Title: Causal-model approach to extended contextuality

Speakers: Matt Jones

Collection: Causal Inference & Quantum Foundations Workshop

Date: April 20, 2023 - 2:30 PM

URL: https://pirsa.org/23040131

Abstract: There has been recent interest in extending the concept of contextuality to cases of disturbance or inconsistent connectedness. This talk will describe an approach using probabilistic causal models, which generalize the hidden-variables models of Bell and Kochen & Specker, following recent work by Cavalcanti. I first prove an equivalence between three conditions on an arbitrary measurement system: (1) existence of a model minimizing all causal influences of context upon measurement outcomes, (2) prohibition of a form of "hidden" causal influence, and (3) noncontextuality as defined in the Contextuality-by-Default (CbD) theory of Dzhafarov and Kujala. The no-hidden-influence principle thus confers a physical interpretation to CbD-contextuality, paralleling Bell's local causality and Kochen & Specker's classical embeddability. I then extend this analysis to other causal graph topologies, showing that different graphs yield different notions of contextuality, but only the one corresponding to CbD agrees with traditional contextuality when restricted to non-disturbing systems.

Causal-model Approach to Contextuality for Disturbing Systems

Matt Jones

University of Colorado

Causal Inference & Quantum Foundations Workshop
Perimeter Institute
20 Apr 2023

Contextuality for Disturbing Systems

- Bell nonlocality, Kochen-Specker contextuality
  - Data distributions incompatible with causal model embodying observed conditional independences
- Applies only without signaling or disturbance
  - Bell (1964): spacelike separation
  - Kochen & Specker (1967): operator algebras
- Interest in disturbing systems: Extended contextuality (Kujala et al. 2015)
  - Essential to certain domains of interest
  - Difficult to avoid in experimental practice (noise, finite samples)
- Contextuality-by-default (CbD) (Dzhafarov & Kujala 2017)
  - Theory of random variables
  - Physical implications? No-go interpretation?

Model-based Extended Contextuality

- Contextuality ⇒ fine-tuning (Wood & Spekkens 2015; Pearl & Cavalcanti 2021)
  - Causal effects cancel exactly in the marginals
- Extended contextuality ⇒ hidden influence (Jones 2019)
  - Causal effects partially cancel in marginals
- Strengthening of casual Markov/SEM condition

"Direct influences…are not revealed in the distributional differences only under special, precariously set circumstances." (Dzhafarov 2019)

"A causal model should not allow causal connections stronger than needed to explain the observed deviations from the no-disturbance condition." (Cavalcanti 2018)

Outline

- Quantify causal influence and hidden influence in DAGs
- M-contextuality (model-based contextuality)
  - Incompatible with causal model minimizing causal influences
- No-go theorem for extended contextuality

  M-noncontextual

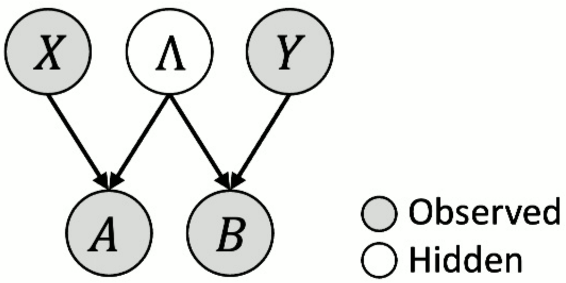
  $\Leftrightarrow$ $\exists$ model with no hidden influences

  $\Leftrightarrow$ CbD-noncontextual

  $\Leftrightarrow$ KS-noncontextual (for nondisturbing systems)
- Distinct variants for different graphical structures

Probabilistic Causal Models

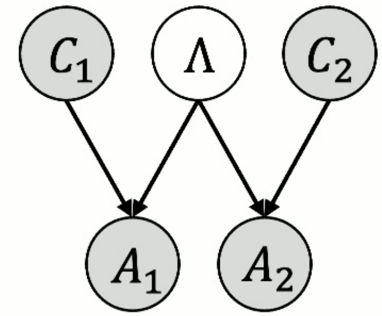
- Powerful tool from statistics, machine learning, cognitive science (Jordan 1999; Pearl 2000; Tenenbaum et al. 2011)
- Extend hidden-variables models (Bell 1964; Kochen & Specker 1967)
- Directed acyclic graph: $(\mathcal{V}, \rightarrow)$
- Joint distribution factors into causal dependencies $\Pr[X_{\mathcal{V}}] = \prod_{v\in\mathcal{V}} \Pr[X_v | Pa(X_v)]$



$$\Pr[\lambda, x, y, a, b] = \Pr[\lambda] \Pr[x] \Pr[y] \Pr[a|x,\lambda] \Pr[b|y,\lambda]$$
$$\Pr[a, b|x, y] = \int \Pr[a|x,\lambda] \Pr[b|y,\lambda] \, \mathrm{d}\lambda$$

Causal Models: Bell Scenarios

G^{Loc}: No Signaling

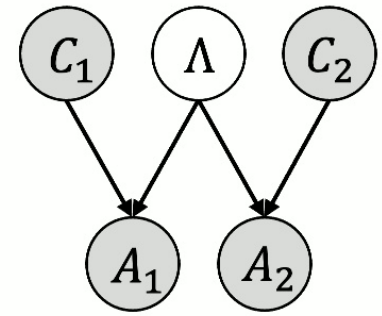


- Notation
  - Observers: $\mathcal{K} = \{k\}$
  - Observables for observer k: $\mathcal{F}_k = \{F_q : q \in \mathrm{Q}_k\}$
  - Setting for observer k: C_k
  - Measurements $A_k = F_{C_k}$
- Local causality (Bell 1976)
  - Distribution $P\big((A_k)_{k\in\mathcal{K}}, (C_k)_{k\in\mathcal{K}}\big)$
  - Markov w.r.t. G^{Loc}
  - No signaling: $A_i \perp_P C_j$ $(j \neq i)$
  - CHSH inequality (Clauser et al 1969)

Causal Models: Bell Scenarios

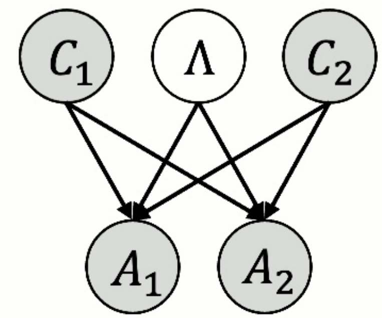
- Notation
  - Observers: $\mathcal{K} = \{k\}$
  - Observables for observer k: $\mathcal{F}_k = \{F_q : q \in \mathrm{Q}_k\}$
  - Setting for observer k: C_k
  - Measurements $A_k = F_{C_k}$
- Local causality (Bell 1976)
  - Distribution $P\big((A_k)_{k\in\mathcal{K}}, (C_k)_{k\in\mathcal{K}}\big)$
  - Markov w.r.t. G^{Loc}
  - No signaling: $A_i \perp_P C_j \ (j \neq i)$
  - CHSH inequality (Clauser et al 1969)

G^{Loc}: No Signaling

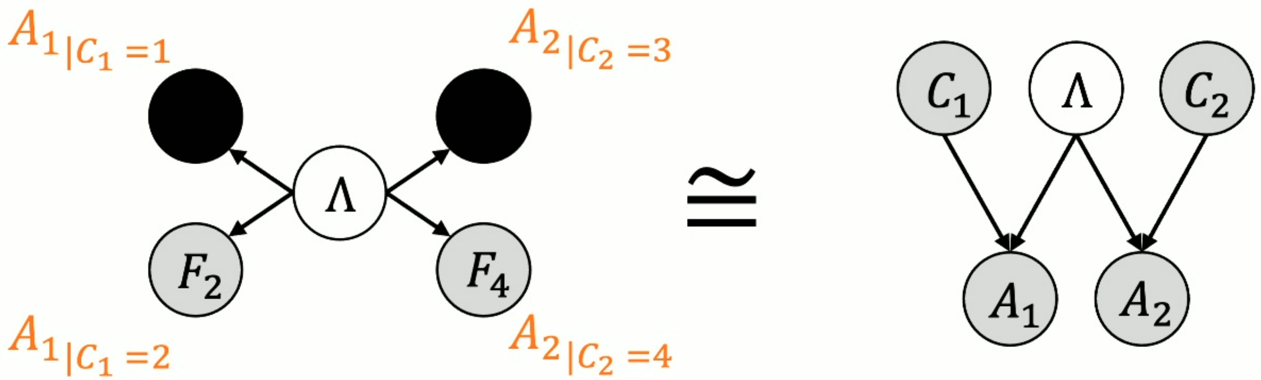


parameter dependence

G^{pd}: Signaling



Kochen-Specker Representation



$$F_q \equiv A_{k|C_k=q}$$

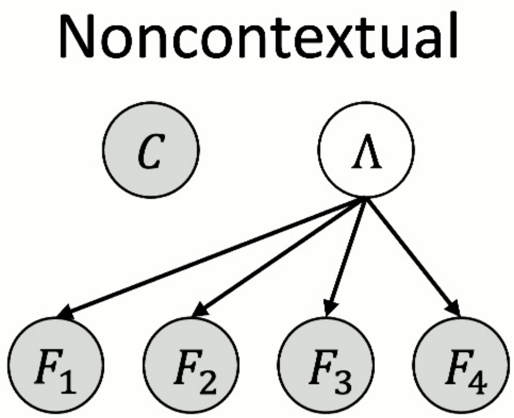
$$P^{\mathrm{KS}}(F_q|\Lambda) = P^{\mathrm{Bell}}(A_k|\Lambda, C_k = q)$$

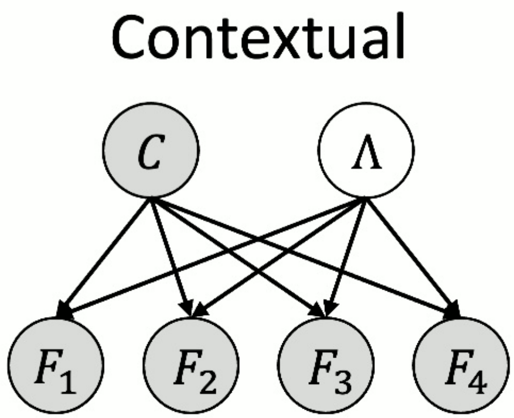
General Kochen-Specker system

- **Measurement system** $M = \{\mathcal{Q}, \mathcal{C}, \prec, (\mu_c)_{c\in\mathcal{C}}\}$
  - Observables $\mathcal{Q} = \{q\}$ countable
  - Contexts $\mathcal{C} = \{c\}$ countable
  - Outcome space $\mathcal{O}_q$ for each q second-countable (e.g. $\mathbb{R}^n$)
  - Measurement F_q for each q
  - Distributions μ_c for $\boldsymbol{F}^c = \{F_q : q \prec c\}$, on $\mathcal{O}^c = \prod_{q\prec c} \mathcal{O}_q$
- **Standard contextuality (KS-contextuality)**
  - No disturbance: $\mu_{c|\mathbf{q}} = \mu_{c'|\mathbf{q}}$ for all $\mathbf{q} \subset \mathcal{Q}$ and $c, c' \succ \mathbf{q}$
  - No global distribution on $\prod_{q\in\mathcal{Q}} \mathcal{O}_q$ compatible with all μ_c

KS (non)contextual Models

- Context, latent state, and observables
  - $\mathcal{M} = \left(C, \Lambda, \{F_q : q \in \mathcal{Q}\}\right)$
  - $Pa(\Lambda) = Pa(C) = \emptyset$
  - $Pa(F_q) = \{\Lambda\}$ or $\{C, \Lambda\}$
  - C treated as index variable
- Deterministic observables (Fine 1982)
  - $F_q(c, \lambda) \in \mathcal{O}_q$
  - No loss of expressive power: push stochasticity into Λ
- Model $\mathcal{M}$ for a system M
  - $\Pr[\boldsymbol{F}^c | C = c] = \mu_c$ for all c



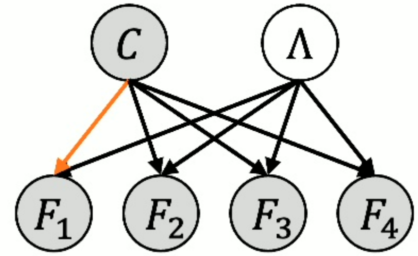


Canonical Causal Models and KS-Contextuality

- **Proposition 1**. For any measurement system M, there exists a (possibly contextual) model $\mathcal{M}$ that is a model for M.

- **Proposition 2 (Fine 1982)**. A nondisturbing measurement system admits a global distribution iff it admits a noncontextual model.

Direct Influence in Causal Models



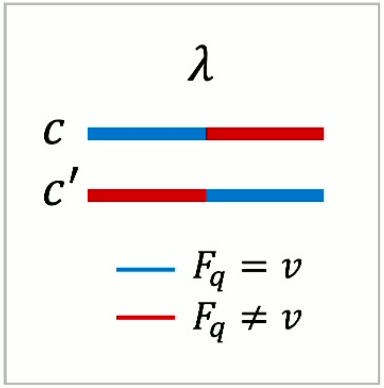
- **Direct influence:** Given $\mathcal{M}$, q, and $\{c, c'\} \succ q$:

$$\Delta_{c,c'}(F_q) \doteq \Pr\left[\left\{\lambda : F_q(\lambda, c) \neq F_q(\lambda, c')\right\}\right]$$

  - Probability of latent state for which context change affects measurement outcome

- **Hidden influences** (finite version)

$$\Pr\left[\left\{\lambda : F_q(\lambda, c) = v, F_q(\lambda, c') \neq v\right\}\right] > 0$$
$$\Pr\left[\left\{\lambda : F_q(\lambda, c) \neq v, F_q(\lambda, c') = v\right\}\right] > 0$$



- **Aligned model**
  - No hidden influences, for any $q, v \in \mathcal{O}_q, \{c, c'\} \succ q$
  - Formalizes no-conspiracy principle: Causal effect matches distribution difference

$$\Delta_{c,c'}(F_q) = \left|P(F_q|c) - P(F_q|c')\right|_{TV}$$

Example: PR Box (Popescu & Rohrlich, 1994)

M

$\Pr[F_q = 1|c]$

	A_1	B_1	A_2	B_2
c_{11}	½	½		
c_{21}		½	½	
c_{22}			½	½
c_{12}	½			½

$\Pr[F_q = F_{q'}]$

q \ q'	B_1	B_2
A_1	1	0
A_2	1	1

$\mathcal{M}$

	λ_1	λ_2
$\Pr[\Lambda = \lambda]$	½	½
$F_{A_1}(c_{11}, \lambda)$	1	-1
$F_{A_1}(c_{12}, \lambda)$	1	-1
$F_{B_1}(c_{11}, \lambda)$	1	-1
$F_{B_1}(c_{21}, \lambda)$	-1	1
$F_{A_2}(c_{21}, \lambda)$	-1	1
$F_{A_2}(c_{22}, \lambda)$	-1	1
$F_{B_2}(c_{12}, \lambda)$	-1	1
$F_{B_2}(c_{22}, \lambda)$	-1	1

$\Delta_{c_{11},c_{12}}(F_{A_1}) = 0$

$\Delta_{c_{11},c_{21}}(F_{B_1}) = 1$

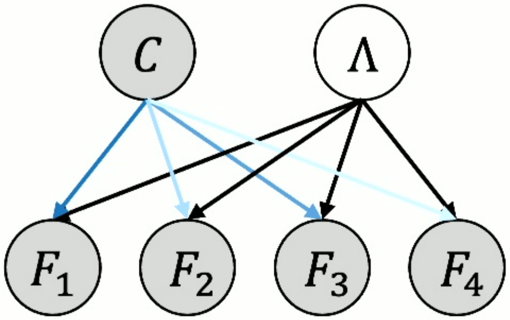
$\Delta_{c_{21},c_{22}}(F_{A_2}) = 0$

$\Delta_{c_{12},c_{22}}(F_{B_2}) = 0$

Model-based Contextuality

- **M-contextual**
  - Given any model $\mathcal{M}$ for M, there exist $q \prec \{c, c'\}$ such that $\Delta_{c,c'}(F_q)$ is greater than necessary
  - Modeling full system requires more direct influence than modeling each observable alone
- **M-noncontextual**
  - $\exists\mathcal{M}$ simultaneously minimizing all direct influences

$$\forall_{\mathcal{M}',q\prec\{c,c'\}}\left[(\mathcal{M}' \vDash P) \Rightarrow \left(\Delta^{\mathcal{M}'}_{c,c'}(F_q) \geq \Delta^{\mathcal{M}}_{c,c'}(F_q)\right)\right]$$



Model-based Contextuality

- **Theorem 1**. For nondisturbing systems, M-contextuality ⇔ KS-contextuality

  Proof: Noncontextual model ⇔ $\forall_{q\prec c,c'}\left[\Delta_{c,c'}\left(F_q\right) = 0\right]$

- **Theorem 2**. An arbitrary measurement system is M-noncontextual iff it admits an aligned model
  - M-contextual iff all models contain hidden influences

  Proof: Hidden influence implies $\Delta_{c,c'}\left(F_q\right)$ can be reduced

  Lemma. Given M and $q \prec c, c'$, the minimum direct influence over all models $\mathcal{M}$ for M is:

$$\min_{\mathcal{M}} \Delta_{c,c'}\left(F_q\right) = |P(q|c) - P(q|c')|_{TV}$$

M-contextuality and CbD-contextuality

- Probabilistic coupling (Kujala et al. 2015; Thorisson 2000)
  - Jointly distributed $\{S_q^c : q \prec c\}$ with $\boldsymbol{S}^c \sim \mu_c$ for all c
- Multimaximal coupling (Dzhafarov & Kujala 2017)
  - Simultaneously minimizes $\Pr[S_q^c \neq S_q^{c'}]$ for all $q \prec \{c, c'\}$
- CbD-contextuality: Non-existence of multimaximal coupling
- **Theorem 3**. M-contextuality ⇔ CbD-contextuality

  Proof: Translation between models and couplings satisfying

$$\Delta_{c,c'}(F_q) = \Pr[S_q^c \neq S_q^{c'}]$$

- **Theorem 4**. A system is CbD-noncontextual iff it admits an aligned model
  - *Proof*: Follows from Theorems 2 and 3

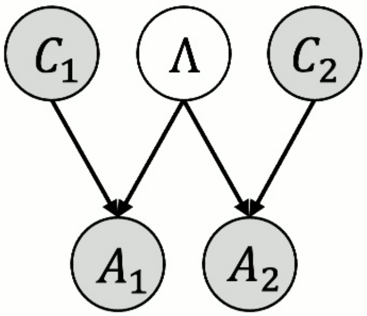
Bell Representation

- Multipartite model
  - $\widetilde{\mathcal{M}} = (\Lambda, (A_k)_{k\in\mathcal{K}}, (C_k)_{k\in\mathcal{K}})$
- Factoring of contexts
  - $C = (C_k)_{k\in\mathcal{K}}$
  - $\mathcal{C} = \prod_k \mathcal{Q}_k$
- Signaling
  - For $c_k = c'_k$:

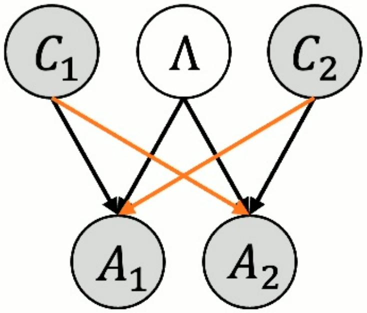
$$\Delta_{c,c'}(A_k) = \Pr[\{\lambda: A_k(\lambda, c) \neq A_k(\lambda, c')\}] = \Delta_{c,c'}(F_{c_k})$$

  - Hidden signaling: cancels in observed marginals
    (cf. Atmanspacher & Filk 2019)

G^{Loc}: No Signaling



G^{pd}: Signaling



Bell systems: Parallel results

Proposition 4 [≃Prop 1]. Any Bell system M admits a model on G^{pd}

Proposition 5. M is M-noncontextual iff it has a model $\widetilde{\mathcal{M}}$ that minimizes $\Delta_{c,c'}(A_k)$ for all k, c, c' ($c_k = c'_k$)

Proposition 6 [≃Prop 2, Thm 1]. A nonsignaling Bell system is KS-noncontextual and M-noncontextual iff it admits a model on G^{loc}

Theorem 5 [≃Thms 2-4]. M is M-noncontextual and CbD-noncontextual iff it admits a model $\widetilde{\mathcal{M}}$ without hidden signals

Interim Summary

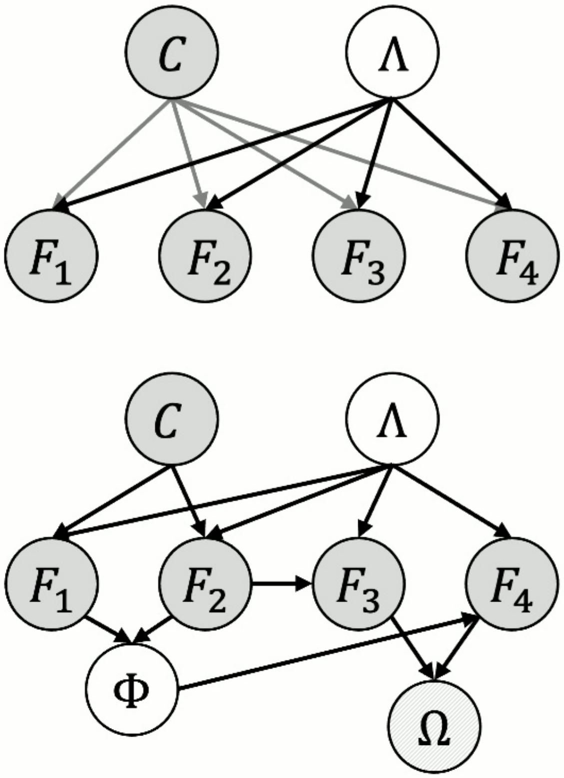
- Quantify direct influence of context
- Define hidden influence
- Equivalence of three definitions of extended contextuality
  - Minimizing direct influence or signaling (M-contextuality)
  - Prohibiting hidden influence or hidden signaling
  - CbD-contextuality
  - ○ Reduce to KS-contextuality for nondisturbing/nonsignaling systems

Interim Summary

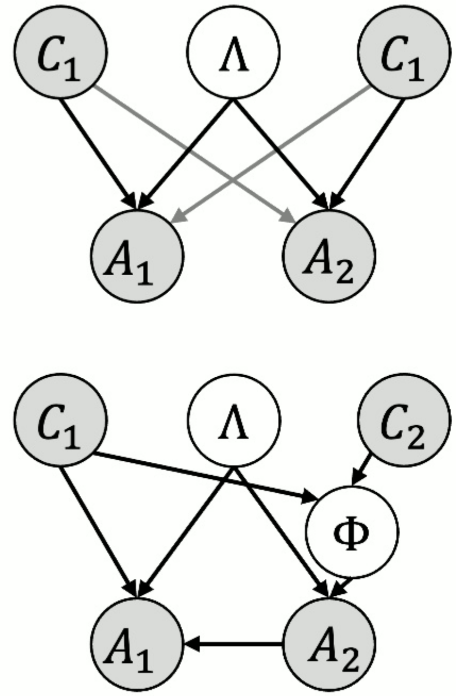
- Quantify direct influence of context
- Define hidden influence
- Equivalence of three definitions of extended contextuality
  - Minimizing direct influence or signaling (M-contextuality)
  - Prohibiting hidden influence or hidden signaling
  - CbD-contextuality
  - ○ Reduce to KS-contextuality for nondisturbing/nonsignaling systems
- Two forms of graphical models
  - Kochen-Specker representation
  - Bell representation

Varieties of Causal Models

KS representation



Bell representation



Motivating example

$P(XY|c_1)$:

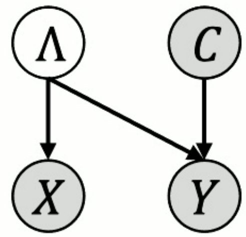
	$Y=-1$	$Y=1$
$X=-1$	$1/2$	0
$X=1$	0	$1/2$

$P(XY|c_2)$:

	$Y=-1$	$Y=1$
$X=-1$	0	$1/2$
$X=1$	$1/2$	0

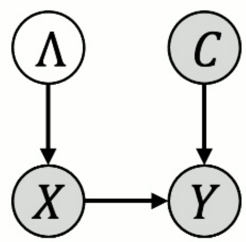
Dzhafarov & Kujala (2016)

	λ_1	λ_2
$X(\Lambda)$	-1	1



$Y(\Lambda, C)$:

	λ_1	λ_2	
c_1	-1	1	$\langle Y(c_1)\rangle = 0$
c_2	1	-1	$\langle Y(c_2)\rangle = 0$



	λ_1	λ_2
$X(\Lambda)$	-1	1

$Y(X, C)$:

	$X=-1$	$X=1$
c_1	-1	1
c_2	1	-1

Questions

- Do different choices of G yield different variants of contextuality?
  - Does consideration of different causal paths change what is classified as conspiracy?
  - Or, do all representations of causality yield the same notion of contextuality?
  - Pearl & Cavalcanti (2021): No-fine-tuning theorem holds for arbitrary graphs
- When is G-contextuality sensible or useful?
  - Minimal requirement: agree with KS-contextuality on nondisturbing systems
- Main conclusions
  - Other graph structures yield alternative extensions of contextuality
  - but they're degenerate or unreasonable, failing to agree with KS-contextuality
  - Canonical graph structure (G^{pd}) and M-contextuality/CbD-contextuality seem to be the only well-behaved notion of extended contextuality within this space

Discussion

- No-go interpretation of extended contextuality
- Translation between theoretical approaches
- Finite datasets and model selection
- Essential role of restrictions on direct influence
- No-hidden-influence as a physical principle
- A priori exclusion of outcome dependence
- Domain-specific considerations for causal graphs

References

Abramsky S & Brandenburger A (2011). *New Journal of Physics*, **13**: 113036.

Atmanspacher H & Filk T (2019). *Quanta and Mind*, **414**: 117-127.

Bell JS (1964). *Physics,* **1**: 195–200.

Bell JS (1976). *Epistemological Letters*, **9**: 11–24.

Cavalcanti EG (2018). *Physical Review X,* **8**: 021018.

Clauser JF, Horne MA, Shimony A & Holt RA (1969). Proposed experiment to test local hidden-variable theories. *Physical review letters, 23*(15), 880.

Dzhafarov EN & Kujala JV (2016). *Journal of Mathematical Psychology,* **74**: 11-33.

Dzhafarov EN & Kujala JV (2017). *Lecture Notes in Computer Science,* ***10106****:* 16-32.

Fine A (1982). *Physics Review Letters*, **48**: 291-295.

Jarrett JP (1984). *Noûs*, **18**: 569–589.

Jones M (2019). *Philosophical Transactions of the Royal Society A,* **377**: 20190133.

Jordan MI (1999). *Learning in graphical models*. MIT Press.

Kochen S & Specker EP (1967). *Journal of Mathematics and Mechanics*, **17**: 59–87.

Kujala JV, Dzhafarov EN & Larsson J-A (2015). Physical Review Letters, **115**: 150401.

Pearl J. (1999). *Causality: models, reasoning and inference*. Cambridge University Press.

Pearl JC & Cavalcanti EG (2021). *Quantum*, **5**: 518.

Popescu S & Rohrlich D (1994). *Foundations of Physics,* **24**: 379–385.

Shimony A (1986). In *Quantum Concepts in Space and Time* (pp. 182–203). Oxford University Press.

Tenenbaum JB, Kemp C, Griffiths TL & Goodman ND (2011). *Science*, **331**: 1279-1285.

Thorisson H. (2000). *Coupling, Stationarity, and Regeneration*. Springer.

Wood CJ & Spekkens RW (2015). *New Journal of Physics,* **17**: 033002.

Motivating example

$P(XY|c_1)$:

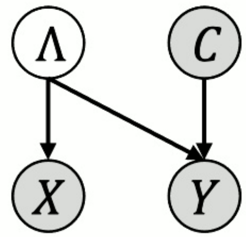
	$Y=-1$	$Y=1$
$X=-1$	$1/2$	0
$X=1$	0	$1/2$

$P(XY|c_2)$:

	$Y=-1$	$Y=1$
$X=-1$	0	$1/2$
$X=1$	$1/2$	0

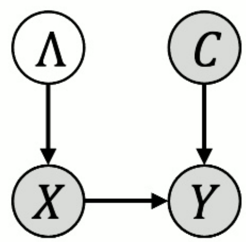
Dzhafarov & Kujala (2016)

	λ_1	λ_2
$X(\Lambda)$	-1	1



$Y(\Lambda, C)$:

	λ_1	λ_2	
c_1	-1	1	$\langle Y(c_1)\rangle = 0$
c_2	1	-1	$\langle Y(c_2)\rangle = 0$



	λ_1	λ_2
$X(\Lambda)$	-1	1

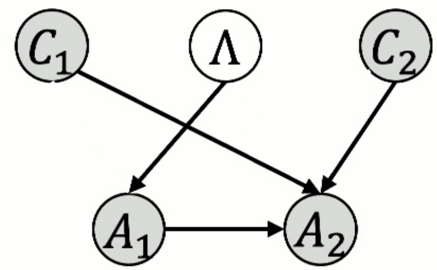
$Y(X, C)$:

	$X=-1$	$X=1$
c_1	-1	1
c_2	1	-1

Parameter and outcome dependence, $\mathcal{G}^{\mathrm{od,pd}}$

- **Proposition 10.** *KS-contextuality* $\nRightarrow$ $\mathcal{G}^{\mathrm{od,pd}}$*-contextuality*

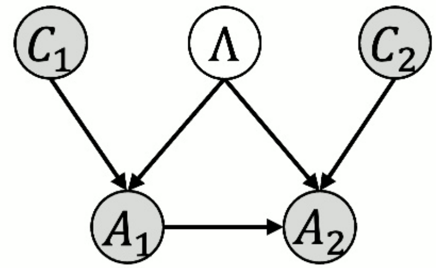
  Proof: PR Box



$P(A_1A_2|C_1C_2)$:

		$C_2 = 1$		$C_2 = 2$	
		$A_2 = -1$	$A_2 = 1$	$A_2 = -1$	$A_2 = 1$
$C_1 = 1$	$A_1 = -1$	$1/2$	0	$1/2$	0
	$A_1 = 1$	0	$1/2$	0	$1/2$
$C_1 = 2$	$A_1 = -1$	$1/2$	0	0	$1/2$
	$A_1 = 1$	0	$1/2$	$1/2$	0

Outcome dependence, $\mathcal{G}^{\mathrm{od}}$



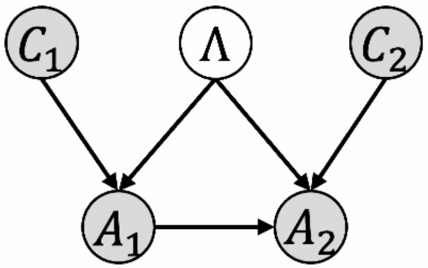
- Also too flexible
- **Proposition 11.** *KS-contextuality does not imply $\mathcal{G}^{\mathrm{od}}$-contextuality*

  Proof: PR box again

A_1:	λ_1	λ_2
$C_1 = 1$	-1	1
$C_1 = 2$	1	-1

A_2:	λ_1	λ_2
$C_2 = 1$, $A_1 = -1$	-1	-1
$C_2 = 1$, $A_1 = 1$	1	1
$C_2 = 2$, $A_1 = -1$	-1	1
$C_2 = 2$, $A_1 = 1$	-1	1

Outcome dependence, $\mathcal{G}^{\mathrm{od}}$



- Also too flexible
- **Proposition 11.** *KS-contextuality does not imply $\mathcal{G}^{\mathrm{od}}$-contextuality*
  Proof: PR box again

A_1:

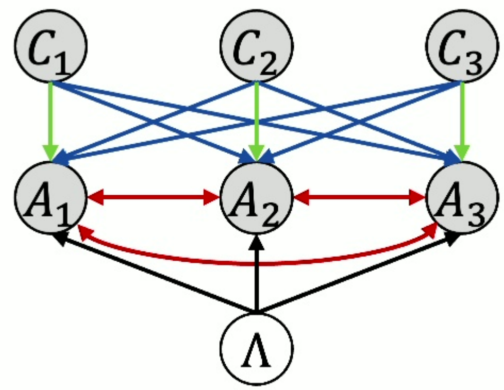
	λ_1	λ_2
$C_1 = 1$	-1	1
$C_1 = 2$	1	-1

A_2:

	λ_1	λ_2
$C_2 = 1$, $A_1 = -1$	-1	-1
$C_2 = 1$, $A_1 = 1$	1	1
$C_2 = 2$, $A_1 = -1$	-1	1
$C_2 = 2$, $A_1 = 1$	-1	1

Graph reduction

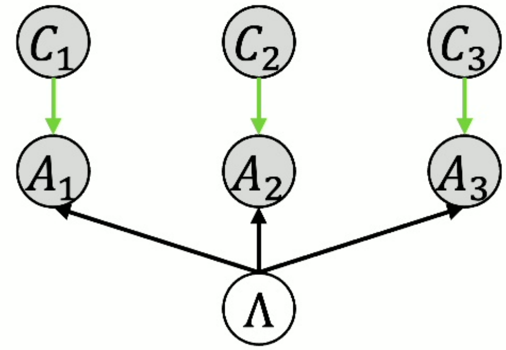
- Restrict to partitioned models (Bell representation)
  - No loss of generality: can assign a different observer to each observable
- Three types of nodes: C_k (setting), A_k (outcome), Λ (hidden)
- Eliminate and combine hidden variables into one, with $Pa(\Lambda) = \emptyset$
- Exclude superdeterminism: $Pa(C_k) = \emptyset$
- Possible connections
  - $C_k \to A_k$ (local settings)
  - $\Lambda \to A_k$ (source dependence)
  - $C_k \to A_{k'}$ (parameter dependence)
  - $A_k \to A_{k'}$ (outcome dependence)



Possible graph structures

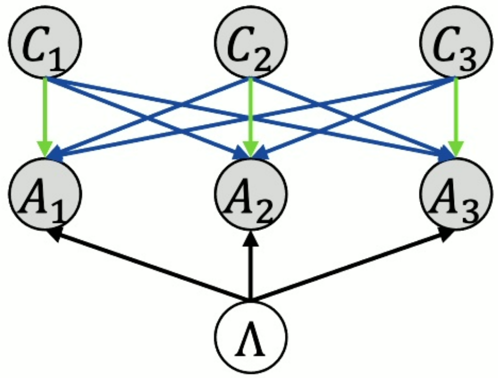
G^{Loc}

(Bell local)



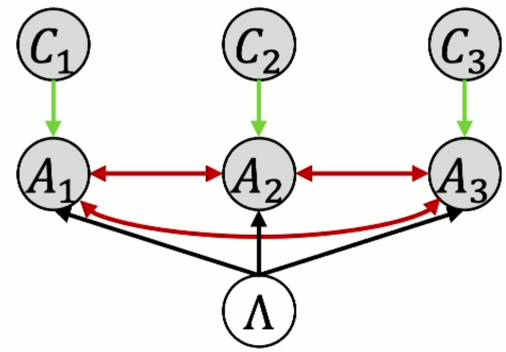
G^{pd}

(parameter dependence)



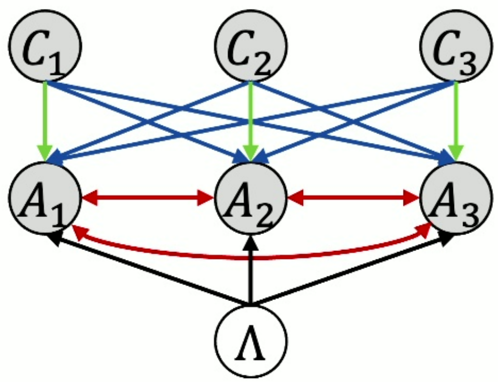
$\mathcal{G}^{\text{od}}$

(outcome dependence)



$\mathcal{G}^{\text{od,pd}}$

(outcome + parameter dependence)



etc.