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 $F_q\left(\lambda,c\right)=\lambda_{c,q}$. For any c, the conditional distribution $\Pr\left[\{F_q:q\prec c\}|C=c\right]$ is equal to the distribution obtained from μ_Λ by projecting $\prod_{c'\in\mathcal{C}}\left(\prod_{q\in\mathcal{Q}}\mathcal{O}_q\right)\rightarrow\prod_{q\prec c}\mathcal{O}_q$ (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 $F_q\left(\lambda\right)=\lambda_q$ for every q and λ. Then the joint distribution $\Pr\left[\{F_q:q\in\mathcal{Q}\}\right]$ equals μ, and thus for any context c, the distribution $\Pr\left[\{F_q:q\prec c\}\right]$ equals μ_c.

Necessity: If $\mathcal{M}$ is a context-free model for M, then μ_c is the same distribution as $\Pr\left[\{F_q:q\prec c\}|C=c\right]$, which in turn is the same as $\Pr\left[\{F_q:q\prec c\}\right]$ because the F_q do not depend on C. Therefore μ_c equals the projection of $\Pr\left[\{F_q:q\in\mathcal{Q}\}\right]$ 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 $\Pr\left[F_q|C=c\right]=\Pr\left[F_q\right]$ (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 $F_q\left(\lambda\right)=\lambda$ 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 $\Pr\left[\{\lambda:F_q\left(\lambda,c\right)\in E\}\right]>0$, $\Pr\left[\left\{\lambda:F_q\left(\lambda,c'\right)\in E\right\}\right]>0$, and for every measurable subset $E'\subset E$, either $\Pr\left[\left\{\lambda:F_q\left(\lambda,c\right)\in E',F_q\left(\lambda,c'\right)\notin E'\right\}\right]>0$ and $\Pr\left[\left\{\lambda:F_q\left(\lambda,c\right)\notin E',F_q\left(\lambda,c'\right)\in E'\right\}\right]>0$, or else $\Pr\left[\left\{\lambda:F_q\left(\lambda,c\right)\in E'\right\}\right]=\Pr\left[\left\{\lambda:F_q\left(\lambda,c'\right)\in E'\right]\right\}=0$. 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 $\Delta_{c,c'}\left(F_q\right)=0$ 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 $\mu^+\left(\mathcal{O}_q^-\right) = \mu^-\left(\mathcal{O}_q^+\right) = 0$ 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 $E \subset \mathcal{O}_q$,

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

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

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

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 $\{k, c, c'\}$ 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 $V_q \subset \mathcal{O}_q$,

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

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

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

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.