Impossibility of Superluminal Signaling in Minkowski Spacetime Does Not Rule Out Causal Loops

V. Vilasini[1,2] and Roger Colbeck[2]
[1]Institute for Theoretical Physics, ETH Zurich, 8093 Zürich, Switzerland
[2]Department of Mathematics, University of York, Heslington, York YO10 5DD, United Kingdom

(Received 6 January 2022; revised 20 May 2022; accepted 24 June 2022; published 6 September 2022)

Causality is fundamental to science, but it appears in several different forms. One is relativistic causality, which is tied to a spacetime structure and forbids signaling outside the future. A second is an operational notion of causation that considers the flow of information between physical systems and interventions on them. In [V. Vilasini and R. Colbeck, General framework for cyclic and fine-tuned causal models and their compatibility with space-time, Phys. Rev. A **106**, 032204 (2022).], we propose a framework for characterizing when a causal model can coexist with relativistic principles such as no superluminal signaling, while allowing for cyclic and nonclassical causal influences and the possibility of causation without signaling. In a theory without superluminal causation, both superluminal signaling and causal loops are not possible in Minkowski spacetime. Here we demonstrate that if we only forbid superluminal signaling, superluminal causation remains possible and show the mathematical possibility of causal loops that can be embedded in a Minkowski spacetime without leading to superluminal signaling. The existence of such loops in the given spacetime could in principle be operationally verified using interventions. This establishes that the physical principle of no superluminal signaling is not by itself sufficient to rule out causal loops between Minkowski spacetime events. Interestingly, the conditions required to rule out causal loops in a spacetime depend on the dimension. Whether such loops are possible in three spatial dimensions remains an important open question.

DOI: 10.1103/PhysRevLett.129.110401

Introduction.—Understanding cause-effect relations is central to the scientific method, yet there are several inequivalent notions of causality. Often, it is defined with respect to a background spacetime structure, after which causal structure and spacetime structure are treated synonymously. An alternative is to define causality operationally and independently of spacetime. One way to do this is through causal models, which are based on intervening on physical systems and analyzing the resulting correlations [1,2]. This is the approach we take here. Causal models have been extensively applied to situations involving classical variables, being used, for instance, for medical testing [3,4], economic predictions [1,5], and machine learning [6–8].

Bell's theorem [9] demonstrates that classical causal models cannot explain quantum correlations within the causal structure that is naturally associated with a Bell experiment [10]. This has fueled several approaches for providing causal explanations to quantum and more general nonclassical correlations. One approach is to develop causal models [10–25] that go beyond classical random variables and allow quantum or even post-quantum systems [26,27] to be causes. Other approaches suggest modifying the causal structure itself without necessarily considering nonclassical causes, such as allowing for additional causal influences that go outside the future light cone (e.g., nonlocal hidden variable theories [28]) or against the direction of time (retrocausality [29]). Such causal influences must remain hidden in order to prevent superluminal signaling at the observed level. More radical approaches lean towards giving up the standard understanding of causation as being acyclic, and replacing it with a suitable notion of logical consistency [30–32]. While these alternatives correspond to different descriptions of the underlying causal model, they are all compatible with the impossibility of superluminal signaling in Minkowski spacetime [33].

In our associated paper [34], we have developed a general framework for causation that can describe nonclassical and cyclic causal influences as well as causation without signaling. We do so by keeping the direction of causation and the order of events in spacetime distinct. The former is modeled operationally and we characterize when the two are *compatible* with each other, i.e., when we can assign spacetime locations to the random variables in the causal model without leading to signaling outside the spacetime's future (we call this assignment an embedding). Here we ask: does the ability to compatibly embed a causal model in an acyclic spacetime (such as Minkowski spacetime) imply that the operational predictions of the causal model can be reproduced within an acyclic causal structure? If so, it would not be necessary to consider cyclic causation.

Within relativistic physics, causal influences are taken to flow within the light cone, making both causal loops and

superluminal signaling impossible in Minkowski spacetime. Here we relax this assumption, and require only that observable signaling is limited by the light cone structure. In scenarios in which there is causation without correlation (i.e., a *fine-tuned* influence), answering the above question is more challenging since fine-tuned influences can act outside the lightcone without leading to signaling. Our framework allows treatment of both correlations and interventions on physical systems in general scenarios with cyclic causal influences. It can model fine-tuned causal influences as well as latent quantum and post-quantum causes, and be used to characterize conditions under which a causal model can be compatibly embedded in a spacetime. Here, we apply this framework to demonstrate the mathematical possibility of causal loops between events in $(1+1)$-Minkowski spacetime where these loops can be operationally detected without superluminal signaling, providing an explicit example. We also show that the observable predictions of this loop cannot be reproduced in any acyclic causal model, answering the aforementioned question in the negative. Interestingly, the same example fails to be embeddable in $(3+1)$-Minkowski spacetime. In the associated paper [34], we have characterized a large class of operationally detectable causal loops within our framework. Whether any of these may be embeddable in $(3+1)$ dimensions remains an interesting open question.

Fine-tuned causal explanations are often undesirable as fundamental explanations of physical phenomena [35], but can be crucial in practical information processing tasks. For instance, cryptographic protocols (such as the one time pad) rely on fine-tuned correlations, and the security of relativistic cryptographic protocols [36,37] combines both relativistic notions of causality and information-theoretic concepts. Hence, understanding the extent to which compatibility with Minkowski spacetime restricts the possible operational causal models also has practical implications.

Causal models: Correlations, interventions and fine-tuning.—We begin by reviewing the essentials of the causal modeling framework (see Ref. [34] for details). A *causal structure* is modeled as a directed graph $\mathcal{G}$ whose nodes correspond to observed or unobserved systems and directed edges $\rightsquigarrow$ denote causation between these systems [38]. The set of observed nodes (denoted S_{obs}) comprises classical random variables (RVs) such as measurement settings or outcomes, while unobserved nodes may be classical, quantum, or post-quantum systems (such as those from a generalized probabilistic theory).

Implicitly the nodes are associated with causal mechanisms that specify how information from their incoming edges is processed. For instance, in the classical case this processing corresponds to a function f_{N_i} from the parents of the node and possibly an additional, parentless [39] RV E_{N_i} (to allow for situations where N_i is not deterministically dependent on its parents) to the node variable N_i itself. "A is a direct cause of B" then corresponds to the function f_B having A as a (nontrivial) argument. [In the nonclassical case, these functions would be replaced by valid maps between systems in the theory (e.g., CPTP maps in quantum theory)].

Often these causal mechanisms are not explicitly known and hence our treatment has to work at the level of observed probability distributions rather than the causal mechanisms. The causal structure imposes constraints on the possible distributions that may arise over the observed nodes. One set of such constraints can be expressed using the notion of *d-separation*. For two disjoint subsets X and Y of observed nodes S_{obs} of a causal structure $\mathcal{G}$, X, and Y are said to be d separated, denoted $(X\perp^d Y)_{\mathcal{G}}$, if there are no directed paths between variables in X and Y and if no variables in X and Y have common ancestors in $\mathcal{G}$. More generally, d-separation $(X\perp^d Y|Z)_{\mathcal{G}}$ is defined for three disjoint subsets X, Y, and Z of the observed nodes (see Sec. I of the Supplemental Material [40] for details).

We say that a distribution $P_{\mathcal{G}}(S_{\text{obs}})$ satisfies the d-separation property with respect to a causal structure $\mathcal{G}$ if whenever we have d-separation between observed nodes in $\mathcal{G}$, then we have a corresponding conditional independence in the observed distribution $P_{\mathcal{G}}(S_{\text{obs}})$, i.e., $(X\perp^d Y|Z)_{\mathcal{G}}$ implies $P_{\mathcal{G}}(XY|Z) = P_{\mathcal{G}}(X|Z)P_{\mathcal{G}}(Y|Z)$. If the converse holds for all disjoint subsets of observed nodes, then the distribution is said to be *faithful* or, equivalently, not *fine-tuned*. In the present Letter, we adopt a minimal definition of a causal model which corresponds to a directed graph $\mathcal{G}$ and an observed distribution $P_{\mathcal{G}}(S_{\text{obs}})$ that satisfies the d-separation property with respect to $\mathcal{G}$ [45]. In this Letter we allow fine-tuned causal models in which conditional independences can occur in $P_{\mathcal{G}}(S_{\text{obs}})$ without the corresponding d-separation in $\mathcal{G}$ (i.e., we allow causation without correlation).

In cases where we know the causal mechanisms, a causal model can be specified by a causal structure $\mathcal{G}^C$, causal mechanisms $\{f_{N_i}, P_{\mathcal{G}^C}(E_{N_i})\}_i$, and an associated observed distribution $P_{\mathcal{G}^C}(S_{\text{obs}})$ for $S_{\text{obs}} \subseteq \{N_i\}_i$. Here the observed distribution must be consistent with relationships specified by the causal mechanisms. Methods for modeling operational causal structures beyond the classical case are proposed in our associated paper [34]. In the present Letter, we restrict to the classical case, which suffices to illustrate our main claims.

So far, we have only discussed the possible correlations that fit with a causal structure. Inferring causation requires more, and the concept of an intervention has been introduced to deal with this [1]. If intervening on X changes the distribution on Y, then we can deduce that X is a cause of Y.

In the case where X is parentless in a causal structure $\mathcal{G}$, correlation between X and another variable Y, i.e., $P_{\mathcal{G}}(Y|X) \neq P_{\mathcal{G}}(Y)$ suffices to conclude that X is a cause of Y. More generally, an intervention on X corresponds to forcing X to take a certain value, x, irrespective of its parents. This results in a postintervention causal structure

$\mathcal{G}_{\mathrm{do}(X)}$ obtained from $\mathcal{G}$ by removing all the incoming edges to X, while maintaining the causal mechanisms for $\mathcal{G}_{\mathrm{do}(X)}$ from $\mathcal{G}$, except the mechanism for the intervened node X, which is replaced by $X = x$. We say that X *affects* Y if there exist values x and y such that

$$P_{\mathcal{G}_{\mathrm{do}(X)}}(Y = y|X = x) \neq P_{\mathcal{G}}(Y = y). \tag{1}$$

In general $P_{\mathcal{G}_{\mathrm{do}(X)}}(Y|X)$ is not the same as $P_{\mathcal{G}}(Y|X)$ (but they are equal when X is parentless in $\mathcal{G}$).

Acyclic and cyclic causal models embedded in Minkowski spacetime.—The affects relation defined in (1) naturally extends to joint interventions on sets of observed nodes. In the presence of fine-tuning, a set of RVs can jointly affect another set without any affects relations between individual pairs of elements in the sets. We will illustrate three such examples here and use these to establish our main claim. All three causal models have observed nodes $S_{\mathrm{obs}} := \{A, B, C\}$ which we will take to be classical and binary. The same observed correlations will be used in all three cases. The distinguishing feature at the observed level will be the affects relations which model what happens under intervention. The causal structures and compatible spacetime embeddings for the three examples are given in Fig. 1.

Example 1 (The one-time pad) Consider the causal structure of Fig. 1(a), $\mathcal{G}^{\mathrm{OTP}}$ with the causal mechanisms $A = E_A$, $C = E_C$, and $B = A \oplus C$, with E_A and E_C being independent and uniformly distributed. Then, B is uniformly distributed and we have $P_{\mathcal{G}^{\mathrm{OTP}}}(B|AC) \neq P_{\mathcal{G}^{\mathrm{OTP}}}(B)$, $P_{\mathcal{G}^{\mathrm{OTP}}}(B|A) = P_{\mathcal{G}^{\mathrm{OTP}}}(B)$, and $P_{\mathcal{G}^{\mathrm{OTP}}}(B|C) = P_{\mathcal{G}^{\mathrm{OTP}}}(B)$. Since A and C are parentless, these statements are equivalent to $\{A, C\}$ affects B, A does not affect B, and C does not affect B [46]. A Minkowski spacetime embedding of the RVs that does not enable any signaling outside the future is one in which B is assigned a spacetime location in the joint future of the spacetime locations of A and C, as illustrated in Fig. 1(a).

Example 2 (A simplified jamming scenario; cf. [47]) Consider the causal structure of Fig. 1(b), $\mathcal{G}^{\mathrm{jam}}$ with the observed nodes $\{A, B, C\}$ and a classical unobserved node Λ. Suppose we have the causal mechanisms $\Lambda = E_\Lambda$, $A = \Lambda$, $B = E_B$, and $C = B \oplus \Lambda$ with E_Λ and E_B uniformly distributed. This gives the same observed correlations as the previous example, with $B = A \oplus C$ and with all observed variables uniformly distributed. Additionally, $P_{\mathcal{G}^{\mathrm{jam}}}(AC|B) \neq P_{\mathcal{G}^{\mathrm{jam}}}(AC)$, $P_{\mathcal{G}^{\mathrm{jam}}}(A|B) = P_{\mathcal{G}^{\mathrm{jam}}}(A)$ and $P_{\mathcal{G}^{\mathrm{jam}}}(C|B) = P_{\mathcal{G}^{\mathrm{jam}}}(C)$. Since B is parentless, this is equivalent to B affects $\{A, C\}$, B does not affect A, and B does not affect C. A compatible Minkowski spacetime embedding in this case requires that the joint future of the spacetime locations of A and C is contained entirely within the spacetime future of B, as shown in Fig. 1(b). This is because B affects $\{A, C\}$ can only be verified when A and C are brought together, which is possible only in their joint future. Since we have no pairwise affects relations, there is no pairwise signaling between the RVs and no RV is required to be in the future of any other. Note that the causal influence $B \rightsquigarrow C$ is superluminal, even though there is no superluminal signaling.

Example 3 (A fine-tuned causal loop) Consider the causal structure of Fig. 1(c), $\mathcal{G}^{\mathrm{loop}}$, with the same observed and unobserved nodes as the previous example, but with the causal mechanisms $\Lambda = E_\Lambda$, $A = \Lambda$, $C = B \oplus \Lambda$, and $B = A \oplus C$, with E_Λ uniformly distributed. Note that these causal mechanisms do not admit a unique solution. Nevertheless, in Sec. II of the Supplemental Material [40] we apply a method proposed in [34] to uniquely determine the observed distribution based on these mechanisms, and show that the same distribution as the previous two examples is obtained. The effect of interventions in this case cannot be directly inferred from the observed correlations since none of the observed nodes are parentless. The postintervention causal structure $\mathcal{G}^{\mathrm{loop}}_{\mathrm{do}(AC)}$ is identical to $\mathcal{G}^{\mathrm{OTP}}$ and the postintervention causal structure $\mathcal{G}^{\mathrm{loop}}_{\mathrm{do}(B)}$ is identical to $\mathcal{G}^{\mathrm{jam}}$. Applying (1), we find $\{A, C\}$ affects B *and* B

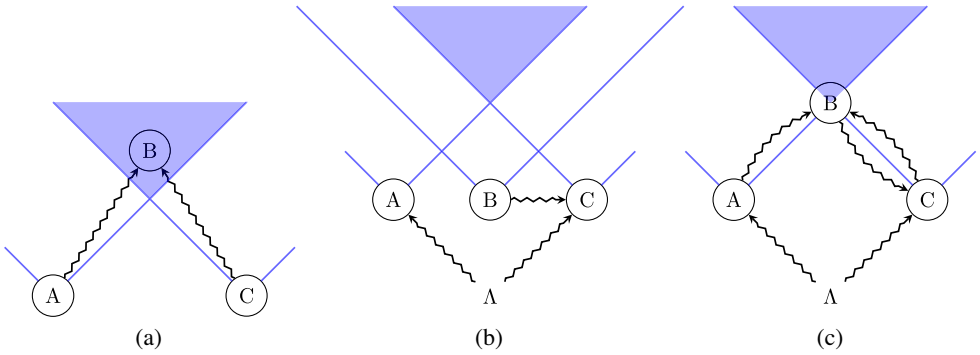

FIG. 1. Three examples of causal models and their compatible embeddings in $(1 + 1)$-Minkowski spacetime. In each case, the operational causal structure associated with the model is given in black, circled variables are observed nodes, while uncircled ones are not and the black arrows denote causation. Space-time information is given in blue with time along the vertical and space along the horizontal axis. The solid lines represent lightlike surfaces and the shaded region corresponds to the joint future of A and C in all cases.

affects $\{A, C\}$ and no pairwise affects relations between A, B, and C. The spacetime embedding must satisfy the compatibility conditions of both the previous examples, and Fig. 1(c) illustrates an embedding with these properties, i.e., this causal loop can be compatibly embedded in $(1+1)$-Minkowski spacetime.

All three examples above lead to the same correlations $B = A \oplus C$, where B is correlated jointly with A and C but not individually, hence there is fine-tuning. Correlation between B and $\{A, C\}$ implies (by the d-separation property) that B and $\{A, C\}$ must be d connected, i.e., B is d connected with A and/or C. However, the RVs A, B, and C are pairwise uncorrelated so there must be a pair of variables that are d connected and yet independent, which constitutes fine-tuning (these independences disappear for small changes in the distribution of one of the variables, e.g., if Λ is nonuniform in examples 2 and 3).

The causal loop of example 3 exhibits many curious features. It is an operationally detectable causal loop, i.e., any causal model that gives rise to the affects relations of this example must necessarily be associated with a cyclic causal structure. This is proven in Sec. III of the Supplemental Material [40], but the intuition is relatively simple: Consider three parties in possession of the 3 observed RVs A, B, and C and two types of experiment: (E1) Alice and Charlie perform all possible interventions on A and C and Bob observes B without intervening; (E2) Bob performs all possible interventions on B while Alice and Charlie observe A and C without intervening. After both experiments the parties can get together to verify whether $B = A \oplus C$ holds. Here, "all possible interventions" on a variable corresponds to running through all possible values of that variable, setting these independently of their parents [see (1)], and collecting statistics for each choice. These statistics differ between the three causal models, as they have different sets of affects relations. These interventions do not enable the parties to signal outside the spacetime future as the affects relations of all these causal models are compatible with the given spacetime embedding. They nevertheless allow the parties to operationally verify the existence of a causal loop as we show in the Supplemental Material [40].

In example 3, experiment E1 shows that A and C are causes of B while E2 shows that B is a cause of at least one of A and C. Given the spacetime embedding from Fig. 1(c), these interventions would enable the agents to operationally detect retrocausation. By contrast, for the first model, E2 would correspond to a post-intervention scenario with no edges and therefore lead to no correlations between the RVs, while in the second model, E1 would also lead to no correlations. In other words, these two experiments enable the parties to operationally distinguish between the causal models of Examples 1–3 in spite of them having the same observed correlations.

The mathematical possibility of an operationally detectable causal loop being embedded in Minkowski spacetime without signaling necessarily involves fine-tuning (see Ref. [34]) since in the absence of fine-tuning, signaling and causation are equivalent. Note that in example 3 the locations of the random variables in Minkowski spacetime have to be carefully chosen to allow compatibility: B and C must be lightlike separated, and arbitrarily small adjustments to the location of either B or C will remove compatibility. This requires B to be embedded exactly at the earliest location in the joint future of A and C. In a sense this is another kind of fine-tuning, but at the level of the spacetime embedding. Furthermore, such an embedding is not possible in $(3+1)$-Minkowski spacetime because the intersection of two light cones is not itself a cone and, consequently, there does not exist a frame-independent earliest location in the joint future of two given points, unlike in $(1+1)$-D [48].

Beyond these examples, a causal model may give rise to a complicated set of affects relations between various subsets of the observed nodes, making characterizing compatibility with a spacetime a more complicated task. For instance, if A affects B and $\{A, C\}$ affects B but C does not affect B, does B have to be embedded in the joint future of A and C, or only in the future of A? The novel causal modeling concept required to answer such questions in fine-tuned models is the notion of a *higher-order affects relation*, which we introduce in our framework [34]. Here we have illustrated one example, but there is a general class of physical theories involving cyclic causal structures that are compatible with no superluminal signaling in Minkowski spacetime as discussed in our associated paper [34].

Discussion and outlook.—In relativistic physics, a causal influence from one spacetime location to another implies that the latter is in the future light cone of the former. This means that within relativistic physics, it is not possible to have causal loops or closed timelike curves in spacetimes whose light-cone structures form a partial order. On the other hand, the approach adopted here, where causal structure and spacetime structure are distinct notions, only requires observable signaling (and not causal influences) to stay within the spacetime future for compatibility. Maintaining a clear separation between causation and spacetime structure and characterizing their interdependence is useful for considering formulations of physics without a fixed background spacetime structure (e.g., in quantum gravity [49,50]), as well as practical information processing tasks in spacetime.

We note that a previous work [51] proposes a set of necessary and sufficient conditions for ruling out causal loops in Bell-type scenarios. Our framework [34] identifies implicit assumptions in this claim. In particular, one of the claims in [51] assumes that no superluminal signaling in Minkowski spacetime rules out causal loops, which we have shown does not hold in general.

We have established that the principle of no superluminal signaling in Minkowski spacetime alone is insufficient to

rule out causal loops, through an explicit construction in $(1+1)$ dimensions. We also found that the conditions for ruling out causal loops can depend on the spacetime dimension. Previous works have established that logical consistency [52,53] or familiar quantum properties such as linearity, no-cloning [54], etc. are not sufficient. Finding underlying principles that can do so remains an interesting open problem, a pertinent question being whether the principle of no superluminal signaling rules out causal loops in the case of $(3+1)$-Minkowski spacetime.

V. V. acknowledges support from the Department of Mathematics, University of York and the ETH Postdoctoral Fellowship from ETH Zürich.

[1] J. Pearl, *Causality: Models, Reasoning, and Inference*, 2nd ed. (Cambridge University Press, Cambridge, England, 2009).

[2] P. Spirtes, C. N. Glymour, and R. Scheines, *Causation, Prediction, and Search*, 2nd ed. (The MIT Press, Cambridge, MA, 2001).

[3] S. Kleinberg and G. Hripcsak, A review of causal inference for biomedical informatics, J. Biomed. Inf. **44**, 1102 (2011).

[4] Y. Raita, C. A. Camargo, L. Liang, and K. Hasegawa, Big data, data science, and causal inference: a primer for clinicians, Front. Med. **8**, 998 (2021).

[5] P. Spirtes, Graphical models, causal inference, and econometric models. J. Econ. Methodol. **12**, 3 (2005).

[6] Maya L. Petersen and M. J. van der Laan, Causal models and learning from data, Epidemiology **25**, 418 (2014).

[7] S. Arti, I. Hidayah, and S. S. Kusumawardani, Research trend of causal machine learning method: A literature review, Int. J. Inf. Dev. **9**, 111 (2020).

[8] T. Liu, L. Ungar, and K. Kording, Quantifying causality in data science with quasi-experiments, Nat. Comput. Sci. **1**, 24 (2021).

[9] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, England, 1987).

[10] C. J. Wood and R. W. Spekkens, The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning, New J. Phys. **17**, 033002 (2015).

[11] R. R. Tucci, Quantum Bayesian nets, Int. J. Mod. Phys. B **09**, 295 (1995).

[12] M. S. Leifer, Quantum dynamics as an analog of conditional probability, Phys. Rev. A **74**, 042310 (2006).

[13] K. B. Laskey, Quantum causal networks, arXiv:0710.1200.

[14] M. Leifer and D. Poulin, Quantum graphical models and belief propagation, Ann. Phys. (Amsterdam) **323**, 1899 (2008).

[15] M. S. Leifer and R. W. Spekkens, Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference, Phys. Rev. A **88**, 052130 (2013).

[16] J. Henson, R. Lal, and M. Pusey, Theory-independent limits on correlations from generalized Bayesian networks, New J. Phys. **16**, 113043 (2014).

[17] J. Pienaar and Č. Brukner, A graph-separation theorem for quantum causal models, New J. Phys. **17**, 073020 (2015).

[18] K. Ried, M. Agnew, L. Vermeyden, D. Janzing, R. W. Spekkens, and K. J. Resch, A quantum advantage for inferring causal structure, Nat. Phys. **11**, 414 (2015).

[19] F. Costa and S. Shrapnel, Quantum causal modelling, New J. Phys. **18**, 063032 (2016).

[20] T. Fritz, Beyond Bells theorem II: Scenarios with arbitrary causal structure, Commun. Math. Phys. **341**, 391 (2015).

[21] J.-M. A. Allen, J. Barrett, D. C. Horsman, C. M. Lee, and R. W. Spekkens, Quantum Common Causes and Quantum Causal Models, Phys. Rev. X **7**, 031021 (2017).

[22] M. Weilenmann and R. Colbeck, Analysing causal structures with entropy, Proc. R. Soc. A **473**, 2207 (2017).

[23] J. Barrett, R. Lorenz, and O. Oreshkov, Quantum causal models, arXiv:1906.10726.

[24] J. Pienaar, Quantum causal models via quantum Bayesianism, Phys. Rev. A **101**, 012104 (2020).

[25] D. Schmid, J. H. Selby, and R. W. Spekkens, Unscrambling the omelette of causation and inference: The framework of causal-inferential theories, arXiv:2009.03297.

[26] J. Barrett, Information processing in generalized probabilistic theories, Phys. Rev. A **75**, 032304 (2007).

[27] M. Weilenmann and R. Colbeck, Analysing causal structures in generalised probabilistic theories, Quantum **4**, 236 (2020).

[28] D. Bohm, A suggested interpretation of the quantum theory in terms of, "hidden" variables. I, Phys. Rev. **85**, 166 (1952).

[29] K. B. Wharton and N. Argaman, Colloquium: Bell's theorem and locally mediated reformulations of quantum mechanics, Rev. Mod. Phys. **92**, 021002 (2020).

[30] Ä. Baumeler and S. Wolf, The space of logically consistent classical processes without causal order, New J. Phys. **18**, 013036 (2016).

[31] Ä. Baumeler and S. Wolf, Causality-complexity-consistency: Can space-time be based on logic and computation?, *Tutorials, Schools, and Workshops in the Mathematical Sciences* (Springer International Publishing, 2017), pp. 69–101, 10.1007/978-3-319-68655-4_6.

[32] Ä. Baumeler, J. Degorre, and S. Wolf, Bell correlations and the common future, *Quantum Foundations, Probability and Information* (2018), 10.1007/978-3-319-74971-6_18.

[33] We reference Minkowski space-time here for concreteness, but our results also apply to any other partially ordered space-time.

[34] V. Vilasini and R. Colbeck, General framework for cyclic and fine-tuned causal models and their compatibility with space-time, Phys. Rev. A **106**, 032204 (2022).

[35] Since one then has to explain how the fine-tuning comes about.

[36] A. Kent, Secure classical bit commitment over finite channels, J. Cryptol. **18**, 313 (2005).

[37] R. Colbeck and A. Kent, Variable bias coin tossing, Phys. Rev. A **73**, 032320 (2006).

[38] Such causal structures are formally independent of space-time, although, as we will later discuss, they may be embedded in a separate space-time structure.

[39] The functional dependences are typically described in terms of an augmented causal structure that is formed from the original causal structure by including $E_{N_i} \rightsquigarrow N_i$ for each

node N_i. Note that the E_{N_i} nodes have no other edges associated with them and are parentless in the augmented graph.

[40] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.129.110401 for further details on the causal modeling framework, a method for evaluating the observed distribution of example 3, and for a proof of the operational detectability of this causal loop. The Supplemental Material includes Refs. [41–44].

[41] P. Forré and J. M. Mooij, Markov properties for graphical models with cycles and latent variables, arXiv:1710.08775.

[42] R. M. Neal, On deducing conditional independence from d-separation in causal graphs with feedback (research note), J. Artif. Intell. Res. **12**, 87 (2000).

[43] J. Barrett, R. Lorenz, and O. Oreshkov, Cyclic quantum causal models, Nat. Commun. **12**, 885 (2021).

[44] S. Bongers, P. Forr, J. Peters, and J. M. Mooij, Foundations of structural causal models with cycles and latent variables, Ann. Stat. **49**, 2885 (2021).

[45] Here, we make no assumptions about the causal mechanisms or the physical theory describing them, but instead derive the consequences of imposing d-separation at the observed level. We note that one may consider other top-down methods for describing causal structures without imposing d-separation, but this is yet to be explored.

[46] This is analogous to a one-time pad protocol where a message M (a random classical bit) is encrypted using a uniformly distributed key K to obtain a cipher text $C = M \oplus K$. Even though the message M is a cause of the cipher text C, we have $P(C|M) = P(C)$ and C contains no information about M in the absence of the key. In causal modeling terms, C and M are independent even though they are d connected; the model is fine-tuned.

[47] J. Grunhaus, S. Popescu, and D. Rohrlich, Jamming nonlocal quantum correlations, Phys. Rev. A **53**, 3781 (1996).

[48] Although we have focused here on Minkowski spacetime, the framework of our accompanying Letter [34] can be applied to consider compatible embeddings in any partially ordered spacetime. Applying this framework, one can find other examples of embeddable causal loops in spacetimes (not necessarily Minkowski), where the property of having a frame-independent earliest location in the joint future of any two given points fails. We saw that this property holds in $(1+1)$ but not in $(2+1)$ or $(3+1)$ Minkowski spacetime.

[49] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nat. Commun. **3**, 1092 (2012).

[50] M. Zych, F. Costa, I. Pikovski, and Č. Brukner, Bell's theorem for temporal order, Nat. Commun. **10**, 3772 (2019).

[51] P. Horodecki and R. Ramanathan, The relativistic causality versus no-signaling paradigm for multi-party correlations, Nat. Commun. **10**, 1701 (2019).

[52] D. Deutsch, Quantum mechanics near closed timelike lines, Phys. Rev. D **44**, 3197 (1991).

[53] S. Lloyd, L. Maccone, R. Garcia-Patron, V. Giovannetti, Y. Shikano, S. Pirandola, L. A. Rozema, A. Darabi, Y. Soudagar, L. K. Shalm, and A. M. Steinberg, Closed Timelike Curves via Postselection: Theory and Experimental Test of Consistency, Phys. Rev. Lett. **106**, 040403 (2011).

[54] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, Witnessing causal nonseparability, New J. Phys. **17**, 102001 (2015).