PHILOSOPHICAL TRANSACTIONS A

royalsocietypublishing.org/journal/rsta

Research

Cite this article: Jones M. 2019 Relating causal and probabilistic approaches to contextuality. *Phil. Trans. R. Soc. A* **377**: 20190133.
http://dx.doi.org/10.1098/rsta.2019.0133

Accepted: 4 June 2019

One contribution of 16 to a theme issue 'Contextuality and probability in quantum mechanics and beyond'.

Subject Areas:
mathematical physics

Keywords:
contextuality, probabilistic causal models, contextuality-by-default, direct influence

Author for correspondence:
Matt Jones
e-mail: mcj@colorado.edu

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.4606826.

THE ROYAL SOCIETY PUBLISHING

Relating causal and probabilistic approaches to contextuality

Matt Jones

Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, USA

MJ, 0000-0002-4469-7896

A primary goal in recent research on contextuality has been to extend this concept to cases of inconsistent connectedness, where observables have different distributions in different contexts. This article proposes a solution within the framework of probabilistic causal models, which extend hidden-variables theories, and then demonstrates an equivalence to the contextuality-by-default (CbD) framework. CbD distinguishes contextuality from direct influences of context on observables, defining the latter purely in terms of probability distributions. Here, we take a causal view of direct influences, defining direct influence within any causal model as the probability of all latent states of the system in which a change of context changes the outcome of a measurement. Model-based contextuality (M-contextuality) is then defined as the necessity of stronger direct influences to model a full system than when considered individually. For consistently connected systems, M-contextuality agrees with standard contextuality. For general systems, it is proved that M-contextuality is equivalent to the property that any model of a system must contain 'hidden influences', meaning direct influences that go in opposite directions for different latent states, or equivalently signalling between observers that carries no information. This criterion can be taken as formalizing the 'no-conspiracy' principle that has been proposed in connection with CbD. M-contextuality is then proved to be equivalent to CbD-contextuality, thus providing a new interpretation of CbD-contextuality as the non-existence of a model for a system without hidden direct influences.

This article is part of the theme issue 'Contextuality and probability in quantum mechanics and beyond'.

1. Introduction

Probabilistic contextuality describes an empirical system of measurements wherein a set of observables can be measured in different subsets in different contexts, and even though each observable has the same distribution in all contexts in which it is measured, the joint distributions of measurements within each context cannot be pieced together into a global joint distribution. The original significance lies in Bell's theorem and related results [1,2], which state that certain contextual systems cannot be explained by any local hidden-variables theory, although such systems are predicted by quantum mechanics and have been experimentally confirmed [3,4].

The standard formulation of contextuality applies only to cases where the distribution of each observable is identical across contexts. This property of a system of measurements is called consistent connectedness [5], marginal selectivity [6,7] or no-disturbance [8]. Dzhafarov and Kujala have made persuasive arguments for extending contextuality to inconsistently connected systems, including the fact that real experiments never eliminate all sources of contamination, and that sample frequencies in finite datasets will generally not be equal even if the true generating probabilities are [9–11]. Beyond these pragmatic considerations, it is of interest to know whether contextuality can be usefully defined in cases where the distributions of observables are truly different across contexts. This might enable contextuality analysis to be applied to other domains, such as human cognition or behaviour [12–15].

The contextuality-by-default (CbD) theory offers one approach for defining contextuality for inconsistently connected systems [5,11,16]. CbD treats measurements of each observable in different contexts as different random variables, by default. It then asks whether the distributions of random variables in each context are compatible with a global distribution in which all variables for each observable are made as equal as possible, in a rigorous sense based on probabilistic couplings [17]. If not, then the system is CbD-contextual (we use this term to distinguish from standard contextuality and the model-based M-contextuality introduced below). For consistently connected systems, CbD-contextuality agrees with standard contextuality.

In addition to applying beyond consistently connected systems, CbD departs from previous approaches to contextuality in that it is a purely probabilistic theory of random variables, not grounded in theories of the physical system generating the measurements. As such, it is unclear what CbD-contextuality indicates about that system. Does establishing that an inconsistently connected set of measurements is CbD-contextual imply anything about viable theories of the physical system, similar to how standard contextuality implies a system cannot be described by any local hidden-variable theory?

The present article proves an affirmative answer to this question, based on a characterization of contextuality recently advanced by Cavalcanti [18] in terms of probabilistic causal models. Probabilistic causal models are widely used in statistics, computer science, machine learning and psychology and are well suited for situations involving stochastic latent structure [19,20]. They provide a useful generalization of hidden-variable theories in physics, enabling a physical system to be described by context variables controlled by the experimenter, unobservable variables representing theoretical latent (hidden) states of the system and measured observables. Within this framework, we propose a definition of model-based contextuality (M-contextuality) that applies to both consistently and inconsistently connected systems, and we prove that it is equivalent both to CbD-contextuality and to the non-existence of a certain type of model for the system under investigation.

Our approach builds on two of the main principles that have motivated CbD: the distinction between contextuality and direct influence [6,11,14], and Cervantes and Dzhafarov's *no-conspiracy* principle prohibiting 'hidden' direct influences [14]. Early work on CbD showed that contextuality can be defined as a context-dependence of the identity of random variables over and above the dependence due to direct influences. Although the notion of direct influence was founded on the theory of selective influences in probabilistic causal models [21], subsequent developments of CbD have defined direct influence in purely probabilistic terms, to refer to the difference in an observable's distribution across different contexts [11,14]. Here we refer to such distributional differences as inconsistent connectedness, reserving direct influence to refer to a causal effect of context on the values of observables. Whereas inconsistent connectedness is an empirical (statistical) property of the measurements, direct influence as defined here is a theoretical (model-dependent) property of the physical system.

This article proposes a quantitative definition of direct influence within any probabilistic causal model, as the probability of all latent (hidden) states of the model in which a change of context changes the value of an observable. The degree of inconsistent connectedness of a measurement system imposes a minimal amount of direct influence needed to model each observable in any pair of contexts in which it is measured. We define a measurement system as M-contextual if modelling the full system requires direct influences stronger than these minimum values. This formalizes a proposal by Cavalcanti [18] that 'A causal model should not allow causal connections stronger than needed to explain the observed deviations from the no-disturbance condition' (p. 6). For consistently connected systems, the minimal direct influences are zero, and the definition of M-contextuality coincides with that of standard contextuality.

Concerning the no-conspiracy principle, Ehtibar Dzhafarov (E Dzhafarov 2018, personal communication) gives the following philosophical-level statement of the principle: 'Direct influences of a reasonable substantive theory (in physics or psychology) are not revealed in the distributional differences only under special, precariously set circumstances. As a rule, there are no "hidden" direct influences.' A conceptually similar and logically weaker principle is that of *no-fine-tuning* introduced by Wood & Spekkens [22] and elaborated by Cavalcanti [18], which holds that empirical conditional independence between measurement outcomes arises only when there is no causal connection: causal parameters cannot be fine-tuned such that their effects exactly balance out. Building on the present definition of direct influence, we formalize the idea of hidden direct influences as direct influences that work in opposite directions for different latent states, thus leaving the marginal distributions of observables unaffected. We then interpret the no-conspiracy principle as a prohibition against models with hidden influences. The primary results of this article are proofs that M-contextuality and CbD-contextuality, as properties of a measurement system, are both equivalent to the non-existence of a model of that system without hidden influences (theorems 5.3, 8.5 and 8.6).

The definition proposed here for hidden direct influence agrees with that of non-communicating signalling given in Atmanspacher and Filk's recent criticism of CbD [23]. Likewise, their observation that the criterion of CbD-contextuality accounts for communicating but not non-communicating signalling anticipates the result of the present article that a CbD-contextual system is one that cannot be modelled without hidden direct influences. We discuss in the concluding section how the formalism offered here reconciles the position of Atmanspacher and Filk with that of Dzhafarov, Kujala and colleagues, at least at a mathematical level. More generally, the value of the present results is that they show a formal equivalence between three conceptually different approaches to contextuality: (i) the assumption underlying M-contextuality that direct influence in causal models is limited to that implied by inconsistent connectedness, (ii) the no-conspiracy and no-fine-tuning principles and (iii) the probabilistic couplings approach of CbD. This correspondence will hopefully facilitate understanding and further development of both CbD- and model-based approaches to contextuality.

The remainder of this article is organized as follows. Section 2 gives notation and definitions for standard contextuality. Section 3 describes causal probabilistic models and their relationship to standard contextuality. Section 4 defines a quantitative measure of direct influence in causal models, defines hidden direct influences and offers a formalization of the no-conspiracy principle. Section 5 defines M-contextuality, proves that it agrees with standard contextuality for consistently connected systems (theorem 5.2) and proves that regardless of consistent connectedness, M-contextuality is equivalent to the non-existence of a model without hidden influences (theorem 5.3). Section 6 recasts the preceding results for systems defined by a set of separate observers, as in Bell scenarios, relating direct influence to signalling among observers. Section 7 gives examples. Section 8 derives a translation between the model-based approach and CbD and proves the final main result (theorem 8.5), that M-contextuality and CbD-contextuality are equivalent. As a corollary (theorem 8.6), we also show that CbD-contextuality can be given a causal interpretation, in that a CbD-contextual system is one that is incompatible with a particular class of probabilistic causal models, namely those without hidden influences.

2. Standard contextuality

Definition 2.1 (Measurement system). A measurement system consists of a set of observables $\mathcal{Q} = \{q\}$, a set of possible values $\mathcal{O}_q$ for each observable, a set of contexts $\mathcal{C} = \{c\}$, a relation $\prec$ with $q \prec c$ indicating that observable q is measured in context c, and a set of random variables $M = \{M_q^c : q \in \mathcal{Q}, c \in \mathcal{C}, q \prec c\}$. The subset $M^c = \{M_q^c : q \prec c\}$ is jointly distributed with distribution μ_c for each c, and M_q^c and $M_{q'}^{c'}$ are stochastically unrelated (i.e. are not measured together) whenever $c \neq c'$. Note the specification of $M = \{M_q^c\}$ determines $\mathcal{Q}$, $\{\mathcal{O}_q\}$, $\mathcal{C}$, $\prec$ and $\{\mu_c\}$, and therefore we can refer to the entire measurement system as M. The only technical requirements for the present results to hold are that $\mathcal{Q}$ and $\mathcal{C}$ are both countable (i.e. no larger than the infinite set of natural numbers) and that each $\mathcal{O}_q$ is Hausdorff and second-countable (this includes finite outcome spaces, n-dimensional Cartesian space $\mathbb{R}^n$ and separable Hilbert space).

Although most literature on contextuality treats the μ_c as known distributions, in empirical practice one has access only to samples from those distributions. Therefore, one might argue we should refer not to random variables M_q^c but to individual observations, $M_q^{c,i}$, where i $(1 \le i \le n_c)$ indexes the instances in which the experiment was performed in condition c. One advantage of the model-based approach is that it explicitly distinguishes the physical measurements M from theoretical random variables (denoted as F_q below) used to model those measurements. This distinction makes the model-based approach naturally suited to handling sampling error, by standard model-evaluation methods of null-hypothesis significance testing or Bayesian model comparison. For ease of exposition, we set aside sampling error for the majority of the article, treating the μ_c as exactly known and referring to random variables M_q^c rather than specific measurements $M_q^{c,i}$, and comment on model fitting and evaluation in the concluding section.

Definition 2.2 (Consistent connectedness). A measurement system M is consistently connected if each observable has the same marginal distribution within every context in which it is measured. That is, $M_q^c \sim M_q^{c'}$ whenever $q \prec c, c'$, where $\sim$ indicates agreement in distribution.

Definition 2.3 (Standard contextuality). A measurement system is contextual in the standard sense if it is consistently connected but the distributions μ_c are not compatible with a joint distribution over all the observables. More precisely, each μ_c is a probability measure on the Cartesian product $\prod_{q \prec c} \mathcal{O}_q$. A joint distribution μ over all the observables is a probability measure on $\prod_{q \in \mathcal{Q}} \mathcal{O}_q$, and for each c it implies a marginal distribution on the observables measured in that context, given by the push-forward measure $\pi_*^c(\mu)$ where π^c is the natural projection $\prod_{q \in \mathcal{Q}} \mathcal{O}_q \to \prod_{q \prec c} \mathcal{O}_q$. If there exists a μ such that $\pi_*^c(\mu) = \mu_c$ for all c, then the system is noncontextual; otherwise, it is contextual.

The intuitive idea behind contextuality is that the distribution of each observable is unaffected by the context (consistent connectedness), but nevertheless context exerts some sort of effect on the observables that prevents them from being pieced together into a single jointly distributed system. Importantly, the existence of a global distribution in the sense of definition 2.3 immediately implies the system is consistently connected. Therefore, consistent connectedness is a necessary property of any traditionally noncontextual system. For a system that is inconsistently connected, the standard notion of contextuality does not apply.

3. Causal-model characterization of contextuality

Following Cavalcanti [18], we analyse contextuality of a measurement system in terms of how it can be explained by probabilistic causal models [20].

Definition 3.1 (Causal probabilistic model). A causal probabilistic model is a set of jointly distributed random variables $\mathcal{X} = \{X_i\}$, with a dependency structure whereby each variable X_i has a set of parents denoted $Pa(X_i) \subset \mathcal{X}$ (possibly $Pa(X_i) = \emptyset$). The relationship between X_i and $X_{i'}$ given by $X_i \in Pa(X_{i'})$ defines a directed acyclic graph, and the model's joint distribution factors as $\Pr[\mathcal{X}] = \prod_i \Pr[X_i | Pa(X_i)]$.

For present purposes, we are interested in models of the physical system that generate some set of measurements M. For any such model, we can classify its variables into three types: variables the experimenter sets (context), variables that are measured (observables) and unobserved latent variables representing theoretical constructs of the model. This classification enables a causal model of a physical system to be put into a simple canonical form that underlies most of the analysis in this article, as follows. Context variables can be collected into a single variable C ranging over the set of contexts $\mathcal{C}$, with $Pa(C) = \emptyset$ because context is an independent variable set by the experimenter. Observable variables are denoted F_q for each observable q. Those latent variables that do not depend on anything else in the model, $\{X_i : Pa(X_i) = \emptyset\} \setminus \{C\}$, can be collected into a single random variable Λ, which we variously refer to as the source state, latent state or hidden state of the system prior to measurement. Any other latent variables, meaning intermediate ones that depend on Λ and mediating ones that depend on C or $\{F_q\}$, may play an explanatory role in interpreting the theory but are unnecessary for predictions, $\Pr[\{F_q\}|C]$. That is, we can marginalize over these other variables, so that Λ encompasses all the internal mechanisms one might theorize for the physical system, on which the observables might depend. The result is a model described fully by Λ, C and $\{F_q\}$.

This canonical structure can be further simplified in two ways. First, because the determination of context is assumed to be under control of the experimenter and only distributions conditioned on C are of interest, we can dispense with C as a random variable and treat it more simply as an index variable, meaning without any probabilities associated to its values. Second, we assume each F_q is a deterministic function of Λ and C. This assumption incurs no loss of expressive power because one can always incorporate all stochasticity in the model into the definition of Λ, by replacing it with the underlying sample space. More precisely, because all variables in a probabilistic model are jointly distributed, they can be described as functions on a probability space (Ω, Σ, P). Then Λ is a function on Ω and each F_q is a function on $\Omega \times \mathcal{C}$. By replacing Λ with the identity function on Ω, we can write each F_q as a function of Λ and C. We call models of the resulting structure canonical causal models and use them as our primary focus here.

Definition 3.2 (Canonical causal model). A canonical causal model (or simply canonical model) comprises a random variable Λ representing the hidden state of the system being modelled, an index variable C representing the contexts in which measurements can be made, and a set of functions $F_q(\Lambda, C)$ taking values in $\mathcal{O}_q$ and representing the measurement outcome for each observable. The dependency structure is thus $Pa(F_q) = \{\Lambda, C\}$, $Pa(\Lambda) = Pa(C) = \emptyset$, as shown in figure 1*a*.

It should be apparent that canonical causal models extend the class of hidden-variables models used in classic work on contextuality (e.g. [1]), by allowing context to directly influence the observables (i.e. $C \in Pa(F_q)$). Prohibiting such dependencies yields the class of (noncontextual) hidden-variables models, which we refer to here as context-free models.

Definition 3.3 (Context-free causal model). A context-free causal model (or simply context-free model) is a canonical model in which each F_q is independent of C. That is, $Pa(F_q) = \{\Lambda\}$, and $F_q(\lambda, c) = F_q(\lambda)$ for all c (figure 1*b*).

We can now formalize the relationship between canonical models and measurement systems:

Definition 3.4 (Model for a system). A canonical causal model $\mathcal{M}$ is a model for a measurement system M if it matches all of the individual contexts' data distributions. That is, for

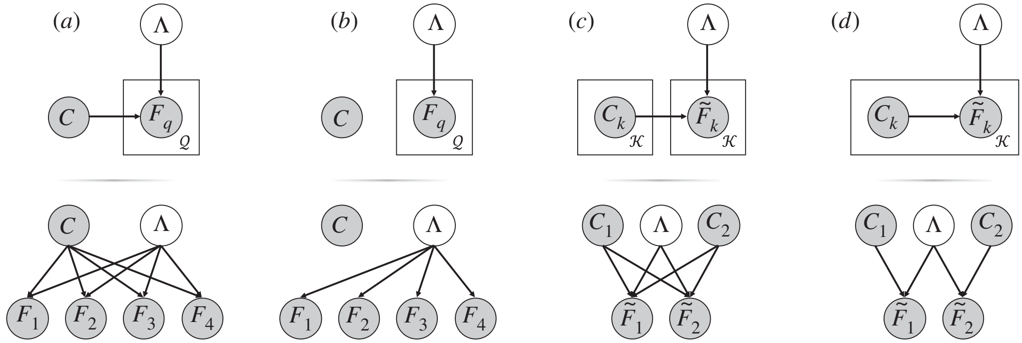

Figure 1. (*a*) Structure of a general canonical causal model. (*b*) Structure of a context-free causal model. (*c*) Structure of a general partitioned model. (*d*) Structure of a partitioned model with no signalling. Filled circles represent observed variables, unfilled circles represent unobserved variables and arrows represent dependencies. The upper diagram in each case uses plate notation: the variable in a plate is replicated over the indexing set in the lower right (e.g. a different F_q for every $q \in \mathcal{Q}$), and an arrow into or out of a plate represents an arrow to or from every copy. The lower diagram in each case explicitly depicts a model for a system with four observables or two observers, such as the examples in §7.

each context c, $\Pr[\{F_q : q \prec c\}|C = c] = \mu_c$, where the equality here is an equality of distributions, as probability measures on $\prod_{q\prec c} \mathcal{O}_q$.

The definition of a model for a system is closely related to the concept of a coupling as used in CbD [5]. A coupling for M is a set of jointly distributed random variables $T = \{T_q^c : q \in \mathcal{Q}, c \in \mathcal{C}, q \prec c\}$ such that the subset $T^c = \{T_q^c : q \prec c\}$ is distributed according to μ_c for each c (see definition 8.1 in §8). We prove in §8 (proposition 8.4) that there exists a natural translation between models and couplings for any measurement system that preserves their essential properties regarding contextuality. However, we suggest the model-based approach offers two conceptual advantages. First, the model-based approach emphasizes the ontological distinction between the measurements M_q^c as physical events, and the random variables F_q as theoretical constructs meant to explain those physical events (when taken together with the other components of $\mathcal{M}$) [18]. Second, a canonical causal model contains explicit causal structure via the latent variable Λ, which enables formal definition of direct influences of C on F_q (see §4). This in turn enables us to distinguish direct influence, as a theoretical property of the physical system and its dynamics (i.e. of the process generating the data), from inconsistent connectedness, as a purely statistical property of the data distributions.

The following three results summarize the relationship between causal models and standard contextuality. Proposition 3.5 states that the framework of canonical causal models constitutes a universal language capable of describing any measurement system (contextual or not). Proposition 3.6 recapitulates Fine's theorem [24] that noncontextuality is equivalent to compatibility with a hidden-variable theory (i.e. a context-free model). Proposition 3.7 states that consistent connectedness is equivalent to the analogous property of the individual observables. The proofs of these and all subsequent propositions and theorems are provided in the electronic supplementary material.

Proposition 3.5. *For any measurement system M, there exists a canonical causal model $\mathcal{M}$ such that $\mathcal{M}$ is a model for M.*

Proposition 3.6 (Fine [24]). *A measurement system is noncontextual iff there exists a context-free model of that system.*

Proposition 3.7. *A measurement system M is consistently connected iff there exists a context-free model for the single-observable subsystem $M_q = \{M_q^c : q \prec c\}$ for all q.*

4. Direct influence in causal models

Proposition 3.6 shows how standard contextuality can be understood in the framework of causal models. If a system of measurements can be explained by a causal model in which the value of each observable depends only on the state of the physical system prior to measurement, and not on the context (including which other observables are measured), then the system is noncontextual. If a measurement system is consistently connected but cannot be modelled without assuming the observables depend on context, then it is contextual. Furthermore, proposition 3.7 shows that consistent connectedness has a similar relationship to models of individual observables: a system is consistently connected iff each separate observable can be modelled in a way that it is not dependent on context. Therefore, a contextual system is one in which some dependence of the observables on the context is required to model the full system, even though no such dependence is needed to model any observable on its own.

These observations suggest an extension of the concept of contextuality to inconsistently connected systems. Specifically, we define a formal measure of *direct influence* of C on each F_q within any canonical causal model. We then propose to define any measurement system as contextual if modelling the complete system requires one or more direct influences to be greater than is necessary when they are considered separately. Zero direct influence will correspond to a context-free model. Therefore, our extended definition of contextuality coincides with the standard one for the case of consistently connected systems: in that case, each separate observable can be modelled with zero direct influence, and the system is contextual iff the full system can also be modelled with zero direct influence.

In a context-free model, a change between two contexts never changes the outcome, regardless of the state of the physical system: $F_q(\lambda, c) = F_q(\lambda, c')$ for all q, λ and $c, c' \succ q$. We build on this property to define a quantitative measure of direct influence in any canonical causal model.

Definition 4.1 (Direct influence). Given a canonical model $\mathcal{M} = (\Lambda, C, \{F_q : q \in \mathcal{Q}\})$, the direct influence on each F_q for any pair of contexts c and c' is defined as $\Delta_{c,c'}(F_q) = \Pr[\{\lambda : F_q(\lambda, c) \neq F_q(\lambda, c')\}]$. Thus $\Delta_{c,c'}(F_q)$ represents the probability that a change of context between c and c' would change the value of observable q, with respect to the distribution over hidden states.

It is important to note that direct influence as defined here is a characteristic of a model, not of the set of measurements being modelled. Thus, the direct influence attributed to a system is a theoretically relative construct (i.e. it depends on one's theory of the underlying physical system), as opposed to consistent connectedness, which is a purely empirical property of the measurements alone. Nevertheless, we can use direct influence to define contextuality, a property of the measurement system alone, by quantifying over models (see §5).

An important property related to direct influence is whether a model contains direct influences in opposing directions, which we refer to as hidden influences. We provide here the definition as it applies to observables with countably many possible values. The general definition, provided in the electronic supplementary material, encompasses the definition here and is conceptually similar but more technical.

Definition 4.2 (Aligned model versus hidden influences). A canonical model $\mathcal{M} = (\Lambda, C, \{F_q : q \in \mathcal{Q}\})$ is aligned if, for any observable q, value $v \in \mathcal{O}_q$, and pair of contexts $c, c' \succ q$, either $\Pr[\{\lambda : F_q(\lambda, c) = v, F_q(\lambda, c') \neq v\}] = 0$ or $\Pr[\{\lambda : F_q(\lambda, c) \neq v, F_q(\lambda, c') = v\}] = 0$. If, alternatively, both of these sets have positive probability (for some q, v, c, c'), we say the model contains hidden influences. That is, the switch from context c to c' changes the value of observable q to v for some states and away from v for other states.

Given this definition, Cervantes and Dzhafarov's no-conspiracy principle [14] can be formalized as a prohibition against models with hidden direct influences, or equivalently a restriction to aligned models.

5. Model-based contextuality

Given the formal measure defined above of direct influence of context upon an observable within a canonical causal model, we can now propose a definition of model-based contextuality that applies to inconsistently connected systems as well as to consistently connected ones. If a system is inconsistently connected, one can consider how great each direct influence must be, to model the difference in distribution for any observable between any pair of contexts. One can then ask whether these minimal direct influences are mutually compatible: That is, does modelling the full system require direct influences to be greater than is necessary individually? This approach formalizes the generalized no-fine-tuning principle proposed by Cavalcanti [18] that direct influences should be no greater than required by violations of consistent connectedness.

Definition 5.1 (M-contextuality). A measurement system M is M-noncontextual if there exists a canonical model $\mathcal{M}$ for M that simultaneously minimizes all direct influences. That is, for each q and $c, c' \succ q$, $\mathcal{M}$ achieves the minimum value of $\Delta_{c,c'}(F_q)$ over all models for M. If such a model does not exist, M is M-contextual.

For consistently connected systems, this definition reduces to that of standard contextuality. Indeed, the minimal direct influences for a consistently connected system are all zero, and consequently M-contextuality becomes equivalent to the non-existence of a context-free model, which is equivalent to standard contextuality by proposition 3.6.

Theorem 5.2. *For consistently connected systems, M-contextuality is equivalent to standard contextuality.*

Our main result regarding M-contextuality is a characterization in terms of the existence of aligned models. The proof is based on explicit determination of the lower bound on direct influence for each observable and context pair and on showing that the models meeting that bound are precisely those without hidden direct influences.

Theorem 5.3. *A measurement system M is M-contextual iff there does not exist an aligned canonical model for M.*

According to theorem 5.3, an M-contextual system is one for which all models must contain hidden direct influences. Thus, M-contextuality embodies the no-conspiracy principle: if one excludes hidden influences *a priori*, then the criterion of M-noncontextuality is simply that a system can be modelled by a classical probabilistic model. Put differently, if there is no aligned model for a system, then, following the logic of the no-conspiracy principle [14], we conclude the system cannot be explained by direct influences alone and that therefore there are contextual influences in addition to the direct ones.

6. Partitionable systems

In many measurement systems of interest, the observables can be partitioned such that exactly one observable from each subset is measured in any context. This situation arises in quantum physics when there are multiple observers and each observer can measure one out of a set of pairwise incompatible observables, as in Bell scenarios.

Definition 6.1 (Partitionable measurement system). A partitionable measurement system is one in which the set of observables can be partitioned as $\mathcal{Q} = \bigsqcup_{k\in\mathcal{K}} \mathcal{Q}_k$, the set of contexts is $\mathcal{C} = \prod_{k\in\mathcal{K}} \mathcal{Q}_k$, and $q \prec c$ iff $c_k = q$ for each $q \in \mathcal{Q}_k$. That is, every context c corresponds to a choice of exactly one observable (denoted c_k) from each subset $\mathcal{Q}_k$. We say k indexes observers, and $\mathcal{Q}_k$ is the set of (pairwise incompatible) observables available to observer k.

When a measurement system is partitionable, it admits an alternative form of causal model, with one outcome variable per observer that depends on which measurement that observer chooses as well as (potentially) on the choices of all other observers. We refer to this as a partitioned model.

Definition 6.2 (Partitioned model). Given a partitionable measurement system M and a canonical model $\mathcal{M} = (\Lambda, C, \{F_q\})$ for M, the corresponding partitioned model $\tilde{\mathcal{M}} = (\Lambda, \{C_k\}, \{\tilde{F}_k\})$ is defined as follows. The hidden state of the system is modelled by the same random variable, Λ, as in $\mathcal{M}$. The context is decomposed into a set of variables $\{C_k : k \in \mathcal{K}\}$, with C_k ranging over $\mathcal{Q}_k$ and indicating which observable is measured by observer k. The measurement outcome for observer k is represented by a variable $\tilde{F}_k$ defined by $\tilde{F}_k(\lambda, c) = F_{c_k}(\lambda, c)$.

The general dependency structure of $\tilde{\mathcal{M}}$ is $Pa(\Lambda) = Pa(C_k) = \emptyset$ and $Pa(\tilde{F}_k) = \{\Lambda\} \cup \{C_{k'} : k' \in \mathcal{K}\}$, as illustrated in figure 1*c*. Note that the range of $\tilde{F}_k$ is the union of outcome spaces $\bigcup_{q \in \mathcal{Q}_k} \mathcal{O}_q$, although the image of $\tilde{F}_k(\cdot, c)$ is contained in $\mathcal{O}_{c_k}$ for any c. Therefore, $\Pr[\{\tilde{F}_k\} | C = c]$ can be taken as a distribution on $\prod_k \mathcal{O}_{c_k}$, which under definition 6.2 matches the distribution $\Pr[\{F_q : q \prec c\} | C = c]$ from $\mathcal{M}$. Therefore, $\tilde{\mathcal{M}}$ is a model for M whenever $\mathcal{M}$ is.

A partitioned model $\tilde{\mathcal{M}}$ expresses the same causal theory of a system as the corresponding canonical model $\mathcal{M}$, and therefore notions of direct influence and aligned models can be straightforwardly translated to partitioned models. The direct influence of interest here is the influence on one observer's outcome due to the other observers' choices of observables, which we refer to as signalling.

Definition 6.3 (Signalling). Given a partitioned model $\tilde{\mathcal{M}}$, an observer k, and contexts c and c' with $c_k = c'_k$, signalling to observer k is defined as $\tilde{\Delta}_{c,c'}(\tilde{F}_k) = \Pr[\{\lambda : \tilde{F}_k(\lambda, c) \neq \tilde{F}_k(\lambda, c')\}]$. That is, $\tilde{\Delta}_{c,c'}(\tilde{F}_k)$ is the probability of a latent state in which switching the measurement choices of other observers (i.e. changing $C_{k'}$ for one or more $k' \neq k$) changes the outcome for observer k.

It is easy to see that signalling in a partitioned model $\tilde{\mathcal{M}}$ is equal to direct influence in the corresponding canonical model $\mathcal{M}$. That is, $\tilde{\Delta}_{c,c'}(\tilde{F}_k) = \Delta_{c,c'}(F_q)$ whenever $c_k = c'_k = q$. We can also define hidden signals versus aligned partitioned models, paralleling the definition of hidden direct influences and aligned canonical models. As with definition 4.2, the definition given here applies to observables with countably many possible values; the general definition is provided in the electronic supplementary material.

Definition 6.4 (Aligned versus hidden signals). A partitioned model $\tilde{\mathcal{M}}$ is aligned if, for any observer k, value $v \in \bigcup_{q \in \mathcal{Q}_k} \mathcal{O}_q$, and pair of contexts c and c' with $c_k = c'_k$, either $\Pr[\{\lambda : \tilde{F}_k(\lambda, c) = v, \tilde{F}_k(\lambda, c') \neq v\}] = 0$ or $\Pr[\{\lambda : \tilde{F}_k(\lambda, c) \neq v, \tilde{F}_k(\lambda, c') = v\}] = 0$. If, alternatively, both sets have positive probability (for some k, v, c, c'), we say $\tilde{\mathcal{M}}$ contains hidden signals. That is, a specific change of choices of observables for other observers can change observer k's measurement outcome to v for some states and away from v for other states.

The distinction between aligned and hidden signals is the same as that between communicating and non-communicating signals [23], because aligned signals affect the marginal distribution of the receiving observer's (k's) measurements, whereas purely hidden signals do not. It is easy to see that a canonical model $\mathcal{M}$ is aligned iff the corresponding partitioned model $\tilde{\mathcal{M}}$ is aligned. Likewise, $\mathcal{M}$ is context-free iff $\tilde{\mathcal{M}}$ has no signalling. The latter condition means that $\tilde{F}_k(\lambda, c) = \tilde{F}_k(\lambda, c')$ whenever $c_k = c'_k$, so that each observer's outcome can be written as $\tilde{F}_k(\lambda, c_k)$, and $\tilde{\mathcal{M}}$ conforms to the simplified dependency structure $Pa(\tilde{F}_k) = \{\Lambda, C_k\}$ for every k (figure 1*d*). These observations imply a partitioned analogue of proposition 3.6, whereby standard contextuality for consistently connected partitionable systems is equivalent to the non-existence of a partitioned model without signalling, as in the original analyses of Bell scenarios [2]:

Proposition 6.5. *A consistently connected partitionable measurement system M is noncontextual iff there exists a partitioned model for M that has no signalling.*

The definition of M-contextuality can also be expressed in terms of partitioned models:

Theorem 6.6. *A partitionable measurement system M is M-noncontextual iff there exists a partitioned model $\tilde{\mathcal{M}}$ that simultaneously minimizes all signalling. That is, for each k, c and c' with $c_k = c'_k$, $\tilde{\mathcal{M}}$ achieves the minimum value of $\tilde{\Delta}_{c,c'}(\tilde{F}_k)$ over all partitioned models for M.*

Finally, the analogue of theorem 5.3 for partitionable systems states that M-contextuality is equivalent to the non-existence of a model without hidden signalling:

Theorem 6.7. *A partitionable measurement system M is M-contextual iff there does not exist an aligned partitioned model for M.*

7. Examples

(a) Popescu–Rohrlich box

The Popescu–Rohrlich (PR) box [25] is a measurement system with four binary observables and four contexts, $\{M_1^1, M_2^1, M_2^2, M_3^2, M_3^3, M_4^3, M_4^4, M_1^4\}$, with marginals $\Pr[M_q^c = 1] = 1/2$ for all $q \prec c$ (table 1), and correlations $\Pr[M_1^1 = M_2^1] = 1$, $\Pr[M_2^2 = M_3^2] = 1$, $\Pr[M_3^3 = M_4^3] = 1$, $\Pr[M_1^4 = M_4^4] = -1$ (table 2). This system is consistently connected and contextual (hence also M-contextual). Consequently, the measurements for each observable, M_q, can be modelled without direct influence, but the system as a whole cannot.

Define a model $\mathcal{M}$ for the PR box by taking Λ to range over $\{-1, 1\}$ with uniform distribution, and defining the F_q by

$$F_1\,(\lambda, 1) = \lambda \quad F_2\,(\lambda, 2) = \lambda \quad F_3\,(\lambda, 3) = \lambda \quad F_4\,(\lambda, 4) = -\lambda$$

and

$$F_2\,(\lambda, 1) = \lambda \quad F_3\,(\lambda, 2) = \lambda \quad F_4\,(\lambda, 3) = \lambda \quad F_1\,(\lambda, 4) = \lambda.$$

Then $\Delta_{c,c'}(F_q) = 0$ for $q = 1, 2, 3$, but $\Delta_{c,c'}(F_4) = 1$ (where c and c' are the two applicable contexts for each q). Furthermore, the direct influence on F_4 is hidden: for $\Lambda = 1$, a switch from context 3 to context 4 changes F_4 from 1 to -1, whereas for $\Lambda = -1$ the same switch of context changes F_4 from -1 to 1. These opposing influences cancel out in the margin. This situation illustrates the general fact that causal models for contextual systems always exist, but they must include hidden direct influences.

The PR box is a partitionable system, and can be written as $\mathcal{K} = \{1, 2\}$ and $\mathcal{Q}_1 = \{1, 3\}$, $\mathcal{Q}_2 = \{2, 4\}$. That is, observer 1 measures observable 1 or 3, and observer 2 measures observable 2 or 4. The contexts are now denoted $\mathcal{C} = \{(1, 2), (3, 2), (3, 4), (1, 4)\}$. The model $\mathcal{M}$ corresponds to a partitioned model $\tilde{\mathcal{M}}$ with outcome variables defined by

$$\tilde{F}_1\,(\lambda, (1, 2)) = \lambda \quad \tilde{F}_1\,(\lambda, (3, 2)) = \lambda \quad \tilde{F}_1\,(\lambda, (3, 4)) = \lambda \quad \tilde{F}_1\,(\lambda, (1, 4)) = \lambda$$

and

$$\tilde{F}_2\,(\lambda, (1, 2)) = \lambda \quad \tilde{F}_2\,(\lambda, (3, 2)) = \lambda \quad \tilde{F}_2\,(\lambda, (3, 4)) = \lambda \quad \tilde{F}_2\,(\lambda, (1, 4)) = -\lambda.$$

Signalling in this model is zero everywhere except for $\tilde{\Delta}_{(1,4),(3,4)}(\tilde{F}_2) = 1$. That is, observer 1's measurement outcome is unaffected by observer 2's choice of observable, regardless of the state of the physical system, and a similar statement holds for observer 2 when he or she measures observable 2. However, when observer 2 measures observable 4, the outcome depends on which variable observer 1 measures. Moreover, this signalling goes in opposite directions depending on the latent state Λ. That is, $\tilde{\mathcal{M}}$ contains signalling, and that signalling is (perfectly) hidden. In the language of Atmanspacher & Filk [23], this is non-communicating signalling that cannot transmit information between observers.

(b) An M-noncontextual inconsistently connected system

Consider a system with the same observables, contexts and correlation structure as the PR box, but with unbalanced marginals as shown in table 3. The inconsistent connectedness of this system places a lower bound on the direct influences in any model thereof, which can be shown to be $\Delta_{1,4}(F_1) \geq 1/6$, $\Delta_{1,2}(F_2) \geq 1/3$, $\Delta_{2,3}(F_3) \geq 1/3$ and $\Delta_{3,4}(F_4) \geq 1/6$. The question is whether these lower bounds can be achieved simultaneously, and the answer is affirmative. Let $\mathcal{M}$ be the model

Table 1. Marginal distributions for the PR box. Each observable takes values in $\{-1, 1\}$.

$\Pr[M_q^c = 1]$	q			
c	1	2	3	4
1	$\frac{1}{2}$	$\frac{1}{2}$		
2		$\frac{1}{2}$	$\frac{1}{2}$	
3			$\frac{1}{2}$	$\frac{1}{2}$
4	$\frac{1}{2}$			$\frac{1}{2}$

Table 2. Correlation structure of the PR box. For each pair of observables (q, q') that can be measured together, the table shows their probability of equality within the corresponding context (c).

$\Pr[M_q^c = M_{q'}^c]$	q'	
q	2	4
1	1	-1
3	1	1

Table 3. Marginal distributions for an example measurement system that is inconsistently connected and M-noncontextual. The joint distribution in each context is determined by table 2.

$\Pr[M_q^c = 1]$	q			
c	1	2	3	4
1	$\frac{2}{3}$	$\frac{2}{3}$		
2		$\frac{1}{3}$	$\frac{1}{3}$	
3			$\frac{2}{3}$	$\frac{2}{3}$
4	$\frac{1}{2}$			$\frac{1}{2}$

for this system defined by Λ ranging over $\{1, 2, 3, 4, 5, 6\}$ with uniform distribution, and the F_q as given in table 4. This model is immediately seen to achieve the lower bound on all $\Delta_{c,c'}(F_q)$. Note also that the direct influences are all aligned. For example, switching from context 1 to context 2 can change F_2 from 1 to -1 (when $\Lambda \in \{3, 4\}$), but it cannot change F_2 from -1 to 1. These properties of $\mathcal{M}$ show the system is M-noncontextual.

The partitioned model $\tilde{\mathcal{M}}$ corresponding to $\mathcal{M}$ (not shown) has similar properties. Signalling is present for both observers under both measurement settings, and in all cases the signalling is the minimum possible. Furthermore, this signalling is aligned: any change in one observer's choice of observable changes the other observer's measurement outcome only in one direction (or not at all) across all latent states of the system. Consequently, the signalling shows up fully in a change in the latter observer's marginal probability. In the language of Atmanspacher & Filk [23], this is purely communicating signalling that transmits information between observers.

8. Relationship to contextuality-by-default

The CbD approach to contextuality is originally motivated by the observation that the traditional approach of treating M_q^c as one and the same random variable for all contexts c is mathematically

Table 4. A canonical model for the measurement system in table 3. Each entry shows the value of F_q given the values of Λ and C. Λ is uniformly distributed over its six values.

	Λ					
	1	2	3	4	5	6
$F_1(\lambda, 1)$	1	1	1	1	−1	−1
$F_2(\lambda, 1)$	1	1	1	1	−1	−1
$F_2(\lambda, 2)$	1	1	−1	−1	−1	−1
$F_3(\lambda, 2)$	1	1	−1	−1	−1	−1
$F_3(\lambda, 3)$	1	1	−1	−1	1	1
$F_4(\lambda, 3)$	1	1	−1	−1	1	1
$F_4(\lambda, 4)$	1	−1	−1	−1	1	1
$F_1(\lambda, 4)$	−1	1	1	1	−1	−1

incoherent, even for consistently connected systems [11]. CbD offers a rigorous alternative to this traditional approach through the construct of a probabilistic coupling [17].

Definition 8.1 (Probabilistic coupling). A set of jointly distributed random variables $T = \{T_q^c : q \in \mathcal{Q}, c \in \mathcal{C}, q \prec c\}$ is a coupling for the measurement system $M = \{M_q^c : q \in \mathcal{Q}, c \in \mathcal{C}, q \prec c\}$ if $T^c \sim M^c$ for all contexts c, meaning $\Pr[T^c] = \mu_c$ as distributions on $\prod_{q \prec c} \mathcal{O}_q$.

This definition enables the definition of contextuality to be shifted from the mathematically unsound question of a joint distribution for $\{M_q^{\cdot}\}$ (where the superscript '$\cdot$' indicates the context is simply ignored) to well-defined questions regarding the joint distribution of $\{T_q^c\}$. For a consistently connected system $M = \{M_q^c\}$, the relevant coupling (if it exists) satisfies $\Pr[T_q^c = T_q^{c'}] = 1$ for all $c, c' \succ q$. This formalizes the idea of treating M_q^c and $M_q^{c'}$ as the same random variable. The theory also enables the concept of contextuality to be extended to inconsistently connected systems, by considering couplings in which T_q^c and $T_q^{c'}$ are not always equal. We follow the CbD 2.0 version of the theory, which is based on multimaximal couplings [16].

Definition 8.2 (Multimaximal coupling). For a set of stochastically unrelated random variables $M_q = \{M_q^c : c \succ q\}$, a coupling $T_q = \{T_q^c : c \succ q\}$ is multimaximal if, for all c and c', $\Pr[T_q^c = T_q^{c'}]$ is maximal among all couplings of M_q.

Definition 8.3 (CbD-contextuality). A measurement system M is CbD-noncontextual if there exists a coupling T for M such that, for each observable q, T_q is a multimaximal coupling for M_q. Otherwise, M is CbD-contextual.

It is easy to see that CbD-contextuality agrees with standard contextuality for consistently connected systems. If M is consistently connected, a multimaximal coupling for M is one that satisfies $\Pr[T_q^c = T_q^{c'}] = 1$ for all $c, c' \succ q$. Such a coupling is equivalent to a joint distribution over all of the observables (i.e. $\Pr[\{T_q^{\cdot} : q \in \mathcal{Q}\}]$) that has the distributions μ_c as marginals.

The main results of this section rest on a correspondence between couplings and canonical causal models:

Proposition 8.4. *Given a coupling T for a measurement system M, there exists a canonical model $\mathcal{M}$ for M such that $\Delta_{c,c'}(F_q) = \Pr[T_q^c \neq T_q^{c'}]$ for all q and $c, c' \succ q$. Likewise, given a canonical model $\mathcal{M}$ for a measurement system M, there exists a coupling T for M such that the same relationship holds.*

This correspondence implies the equivalence between M-contextuality and CbD-contextuality:

Theorem 8.5. *A measurement system M is M-contextual iff it is CbD-contextual.*

Finally, the equivalence between M-contextuality and CbD-contextuality, together with theorem 5.3, provides an interpretation of CbD-contextuality in terms of aligned models:

Theorem 8.6. *A measurement system M is CbD-contextual iff there does not exist an aligned model for M.*

Thus CbD-contextuality has a simple interpretation in terms of direct influence, specifically that a measurement system is CbD-contextual whenever modelling that system would require hidden direct influences or, equivalently, hidden signalling for partitionable systems. In other words, a CbD-contextual system is one that can be modelled only by violating the no-conspiracy principle.

9. Conclusion

The present results show an equivalence between three approaches to extending contextuality analysis to inconsistently connected systems: Cavalcanti's generalized no-fine-tuning principle [18] for causal probabilistic models, formalized here in terms of minimal direct influence (M-contextuality); Cervantes and Dzhafarov's no-conspiracy principle [14], formalized here as a restriction to aligned causal models; and the purely mathematical framework of CbD [16]. Thus, the requirement of multimaximal couplings in CbD is equivalent to the assumption that direct influences (in the causal sense used here) are no greater than needed to explain the marginal distributions of individual observables, as well as to the assumption that direct influences never cancel out, even partially. Although there is a sense in which any mathematical definition is as good as any other, the fact that three different principles all converge on the same classification of systems into contextual and noncontextual might be taken as a stronger indication of the likely scientific utility of this classification.

Another main contribution of this article, beyond the correspondence between the particular definitions analysed, is the translation between the CbD and model-based approaches. Under this translation, alternative conditions on couplings correspond to alternative criteria for direct influence. For example, the maximal couplings used in earlier versions of CbD [5] correspond to canonical causal models minimizing $\Pr[\{\lambda : \exists c, c' \succ q\ (F_q(\lambda, c) \neq F_q(\lambda, c'))\}]$ for each observable q. Likewise, other measures of causal influence in probabilistic models [26] might translate to interesting criteria on couplings. This strategy of translation might enable insights from either framework to inform development in the other, to better understand the properties of contextual inconsistently connected systems, or to devise more refined definitions.

Although the translation between canonical models and couplings is fairly trivial, with the latent state of the model corresponding to the sample space of the coupling (proposition 8.4), the model-based approach offers a number of conceptual advantages. First, it maintains the standard interpretation of contextuality as the impossibility of explaining a system classically, meaning with a hidden-variables theory in which measurements that were not made are nevertheless well defined [27]. Second, it distinguishes direct influence, a property of the theoretical data-generating process, from inconsistent connectedness, a property of the data distribution [18]. Consequently, third, it provides a formalism for expressing how an observable might directly depend on context, thus enabling potential extensions wherein theoretical assumptions regarding the physical system impose additional constraints on this dependence.

Expanding on the third point, the definition of M-contextuality is based on causal models that may embody rich theories of the physical system underlying a set of measurements. Because the definition quantifies over all such models, it depends only on the measurements themselves (i.e. $\{\mu_c\}$ or samples therefrom). It is thus a meta-theory, providing conditions regarding what types of models are and are not mathematically possible for a given system. However, the model-based approach can also accommodate theoretical considerations directly. That is, one could incorporate domain-specific assumptions regarding direct influences, based on theoretical principles applicable to the physical system under study, and define a system as contextual whenever any model of that system would require stronger direct influences than allowed by the assumed theory. This is the standard logic with Bell scenarios, where the underlying theory is special relativity and the constraint is that signalling between spacelike-separated observers must be absent. The present approach can also accommodate more nuanced constraints, including limits on the magnitude of direct influence or signalling even when it is allowed to be non-zero. A further generalization of the analysis presented here would be to consider model architectures other than the canonical and partitioned ones. For example, rather than allowing direct influence of context on each observable, $C \in Pa(F_q)$, or signalling between observers, $C_k \in Pa(\tilde{F}_{k'})$, one could consider models with direct influence between observables, $F_q \in Pa(F_{q'})$, to see what additional insight they might provide into the mathematical nature of contextuality.

As a fourth advantage, the model-based approach is naturally suited to the fact that contextuality analysis of an empirical measurement system is analysis of a finite dataset. As with any situation of hypothesis testing based on sample data, it requires statistical inference, in this case inference of whether the data could have been generated by a model from some class. As noted in §2, the data form a set $\hat{M} = \{M_q^{c,i} : q \in \mathcal{Q}, c \in \mathcal{C}, q \prec c, 1 \leq i \leq n_c\}$, with each $M_\cdot^{c,i}$ a sample from an unknown distribution μ_c. The set of possibilities for $\{\mu_c : c \in \mathcal{C}\}$ is a product of $(m_c - 1)$-simplices, where $m_c = |\{q : q \prec c\}|$ is the number of observables measured in context c. The subset of possibilities consistent with a context-free model (or equivalently a partitioned model with no signalling) is defined by a set of linear inequalities and hence is a polytope within that space. Likewise, the subset of possibilities consistent with an aligned model (canonical or partitioned) is defined by a different polytope, containing the first. Given a dataset $\hat{M}$, the questions of contextuality, M-contextuality and CbD-contextuality are questions of whether $\hat{M}$ is a sample from a set of true distributions lying in the appropriate polytope, which can be answered with existing inferential methods [28–30]. Thus, the model-based approach can accommodate sampling error in a principled way, without conflating it with direct influence.

The present results help to clarify and qualify Atmanspacher and Filk's argument that CbD-contextuality is an inadequate definition because it accounts only for communicating and not non-communicating (hidden) signalling, whereas 'signalling whatsoever, "hidden" or not, cannot create true quantum contextuality' [23]. Expositions of CbD make clear that this distinction is not meaningful within that framework, since it defines direct influence solely in distributional terms [11,14]. Thus, the two groups are essentially speaking different languages. The present approach bridges this gap, by distinguishing the probabilistic concept of inconsistent connectedness from the causal concept of direct influence, and by showing how one can translate between probabilistic couplings and causal models (proposition 8.4). Using this translation, theorem 8.6 shows Atmanspacher and Filk's statement is correct: CbD-contextuality implies that a system cannot be explained entirely by communicating (aligned) direct influence or signalling, but it does not imply the system cannot be explained by hidden influences or signalling. On the other hand, the latter observation is not specific to CbD but applies to all possible theories of contextuality: if unconstrained, direct influence can always explain any pattern of data (proposition 3.5), as Bohmian mechanics demonstrates for EPR-Bell scenarios [31]. Therefore, contextuality is of interest only under some restriction on direct influence. The no-conspiracy principle is one such restriction. In that regard, theorem 8.6 unifies the positions of Atmanspacher and Filk and of Dzhafarov, Kujala and colleagues: CbD-contextuality is equivalent to the proposition that a system cannot be explained by direct influence alone, if one excludes hidden influences *a priori*. The conceptual justification and scientific utility of this exclusion will likely be a matter of further debate.

Data accessibility. This article has no additional data.
Competing interests. I declare I have no competing interests.
Funding. I received no funding for this study.

References

1. Bell JS. 1964 On the Einstein Podolsky Rosen paradox. *Physics* **1**, 195–200. (doi:10.1103/PhysicsPhysiqueFizika.1.195)
2. Clauser J, Horne M, Shimony A, Holt R. 1969 Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23**, 880–884. (doi:10.1103/PhysRevLett.23.880)
3. Aspect A, Grangier G, Roger G. 1981 Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.* **47**, 460–463. (doi:10.1103/PhysRevLett.47.460)
4. Hensen B *et al.* 2015 Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682–686. (doi:10.1038/nature15759)
5. Kujala JV, Dzhafarov EN, Larsson J-A. 2015 Necessary and sufficient conditions for extended noncontextuality in a broad class of quantum mechanical systems. *Phys. Rev. Lett.* **115**, 150401. (doi:10.1103/PhysRevLett.115.150401)
6. Dzhafarov EN, Kujala JV. 2014 A qualified Kolmogorovian account of probabilistic contextuality. Lecture notes in computer science, vol. 8369. In *Quantum Interaction: 7th International Conference, QI 2013* (eds H Atmanspacher, E Haven, K Kitto, D Raine), pp. 201–212. Berlin, Germany: Springer.
7. Townsend JT, Schweickert R. 1989 Toward the trichotomy method of reaction times: laying the foundation of stochastic mental networks. *J. Math. Psychol.* **33**, 309–327. (doi:10.1016/0022-2496(89)90012-6)
8. Ramanathan R, Soeda A, Kurzynski P, Kaszlikowski D. 2012 Generalized monogamy of contextual inequalities from the no-disturbance principle. *Phys. Rev. Lett.* **109**, 050404. (doi:10.1103/PhysRevLett.109.050404)
9. Dzhafarov EN. 2018 Replacing nothing with something special: contextuality-by-default and dummy measurements. In *Quantum foundations, probability and information* (eds A Khrennikov, T Bourama), pp. 39–44. Berlin, Germany: Springer.
10. Dzhafarov EN, Kujala JV. 2015 Conversations on contextuality. In *Contextuality from quantum physics to psychology* (eds EN Dzhafarov *et al.*), pp. 1–22. Hackensack, NJ: World Scientific.
11. Dzhafarov EN, Kujala JV. 2016 Context-content systems of random variables: the contextuality-by-default theory. *J. Math. Psychol.* **74**, 11–33. (doi:10.1016/j.jmp.2016.04.010)
12. Aerts D, Arguëlles JA, Beltran L, Geriente S, Sassoli de Bianchi M, Sozzo S, Veloz T. 2018 Spin and wind directions I: identifying entanglement in nature and cognition. *Found. Sci.* **23**, 323–335. (doi:10.1007/s10699-017-9528-9)
13. Aerts D, Gabora L, Sozzo S. 2013 Concepts and their dynamics: a quantum-theoretic modeling of human thought. *Top. Cogn. Sci.* **5**, 737–772. (doi:10.1111/tops.12042)
14. Cervantes VH, Dzhafarov EN. 2018 Snow queen is evil and beautiful: experimental evidence for probabilistic contextuality in human choices. *Decision* **5**, 193–204. (doi:10.1037/dec0000095)
15. Dzhafarov EN, Kujala JV, Cervantes VH, Zhang R, Jones M. 2016 On contextuality in behavioral data. *Phil. Trans. R. Soc. A* **374**, 20150234. (doi:10.1098/rsta.2015.0234)
16. Dzhafarov EN, Kujala JV. 2017 Contextuality-by-Default 2.0: systems with binary random variables. Lecture notes in computer science, vol. 10 106. In *Quantum Interaction: 10th International Conference, QI 2016* (eds JA de Barros, B Coecke, E Pothos), pp. 16–32. Berlin, Germany: Springer.
17. Thorisson H. 2000 *Coupling, stationarity, and regeneration*. New York, NY: Springer.
18. Cavalcanti EG. 2018 Classical causal models for Bell and Kochen–Specker inequality violations require fine-tuning. *Phys. Rev. X* **8**, 021018. (doi:10.1103/physrevx.8.021018)
19. Jordan MI. 1999 *Learning in graphical models*. Cambridge, MA: MIT Press.
20. Pearl J. 1999 *Causality: models, reasoning and inference*. Cambridge, UK: Cambridge University Press.
21. Dzhafarov EN. 2003 Selective influence through conditional independence. *Psychometrika* **68**, 7–26. (doi:10.1007/BF02296650)
22. Wood CJ, Spekkens RW. 2015 The lesson of causal discovery algorithms for quantum correlations: causal explanations of Bell-inequality violations require fine-tuning. *New J. Phys.* **17**, 033002. (doi:10.1088/1367-2630/17/3/033002)
23. Atmanspacher H, Filk T. In press. Contextuality revisited–signaling may differ from communicating. In *Quanta and mind: essays on the connection between quantum mechanics and the consciousness* (eds A de Barros, C Montemayor). Berlin, Germany: Springer.

24. Fine A. 1982 Hidden variables, joint probability, and the Bell inequalities. *Phys. Rev. Lett.* **48**, 291. (doi:10.1103/PhysRevLett.48.291)
25. Popescu S, Rohrlich D. 1994 Quantum nonlocality as an axiom. *Found. Phys.* **24**, 379–385. (doi:10.1007/BF02058098)
26. Janzing D, Balduzzi D, Grosse-Wentrup M, Schölkopf B. 2013 Quantifying causal influences. *Ann. Stat.* **41**, 2324–2358. (doi:10.1214/13-AOS1145)
27. Einstein A, Podolsky B, Rosen N. 1935 Can quantum-mechanical description of physical reality be considered complete?. *Phys. Rev.* **47**, 777–780. (doi:10.1103/PhysRev.47.777)
28. Davis-Stober CP. 2009 Multinomial models under linear inequality constraints: applications to measurement theory. *J. Math. Psychol.* **53**, 1–13. (doi:10.1016/j.jmp.2008.08.003)
29. Heck DW, Davis-Stober CP. 2019 Multinomial models with linear inequality constraints: overview and improvements of computational methods for Bayesian inference. *J. Math. Psychol.* **91**, 70–87. (doi:10.1016/j.jmp.2019.03.004)
30. Myung JI, Karabatsos G, Iverson GJ. 2005 A Bayesian approach to testing decision making axioms. *J. Math. Psychol.* **49**, 205–225. (doi:10.1016/j.jmp.2005.02.004)
31. Bohm D. 1952 A suggested interpretation of the quantum theory in terms of 'hidden variables' I & II. *Phys. Rev.* **85**, 166–193. (doi:10.1103/PhysRev.85.166)

Relating Causal and Probabilistic Approaches to Contextuality

Supplementary Material

Matt Jones (mcj@colorado.edu)

University of Colorado Boulder

Proof of Proposition 1 For each c, extend μ_c to a probability measure $\mu^{(c)}$ on $\prod_{q \in \mathcal{Q}} \mathcal{O}_q$, for example by choosing an arbitrary distribution for each observable $q \not\prec c$ and then taking the product measure. Identify the set of possible values for Λ (i.e., the space of hidden states) with the Cartesian product $\prod_{c \in \mathcal{C}} \left(\prod_{q \in \mathcal{Q}} \mathcal{O}_q \right)$, and let μ_Λ be the product measure on this space obtained from the measures $\left\{ \mu^{(c)} : c \in \mathcal{C} \right\}$. Finally, define $F_q(\lambda, c) = \lambda_{c,q}$. For any c, the conditional distribution $\Pr[\{F_q : q \prec c\} | C = c]$ is equal to the distribution obtained from μ_Λ by projecting $\prod_{c' \in \mathcal{C}} \left(\prod_{q \in \mathcal{Q}} \mathcal{O}_q \right) \to \prod_{q \prec c} \mathcal{O}_q$ (taking copy c from the outside product and marginalizing over all $q \not\prec c$ in the inside product), which by construction is μ_c.

Proof of Proposition 2 Sufficiency: Noncontextuality of M implies there exists a distribution μ over all the observables such that its projection to the observables within each context c equals the distribution μ_c. Define a context-free model with Λ ranging over $\prod_{q \in \mathcal{Q}} \mathcal{O}_q$ with probability measure $\mu_\Lambda = \mu$, and with $F_q(\lambda) = \lambda_q$ for every q and λ. Then the joint distribution $\Pr[\{F_q : q \in \mathcal{Q}\}]$ equals μ, and thus for any context c, the distribution $\Pr[\{F_q : q \prec c\}]$ equals μ_c.

Necessity: If $\mathcal{M}$ is a context-free model for M, then μ_c is the same distribution as $\Pr[\{F_q : q \prec c\} | C = c]$, which in turn is the same as $\Pr[\{F_q : q \prec c\}]$ because the F_q do not depend on C. Therefore μ_c equals the projection of $\Pr[\{F_q : q \in \mathcal{Q}\}]$ to $\prod_{q \prec c} \mathcal{O}_q$ for all c, implying M is noncontextual.

Proof of Proposition 3 In any context-free model for M_q, F_q is independent of C, implying $\Pr[F_q | C = c] = \Pr[F_q]$ (as distributions on $\mathcal{O}_q$) for all c. Therefore M_q^c has the same distribution for all c, implying consistent connectedness. Conversely, if M is consistently connected then for each q we can define the probability measure μ_q on $\mathcal{O}_q$ that is the distribution shared by all M_q^c. A model of M_q is then trivially constructed by letting Λ range over $\mathcal{O}_q$ with distribution μ_q and taking $F_q(\lambda) = \lambda$ for all $\lambda \in \mathcal{O}_q$.

General Definition of Aligned Canonical Models and Hidden Influences Given an observable q with arbitrary outcome space $\mathcal{O}_q$ and two contexts $c, c' \succ q$, a canonical causal model $\mathcal{M}$ is said to have hidden direct influences with respect to $\{q, c, c'\}$ when there exists a measurable set $E \subset \mathcal{O}_q$ such that $\Pr[\{\lambda : F_q(\lambda, c) \in E\}] > 0$, $\Pr\left[\left\{\lambda : F_q\left(\lambda, c'\right) \in E\right\}\right] > 0$, and for every measurable subset $E' \subset E$, either $\Pr\left[\left\{\lambda : F_q(\lambda, c) \in E', F_q\left(\lambda, c'\right) \notin E'\right\}\right] > 0$ and $\Pr\left[\left\{\lambda : F_q(\lambda, c) \notin E', F_q\left(\lambda, c'\right) \in E'\right\}\right] > 0$, or else $\Pr\left[\left\{\lambda : F_q(\lambda, c) \in E'\right\}\right] = \Pr\left[\left\{\lambda : F_q\left(\lambda, c'\right) \in E'\right]\right\} = 0$. A model is aligned if it has no hidden direct influences for any q, c, c'. This definition is equivalent to Definition 9 in the main text when $\mathcal{O}_q$ is discrete, as can be seen by identifying E with $\{v\}$.

Proof of Theorem 1 Let M be a consistently connected measurement system. By Proposition 3, for each q there exists a context-free model $\mathcal{M}_q$ for M_q. The model $\mathcal{M}_q$ satisfies $\Delta_{c,c'}(F_q) = 0$ for all $c, c' \succ q$, and it can be arbitrarily extended to a model for the full system. Therefore, M is M-noncontextual iff there exists a model for M with all direct influences equal to zero.

If there exists a context-free model for M, all direct influences in this model are zero and therefore M is M-noncontextual. Conversely, assume M is M-noncontextual a let $\mathcal{M}$ be a model for M with all direct influences equal to zero. For each q and contexts $c, c' \succ q$, define $E_q^{cc'} = \{\lambda : F_q(\lambda, c) = F_q(\lambda, c')\}$. By assumption, $\Pr\left[E_q^{cc'}\right] = 1$. Because $\mathcal{Q}$ and $\mathcal{C}$ are assumed to be countable, $\Pr[E] = 1$, where $E = \bigcap_{q,c,c':q \prec c,c'} E_q^{cc'}$. Now define a new model $\mathcal{M}'$ by restricting the range of Λ and the domain of every F_q (in the first argument) to E. By construction, $\mathcal{M}'$ is a context-free model for M.

Proof of Theorem 2 Fix and q and $c, c \succ q$, and let μ^c and $\mu^{c'}$ respectively be the distributions of M_q^c and $M_q^{c'}$, as probability measures on $\mathcal{O}_q$. By the Hahn-Jordan decomposition theorem applied to the signed measure $\mu^c - \mu^{c'}$, there exist a partition of the outcome space $\mathcal{O}_q = \mathcal{O}_q^+ \sqcup \mathcal{O}_q^-$ and positive measures μ^+ and μ^- such that $\mu^+(\mathcal{O}_q^-) = \mu^-(\mathcal{O}_q^+) = 0$ and $\mu^c - \mu^{c'} = \mu^+ - \mu^-$. Moreover, μ^+ and μ^- are unique. Define $\mu^0 = \mu^c - \mu^+ = \mu^{c'} - \mu^-$, which is necessarily a positive measure, and define $\alpha = \mu^0(\mathcal{O}_q)$. We prove the following three statements:

(i) The minimal direct influence across all models for M is given by $\min_{\mathcal{M}} \Delta_{c,c'}(F_q) = 1 - \alpha$.
(ii) If a model $\mathcal{M}$ for M satisfies $\Delta_{c,c'}(F_q) = 1 - \alpha$, then it contains no hidden influences with respect to $\{q, c, c'\}$.
(iii) Conversely, if a model $\mathcal{M}$ for M contains no hidden influences with respect to $\{q, c, c'\}$, then it satisfies $\Delta_{c,c'}(F_q) = 1 - \alpha$.

Together, these three statements imply that any model $\mathcal{M}$ for M is aligned iff it minimizes all direct influences, which in turn implies the theorem.

Proof of Statement (i). Let $\mathcal{M}$ be any canonical model for M. The direct influence in $\mathcal{M}$ is constrained by

$$\begin{aligned}
\Delta_{c,c'}(F_q) &\geq \Pr\left[\left\{\lambda : F_q(\lambda, c) \in \mathcal{O}_q^+, F_q(\lambda, c') \notin \mathcal{O}_q^+\right\}\right] \\
&\geq \Pr\left[\left\{\lambda : F_q(\lambda, c) \in \mathcal{O}_q^+\right\}\right] - \Pr\left[\left\{\lambda : F_q(\lambda, c') \in \mathcal{O}_q^+\right\}\right] \\
&= \mu^c\left(\mathcal{O}_q^+\right) - \mu^{c'}\left(\mathcal{O}_q^+\right) \\
&= \mu^+\left(\mathcal{O}_q^+\right) - \mu^-\left(\mathcal{O}_q^+\right) \\
&= \mu^+(\mathcal{O}_q) \\
&= \mu^c(\mathcal{O}_q) - \mu^0(\mathcal{O}_q) \\
&= 1 - \alpha.
\end{aligned}$$

Therefore $1 - \alpha$ is a lower bound for $\Delta_{c,c'}(F_q)$. To construct a model meeting this bound, let Λ range over $\mathcal{O}_q \times \mathcal{O}_q$ and define $F_q((v_1, v_2), c) = v_1$ and $F_q((v_1, v_2), c') = v_2$ for all $v_1, v_2 \in \mathcal{O}_q$. Let $\pi^d : \mathcal{O}_q \to \mathcal{O}_q \times \mathcal{O}_q$ be the diagonal embedding $\pi^d(v) = (v, v)$, and define the push-forward measure $\mu^d = \pi_*^d\left(\mu^0\right)$, so that $\mu^d(E) = \mu^0(\{v \in \mathcal{O}_q : (v, v) \in E\})$ for all measurable $E \subset \mathcal{O}_q \times \mathcal{O}_q$. Define a second measure μ^u on $\mathcal{O}_q \times \mathcal{O}_q$, generated by

$$\mu^u(E_1 \times E_2) = \frac{\mu^+(E_1) \cdot \mu^-(E_2)}{1 - \alpha}$$

for all measurable $E_1, E_2 \subset \mathcal{O}_q$. Now define the distribution on Λ by $\Pr[\Lambda] = \mu^d + \mu^u$. For any measurable $E \subset \mathcal{O}_q$,

$$\begin{aligned}
\Pr[F_q \in E | C = c] &= \mu^d (E \times \mathcal{O}_q) + \mu^u (E \times \mathcal{O}_q) \\
&= \mu^0 (E) + \frac{\mu^+ (E) \cdot \mu^- (\mathcal{O}_q)}{1-\alpha} \\
&= \mu^0 (E) + \frac{\mu^+ (E) \cdot \left(\mu^{c'} (\mathcal{O}_q) - \mu^0 (\mathcal{O}_q)\right)}{1-\alpha} \\
&= \mu^c (E) .
\end{aligned}$$

A similar calculation shows $\Pr\left[F_q \in E \middle| C = c'\right] = \mu^{c'} (E)$. Therefore $\mathcal{M}$ is a model for the subsystem $\left\{M_q^c, M_q^{c'}\right\}$, which can be arbitrarily extended to a model for the full system M. The direct influence is given by

$$\begin{aligned}
\Delta_{c,c'} (F_q) &= \mu^d (\mathcal{O}_q \times \mathcal{O}_q) \\
&= \frac{\left(\mu^c (\mathcal{O}_q) - \mu^0 (\mathcal{O}_q)\right) \cdot \left(\mu^{c'} (\mathcal{O}_q) - \mu^0 (\mathcal{O}_q)\right)}{1-\alpha} \\
&= 1 - \alpha.
\end{aligned}$$

Proof of Statement (ii). Assume $\mathcal{M}$ has hidden influences with respect to $\{q, c, c'\}$, and let E be as given above in the General Definition of Hidden Influences. Define $E^+ = E \cap \mathcal{O}_q^+$ and $E^- = E \cap \mathcal{O}_q^-$. Because $\Pr[\{\lambda : F_q(\lambda, c) \in E\}] > 0$ and $\Pr\left[\left\{\lambda : F_q\left(\lambda, c'\right) \in E\right\}\right] > 0$, it cannot be that $\Pr\left[\left\{\lambda : F_q(\lambda, c) \in E^+\right\}\right] = \Pr\left[\left\{\lambda : F_q\left(\lambda, c'\right) \in E^+\right\}\right] = 0$ and $\Pr\left[\left\{\lambda : F_q(\lambda, c) \in E^-\right\}\right] = \Pr\left[\left\{\lambda : F_q\left(\lambda, c'\right) \in E^-\right\}\right] = 0$. Without loss of generality, assume the former equality, $\Pr\left[\left\{\lambda : F_q(\lambda, c) \in E^+\right\}\right] = \Pr\left[\left\{\lambda : F_q\left(\lambda, c'\right) \in E^+\right\}\right] = 0$, is false. Then the definition of hidden influences implies $\Pr\left[\left\{\lambda : F_q(\lambda, c) \notin E^+, F_q\left(\lambda, c'\right) \in E^+\right\}\right] > 0$. Because $E^+ \subset \mathcal{O}_q^+$, the sets $\left\{\lambda : F_q(\lambda, c) \in \mathcal{O}_q^+, F_q\left(\lambda, c'\right) \notin \mathcal{O}_q^+\right\}$ and $\left\{\lambda : F_q(\lambda, c) \notin E^+, F_q\left(\lambda, c'\right) \in E^+\right\}$ are disjoint, so we can bound the direct influence as $\Delta_{c,c'}(F_q) \geq \Pr\left[\left\{\lambda : F_q(\lambda, c) \in \mathcal{O}_q^+, F_q\left(\lambda, c'\right) \notin \mathcal{O}_q^+\right\}\right] + \Pr\left[\left\{\lambda : F_q(\lambda, c) \notin E^+, F_q\left(\lambda, c'\right) \in E^+\right\}\right]$. The proof of Statement 1 shows the former of these probabilities is at least $1 - \alpha$, and therefore we have $\Delta_{c,c'}(F_q) > 1 - \alpha$. Thus we have shown any model with hidden influences cannot satisfy $\Delta_{c,c'}(F_q) = 1 - \alpha$.

Proof of Statement (iii). We first prove that $\Pr\left[\left\{\lambda : F_q(\lambda, c) \notin E, F_q\left(\lambda, c'\right) \in E\right\}\right] = 0$ for any measurable $E \subset \mathcal{O}_q^+$. To see this, assume the contrary, that $\Pr\left[\left\{\lambda : F_q(\lambda, c) \notin E, F_q\left(\lambda, c'\right) \in E\right\}\right] = \varepsilon$ with $\varepsilon > 0$ for some $E \subset \mathcal{O}_q^+$. Using alignment of $\mathcal{M}$, the probability ε can be squeezed into successively smaller subsets of E so as to produce a contradiction. Specifically, define a property S with $S\left(E'\right)$ being the statement that E' is a measurable subset of E with $\Pr\left[\left\{\lambda : F_q(\lambda, c) \notin E, F_q\left(\lambda, c'\right) \in E \setminus E'\right\}\right] = 0$. Note that S is preserved under countable intersection and that $S\left(E'\right)$ implies $\Pr\left[\left\{\lambda : F_q(\lambda, c) \notin E, F_q\left(\lambda, c'\right) \in E'\right\}\right] = \varepsilon$. If we define $\beta = \inf\left\{\Pr\left[\left\{\lambda : F_q(\lambda, c) \in E'\right\}\right] : S\left(E'\right)\right\}$, then the countable intersection property just stated implies there exists a set $E_0 \subset E$ meeting this bound: $\Pr[\{\lambda : F_q(\lambda, c) \in E_0\}] = \beta$ and $\Pr\left[\left\{\lambda : F_q(\lambda, c) \notin E, F_q\left(\lambda, c'\right) \in E_0\right\}\right] = \varepsilon$. If $\beta > 0$, then alignment of $\mathcal{M}$ implies there are no hidden influences within E_0; that is, there exists $E_1 \subset E_0$ such that $\Pr[\{\lambda : F_q(\lambda, c) \in E_1\}] > 0$ or $\Pr\left[\left\{\lambda : F_q\left(\lambda, c'\right) \in E_1\right\}\right] > 0$, and also $\Pr\left[\left\{\lambda : F_q(\lambda, c) \notin E_1, F_q\left(\lambda, c'\right) \in E_1\right\}\right] = 0$ or $\Pr\left[\left\{\lambda : F_q(\lambda, c) \in E_1, F_q\left(\lambda, c'\right) \notin E_1\right\}\right] = 0$. Because $E_1 \subset \mathcal{O}_q^+$, $\Pr[\{\lambda : F_q(\lambda, c) \in E_1\}] \geq \Pr[\{\lambda : F_q\left(\lambda, c'\right) \in E_1\}]$, which implies that the former relation in each of the two disjunctions just given holds: $\Pr[\{\lambda : F_q(\lambda, c) \in E_1\}] > 0$ and $\Pr\left[\left\{\lambda : F_q(\lambda, c) \notin E_1, F_q\left(\lambda, c'\right) \in E_1\right\}\right] = 0$. This in turn implies $S(E_0 \setminus E_1)$ and $\Pr[\{\lambda : F_q(\lambda, c) \in E_0 \setminus E_1\}] < \beta$, contradicting the definition of β. On the other hand, if $\beta = 0$ then $\Pr[\{\lambda : F_q(\lambda, c) \in E_0\}] < \Pr\left[\left\{\lambda : F_q\left(\lambda, c'\right) \in E_0\right\}\right]$, contradicting the fact that $E_0 \subset \mathcal{O}_q^+$. Therefore the supposed set $E \subset \mathcal{O}_q^+$ with $\Pr\left[\left\{\lambda : F_q(\lambda, c) \notin E, F_q\left(\lambda, c'\right) \in E\right\}\right] > 0$ cannot exist.

Next, let $(E_n)_{n\in\mathbb{N}}$ be a countable basis for $\mathcal{O}_q^+$, using the assumption that $\mathcal{O}_q$ is second-countable. Take any $v_1 \in \mathcal{O}_q$ and $v_2 \in \mathcal{O}_q^+$ with $v_1 \neq v_2$. Because $\mathcal{O}_q$ is Hausdorff, there exists be an open neighborhood N of v_2 not containing v_1. Because $(E_n)_{n\in\mathbb{N}}$ is a basis for $\mathcal{O}_q^+$, there exists some E_m with $v_2 \in E_m \subset N \cap \mathcal{O}_q^+$ and hence also $v_1 \notin E_m$. This shows that $\left\{\lambda : F_q\left(\lambda, c'\right) \in \mathcal{O}_q^+, F_q\left(\lambda, c\right) \neq F_q\left(\lambda, c'\right)\right\}$ is a subset of $\bigcup_n \left\{\lambda : F_q\left(\lambda, c\right) \notin E_n, F_q\left(\lambda, c'\right) \in E_n\right\}$. Therefore

$$\begin{aligned}\Pr\left[\left\{\lambda : F_q\left(\lambda, c'\right) \in \mathcal{O}_q^+, F_q\left(\lambda, c\right) \neq F_q\left(\lambda, c'\right)\right\}\right] &\leq \sum_n \Pr\left[\left\{\lambda : F_q\left(\lambda, c\right) \notin E_n, F_q\left(\lambda, c'\right) \in E_n\right\}\right] \\ &= 0,\end{aligned}$$

which in turn implies

$$\begin{aligned}\Pr\left[\left\{\lambda : F_q\left(\lambda, c\right) = F_q\left(\lambda, c'\right) \in \mathcal{O}_q^+\right\}\right] &= \Pr\left[\left\{\lambda : F_q\left(\lambda, c'\right) \in \mathcal{O}_q^+\right\}\right] \\ &= \mu^0\left(\mathcal{O}_q^+\right).\end{aligned}$$

A parallel argument shows $\Pr\left[\left\{\lambda : F_q\left(\lambda, c\right) = F_q\left(\lambda, c'\right) \in \mathcal{O}_q^-\right\}\right] = \mu^0\left(\mathcal{O}_q^-\right)$. Therefore the total direct influence in $\mathcal{M}$ for q, c, c' is given by

$$\begin{aligned}\Delta_{c,c'}\left(F_q\right) &= 1 - \Pr\left[\left\{\lambda : F_q\left(\lambda, c\right) = F_q\left(\lambda, c'\right)\right\}\right] \\ &= 1 - \Pr\left[\left\{\lambda : F_q\left(\lambda, c\right) = F_q\left(\lambda, c'\right) \in \mathcal{O}_q^+\right\}\right] - \Pr\left[\left\{\lambda : F_q\left(\lambda, c\right) = F_q\left(\lambda, c'\right) \in \mathcal{O}_q^-\right\}\right] \\ &= 1 - \mu^0\left(\mathcal{O}_q^+\right) - \mu^0\left(\mathcal{O}_q^-\right) \\ &= 1 - \alpha.\end{aligned}$$

General Definition of Aligned Partitioned Models and Hidden Signals Given an observer k, an observable $q \in \mathcal{Q}_k$ with arbitrary outcome space $\mathcal{O}_q$, and contexts c and c' with $c_k = c'_k = q$, a partitioned model $\tilde{\mathcal{M}}$ is said to have hidden signals with respect to $\left\{k, c, c'\right\}$ when there exists a measurable set $E \subset \mathcal{O}_q$ such that $\Pr\left[\left\{\lambda : \tilde{F}_k\left(\lambda, c\right) \in E\right\}\right] > 0$, $\Pr\left[\left\{\lambda : \tilde{F}_k\left(\lambda, c'\right) \in E\right\}\right] > 0$, and for every measurable subset $E' \subset E$, either $\Pr\left[\left\{\lambda : \tilde{F}_k\left(\lambda, c\right) \in E', \tilde{F}_k\left(\lambda, c'\right) \notin E'\right\}\right] > 0$ and $\Pr\left[\left\{\lambda : \tilde{F}_k\left(\lambda, c\right) \notin E', \tilde{F}_k\left(\lambda, c'\right) \in E'\right\}\right] > 0$, or else $\Pr\left[\left\{\lambda : \tilde{F}_k\left(\lambda, c\right) \in E'\right\}\right] = \Pr\left[\left\{\lambda : \tilde{F}_k\left(\lambda, c'\right) \in E'\right\}\right] = 0$. A partitioned model is aligned if it has no hidden signals for any k, c, c'. This definition is equivalent to Definition 14 in the main text when $\mathcal{O}_q$ is discrete for all $q \in \mathcal{Q}_k$, as can be seen by identifying E with $\{v\}$.

Proof of Proposition 4 If M is noncontextual, then Proposition 2 implies there exists a context-free canonical model $\mathcal{M}$ for M. The corresponding partitioned model $\tilde{\mathcal{M}}$ is easily seen to be a model for M with no signaling. Conversely, if there is a partitioned model $\tilde{\mathcal{M}}$ for M that has no signaling, the corresponding canonical model $\mathcal{M}$ is context-free, and Proposition 2 then implies M is noncontextual.

Proof of Theorem 3 Let $\tilde{\mathcal{M}}$ be any partitioned model for M, with $\mathcal{M}$ the corresponding canonical model. As observed in the main text, direct influence in $\mathcal{M}$ and signaling in $\tilde{\mathcal{M}}$ exactly correspond, in that $\tilde{\Delta}_{c,c'}\left(\tilde{F}_k\right) = \Delta_{c,c'}\left(F_q\right)$ whenever $c_k = c'_k = q$. Therefore $\tilde{\mathcal{M}}$ minimizes all signaling iff $\mathcal{M}$ minimizes all direct influences. The theorem then follows from the definition of M-noncontextuality, as the existence of such an $\mathcal{M}$.

Proof of Theorem 4 If $\tilde{\mathcal{M}}$ is an aligned partitioned model for M, then the corresponding canonical model $\mathcal{M}$ is also aligned, implying M is M-noncontextual by Theorem 2. Conversely, if

M is M-noncontextual, there exists an aligned canonical model $\mathcal{M}$ for M by Theorem 2, and the corresponding partitioned model $\tilde{\mathcal{M}}$ is also aligned.

Proof of Proposition 5 First part: Let (Ω, Σ, P) be the sample space for the jointly distributed random variables composing T, such that each T_q^c is a function $\Omega \to \mathcal{O}_q$. Define $\mathcal{M}$ by letting Λ range over Ω with distribution P and defining each F_q by $F_q(\lambda, c) = T_q^c(\lambda)$ for $c \succ q$ and choosing arbitrary values for $F_q(\lambda, c)$ for $c \not\succ q$ (for all $\lambda \in \Omega$). Then for any context c and measurable subsets $V_q \subset \mathcal{O}_q$,

$$\begin{aligned}\Pr[\forall q \prec c\, (F_q \in V_q) | C = c] &= \Pr[\{\lambda : \forall q \prec c\, (F_q(\lambda, c) \in V_q)\}] \\ &= \Pr\left[\left\{\lambda : \forall q \prec c\, \left(T_q^c(\lambda) \in V_q\right)\right\}\right] \\ &= \Pr\left[\forall q \prec c\, \left(T_q^c \in V_q\right)\right] \\ &= \Pr\left[\forall q \prec c\, \left(M_q^c \in V_q\right)\right].\end{aligned}$$

Therefore $\mathcal{M}$ is a model for M. For any q and $c, c' \succ q$, the claimed equality holds:

$$\begin{aligned}\Delta_{c,c'}(F_q) &= \Pr\left[\left\{\lambda : F_q(\lambda, c) \neq F_q\left(\lambda, c'\right)\right\}\right] \\ &= \Pr\left[\left\{\lambda : T_q^c(\lambda) \neq T_q^{c'}(\lambda)\right\}\right] \\ &= \Pr\left[T_q^c \neq T_q^{c'}\right].\end{aligned}$$

Second part: Given $\mathcal{M} = (\Lambda, C, \{F_q\})$, let $\Omega = \{\lambda\}$ be the range of Λ with $P = \Pr[\Lambda]$ the associated probability measure on Ω and Σ the sigma-algebra of measurable sets of values for Λ. Then (Ω, Σ, P) defines a sample space. For each q and $c \succ q$, define a random variable T_q^c on this sample space by $T_q^c(\lambda) = F_q(\lambda, c)$. Then derivations similar to those above show that $T = \left\{T_q^c\right\}$ is a coupling for M and that $\Delta_{c,c'}(F_q) = \Pr\left[T_q^c \neq T_q^{c'}\right]$ for all q and $c, c' \succ q$.

Proof of Theorem 5 If M is M-noncontextual, then there exists a canonical causal model $\mathcal{M}$ for M that simultaneously minimizes all direct influences. The corresponding coupling T provided by Proposition 5 minimizes $\Pr\left[T_q^c \neq T_q^{c'}\right]$ for all q and $c, c' \succ q$. Therefore T_q is multimaximal for all q, implying M is CbD-noncontextual. Conversely, if M is CbD-noncontextual then there exists a coupling T for M such that T_q is multimaximal for all q, implying $\Pr\left[T_q^c \neq T_q^{c'}\right]$ is minimal for all $c, c' \succ q$. The corresponding canonical model $\mathcal{M}$ provided by Proposition 5 minimizes $\Delta_{c,c'}(F_q)$ for all q and $c, c' \succ q$, implying M is M-noncontextual.

Proof of Theorem 6 The theorem follows directly from Theorems 2 and 5: CbD-contextuality is equivalent to M-contextuality, which is equivalent to the non-existence of an aligned model.