the net, also for any V which is a *finite* union of regions in $\mathscr{T}$. (Note that $\sigma(\mathscr{M})$ may not be a sigma-algebra since the quasilocal algebra $\mathscr{A}$ is not necessarily a von Neumann algebra, so it may not contain projections.)

In sum, any finite set of regions of a local classical theory $\{\mathscr{N}(V), V \in \mathscr{K}\}$ generated by a globally hyperbolic partition of $\mathscr{M}$ defines a pair $(\mathscr{G}, \mathscr{V})$. For certain specific coverings $\mathscr{G}$ will be a directed acyclic graph; in general, however, it will be a mixed graph.

Now, we state and prove the main claim of the paper.

Proposition 1 *Let G be a directed/mixed acyclic graph constructed from a local classical theory $\{\mathscr{N}(V), V \in \mathscr{K}\}$ where $\mathscr{K}$ is generated by a partition $\mathscr{T}$ of $\mathscr{M}$. Suppose that $\{\mathscr{N}(V), V \in \mathscr{K}\}$ is locally causal in the sense of Definition 1. For any V_A and V_B spacelike separated spacetime regions, call a set $\{V_i\} \subset \mathscr{K}$ a shielder-off set of regions for V_A if $\cup_i V_i$ is a shielder-off region for V_A characterized by the criteria L_1, L_2 and L_3^C. Then, any shielder-off set $\{V_i\}$ d-separates/m-separates V_A from V_B.*

Proof To prove Proposition 1, we have to show that $\{V_i\}$ blocks every path connecting V_A and V_B that is on every path there is at least one non-collider in $\{V_i\}$ or there is at least one collider V_E such that $V_E \notin Anc(\{V_i\}) \vee \{V_i\}$.

First consider those paths that contain no colliders. These paths need to pass through the set of common ancestors, $Anc(V_A) \wedge Anc(V_B)$. But due to L_3^C, the shielder-off set $\{V_i\}$ blocks every path connecting V_A and $Anc(V_A) \wedge Anc(V_B)$. Hence, $\{V_i\}$ blocks all the paths which contain no colliders.

So there remain only those paths to be blocked which contain at least one collider. There are two types of such paths: paths avoiding $\{V_i\}$ and path crossing $\{V_i\}$.

Consider first the paths avoiding $\{V_i\}$. Define the set

$$A^{cut} := (Anc(A) \vee A) \setminus (Anc(\{V_i\}) \vee \{V_i\})$$

Now, it is easy to see that no path which starts from V_A, avoids $\{V_i\}$ and contains only non-colliders can leave $Des(A^{cut})$. However, $V_B \notin Des(A^{cut})$, otherwise L_3^C would not hold. Hence, the path connecting V_A and V_B need to contain at least one collider $V_E \in Des(A^{cut})$. But $Des(A^{cut}) \wedge (Anc(\{V_i\}) \vee \{V_i\}) = \emptyset$, hence $V_E \notin Anc(\{V_i\}) \vee \{V_i\}$. Thus, the path is blocked by $\{V_i\}$.

Consider now the paths crossing $\{V_i\}$. Let $\mathscr{P} = (V_A, \ldots, V_D, V_E, \ldots, V_B)$ a path connecting V_A and V_B such that V_D is the last vertex before the path enters $\{V_i\}$ and V_E is the first vertex on the path which already is in $\{V_i\}$. We show that V_E cannot be a collider.

To see this, note that V_D has to be in A^{cut}, otherwise the subpath $\mathscr{P} = (V_A, \ldots, V_D)$ would contain at least one collider in $Des(A^{cut})$ and hence would be blocked. Now, suppose, contrary to our claim, that V_E is a collider. Then there is an arrow pointing from V_D to V_E. Hence, $V_D \in Anc(\{V_i\})$. But if V_D is both in A^{cut} and also in $Anc(\{V_i\})$, then $\{V_i\}$ cannot be a shielder-off set. Contradiction. Thus, V_E is a non-collider in $\{V_i\}$ and the path is blocked.

In sum, $\{V_i\}$ blocks every path connecting V_A and V_B, that is $\{V_i\}$ d-separates V_A from V_B. ■

The converse of Proposition 1 is not true: d-separating sets are not necessarily shielder-off sets. Reference [35] list algorithms to find the so-called *minimal d-separating sets* for two random variables A and B, that is sets that are d-separating but taking away any vertex from the set they will cease to be d-separating. It turns out that any minimal d-separating set is sitting in the *union* of the ancestors of A and B (including also A and B), $Anc(A) \vee Anc(B) \vee A \vee B$. However, a minimal d-separating set need not satisfy relations L_1, L_2 and L_3^C. For example the sets $\mathscr{D}$, $\mathscr{D}'$ and $\mathscr{D}''$ in Fig. 8 are all minimal d-separating sets but not shielder-off regions for A relative to B.

At any event, shielder-off regions are d-separating, and this was to be shown in this paper.

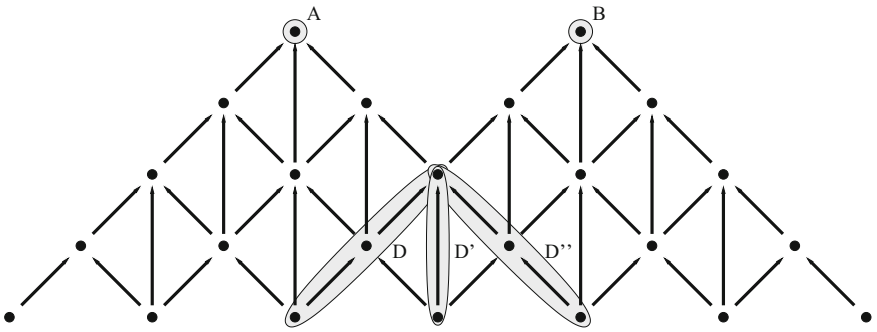

Fig. 8 Minimal d-separating but not shielder-off regions

5 Conclusions

The aim of the paper was to motivate Bell's definition of local causality by means of Bayesian networks. To this aim, first we constructed a causal graph from the covering collection of a spacetime. In certain cases the graph was a directed acyclic graph, in other cases only a mixed acyclic graph. Similarly, we have associated random variables to the local algebras of a local physical theory. By this move shielder-off regions turned out be specific d-separation (m-separating) sets on the causal graph. Hence, Bell's definition of local causality requiring that spacelike separated events should be screened-off by events in a shielder-off region turned out to be a d-separation criterion.

Acknowledgements I wish to thank Péter Vecsernyés for valuable discussions. This work has been supported by the Hungarian Scientific Research Fund, OTKA K-115593 and by the Bilateral Mobility Grant of the Hungarian and Polish Academies of Sciences, NM-104/2014.

References

1. J. S. Bell, "La nouvelle cuisine," in: J. Sarlemijn and P. Kroes (eds.), *Between Science and Technology*, Elsevier, (1990); reprinted in (Bell, 2004, 232-248).
2. J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, (Cambridge: Cambridge University Press, 2004).
3. J. Butterfield, "Stochastic Einstein Locality Revisited," *Brit. J. Phil. Sci.*, **58**, 805-867, (2007).
4. J. Earman and G. Valente, "Relativistic causality in algebraic quantum field theory," *Int. Stud. Phil. Sci.*, **28 (1)**, 1-48 (2014).

5. C. Glymour, "Markov properties and quantum experiments," in W. Demopoulos and I. Pitowsky (eds.) *Physical Theory and its Interpretation*, (Springer, 117-126, 2006).
6. C. Glymour, R. Scheines and P. Spirtes, "Causation, Prediction, and Search," (Cambridge: The MIT Press, 2000).
7. R. Haag, *Local quantum physics*, (Heidelberg: Springer Verlag, 1992).
8. H. Halvorson, "Algebraic quantum field theory," in J. Butterfield, J. Earman (eds.), *Philosophy of Physics, Vol. I*, Elsevier, Amsterdam, 731-922 (2007).
9. D. M. Hausman and J. Woodward, "Independence, invariance and the causal Markov condition," *Brit. J. Phil Sci.*, **50** (4), 521–583 (1999).
10. J. Henson, "Comparing causality principles," *Stud. Hist. Phil. Mod. Phys.*, **36**, 519-543 (2005).
11. J. Henson, "Confounding causality principles: Comment on Rédei and San Pedro's "Distinguishing causality principles"," *Stud. Hist. Phil. Mod. Phys.*, **44**, 17-19 (2013a).
12. J. Henson, "Non-separability does not relieve the problem of Bell's theorem," *Found. Phys.*, **43**, 1008-1038 (2013b).
13. G. Hofer-Szabó, "Local causality and complete specification: a reply to Seevinck and Uffink," in U. Mäki, I. Votsis, S. Ruphy and G. Schurz (eds.) *Recent Developments in the Philosophy of Science: EPSA13 Helsinki*, Springer Verlag, 209-226 (2015a).
14. G. Hofer-Szabó, "Relating Bell's local causality to the Causal Markov Condition," *Found. Phys.*, **45(9)**, 1110-1136 (2015b).
15. G. Hofer-Szabó and P. Vecsernyés, "Reichenbach's Common Cause Principle in AQFT with locally finite degrees of freedom," *Found. Phys.*, **42**, 241-255 (2012a).
16. G. Hofer-Szabó and P. Vecsernyés, "Noncommuting local common causes for correlations violating the Clauser–Horne inequality," *J. Math. Phys.*, **53**, 12230 (2012b).
17. G. Hofer-Szabó and P. Vecsernyés, "Noncommutative Common Cause Principles in AQFT," *J. Math. Phys.*, **54**, 042301 (2013a).
18. G. Hofer-Szabó and P. Vecsernyés, "Bell inequality and common causal explanation in algebraic quantum field theory," *Stud. Hist. Phil. Mod. Phys.*, **44 (4)**, 404-416 (2013b).
19. G. Hofer-Szabó and P. Vecsernyés, "On the concept of Bell's local causality in local classical and quantum theory," *J. Math. Phys*, **56**, 032303 (2015).
20. G. Hofer-Szabó and P. Vecsernyés, "A generalized definition of Bell's local causality," *Synthese* **193(10)**, 3195-3207 (2016).
21. G. Hofer-Szabó, M. Rédei and L. E. Szabó, *The Principle of the Common Cause*, (Cambridge: Cambridge University Press, 2013).
22. T. Maudlin, "What Bell did," *J. Phys. A: Math. Theor.*, **47**, 424010 (2014).
23. T. Norsen, "Local causality and Completeness: Bell vs. Jarrett," *Found. Phys.*, **39**, 273 (2009).
24. T. Norsen, "J.S. Bell's concept of local causality," *Am. J. Phys*, **79**, 12, (2011).
25. J. Pearl, "Causality: Models, Reasoning, and Inference," Cambridge: (Cambridge University Press, 2000).
26. M. Rédei, "Reichenbach's Common Cause Principle and quantum field theory," *Found. Phys.*, **27**, 1309-1321 (1997).
27. M. Rédei, "A categorial approach to relativistic locality," *Stud. Hist. Phil. Mod. Phys.*, **48**, 137-146 (2014).
28. M. Rédei and I. San Pedro, "Distinguishing causality principles," *Stud. Hist. Phil. Mod. Phys.*, **43**, 84-89 (2012).
29. M. Rédei and J. S. Summers, "Local primitive causality and the Common Cause Principle in quantum field theory," *Found. Phys.*, **32**, 335-355 (2002).
30. T. S. Richardson and P. Spirtes, "Ancestral graph Markov models," *Ann. Statist.* **30**, 962-1030 (2002).
31. K. Sadeghi and S. Lauritzen, "Markov properties for mixed graphs," *Bernoulli.* **20/2**, 676-696 (2014).
32. M. P. Seevinck and J. Uffink, "Not throwing our the baby with the bathwater: Bell's condition of local causality mathematically 'sharp and clean'," in: Dieks, D.; Gonzalez, W.J.; Hartmann, S.; Uebel, Th.; Weber, M. (eds.) *Explanation, Prediction, and Confirmation The Philosophy of Science in a European Perspective*, Volume 2, 425-450 (2011).

33. M. Suárez, “Interventions and causality in quantum mechanics,” *Erkenntnis*, **78**, 199-213 (2013).
34. M. Suárez and I. San Pedro “Causal Markov, robustness and the quantum correlations,” in M. Suarez (ed.), *Probabilities, causes and propensities in physics*, 173-193. Synthese Library, 347, (Dordrecht: Springer, 2011)
35. J. Tian, A. Paz, and J. Pearl, “Finding minimal d-separating sets,” *UCLA Cognitive Systems Laboratory, Technical Report* (R-254), (1998).