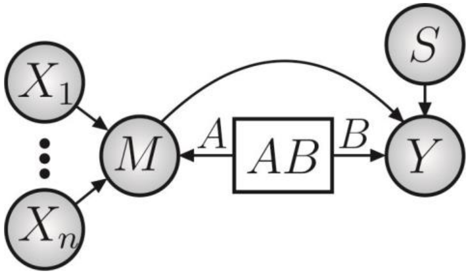
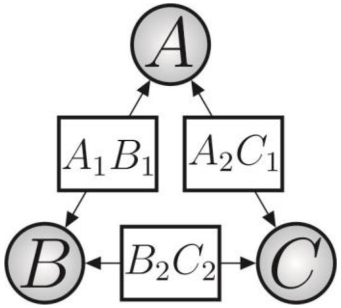
Information-Theoretic Implications of Classical and Quantum Causal Structures

Rafael Chaves

QIP 2015





RC, L. Luft, T. Maciel, D. Gross, D. Janzing, B. Schölkopf (arXiv:1407.2256)

RC, C. Majenz, D. Gross (arXiv:1407.3800)

RC, C. Majenz & D. Gross, Nature Communications 6, 5766 (2015)

RC, L. Luft, T. Maciel, D. Gross, D. Janzing, B. Schölkopf, Proceedings of Uncertainty in Artificial Intelligence 2014

A joint work with

UNI FREIBURG

David Gross

Lukas Luft

Thiago Maciel

MAX-PLANCK-GESELLSCHAFT

Bernhard Schölkopf

Dominik Janzing

UNIVERSITY OF COPENHAGEN

Christian Majenz

Given some empirically observable variables, which *correlations* between them are compatible with a presumed *causal structure*?

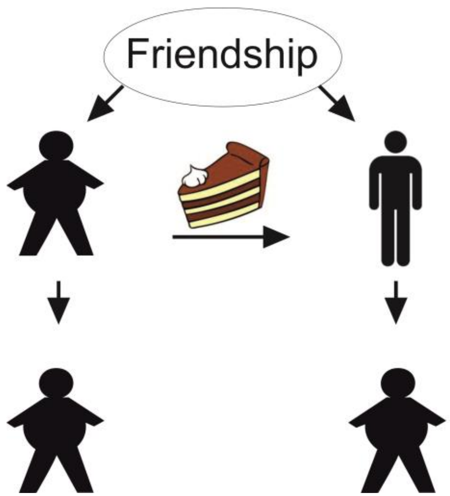
- Distinguishing direct influence from common cause… Is obesity contagious?

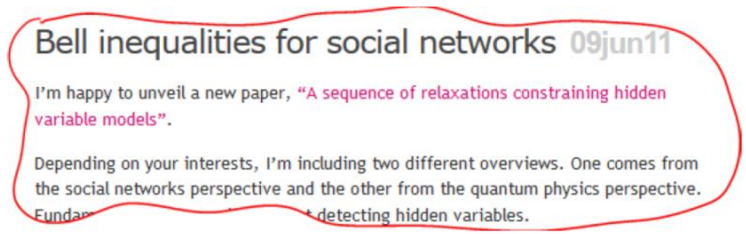
Bell inequalities for social networks 09jun11

I'm happy to unveil a new paper, "A sequence of relaxations constraining hidden variable models".

Depending on your interests, I'm including two different overviews. One comes from the social networks perspective and the other from the quantum physics perspective. Fundam … detecting hidden variables.

- Distinguishing direct influence from common cause… Is obesity contagious?

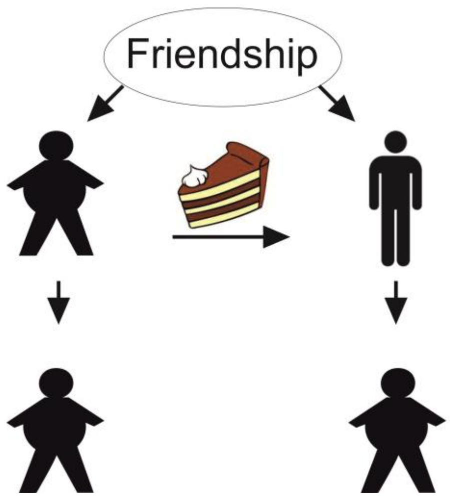


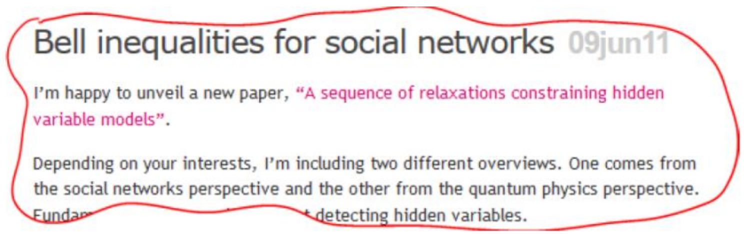

Bell inequalities for social networks 09jun11

I'm happy to unveil a new paper, "A sequence of relaxations constraining hidden variable models".

Depending on your interests, I'm including two different overviews. One comes from the social networks perspective and the other from the quantum physics perspective. Fundam… detecting hidden variables.

- Distinguishing direct influence from common cause… Is obesity contagious?

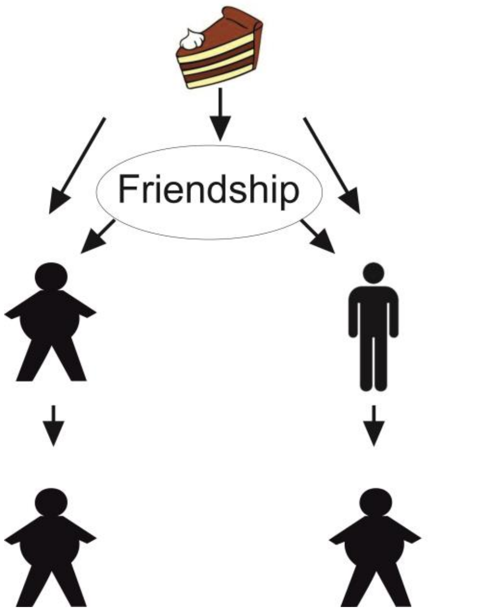




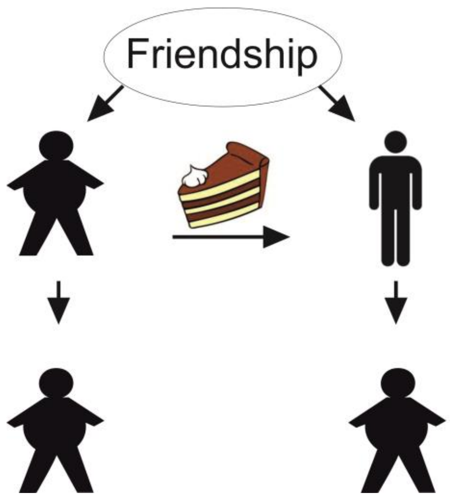
Bell inequalities for social networks 09jun11

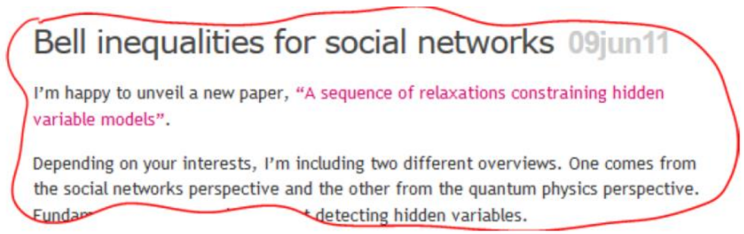
I'm happy to unveil a new paper, "A sequence of relaxations constraining hidden variable models".

Depending on your interests, I'm including two different overviews. One comes from the social networks perspective and the other from the quantum physics perspective. Fundam detecting hidden variables.



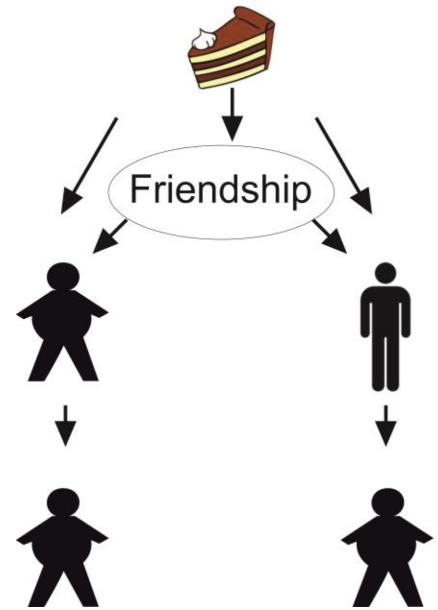
- Distinguishing direct influence from common cause… Is obesity contagious?




Bell inequalities for social networks 09jun11

I'm happy to unveil a new paper, "A sequence of relaxations constraining hidden variable models".

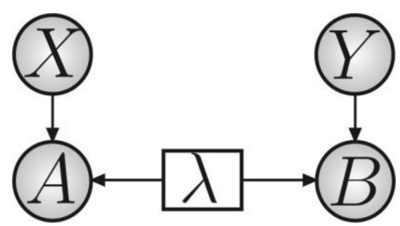
Depending on your interests, I'm including two different overviews. One comes from the social networks perspective and the other from the quantum physics perspective. Fundam… detecting hidden variables.



- Bell's Theorem: Quantum correlations are incompatible with "local realism".

ON THE EINSTEIN PODOLSKY ROSEN PARADOX*

J. S. BELL†



Outline

- Classical causal structures
- The information-theoretic approach to classical causal inference
- The generalization to quantum causal structures
- Where to go from here?

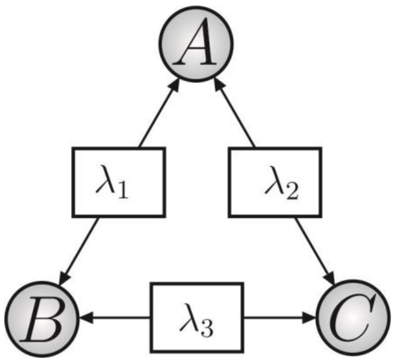
- **Classical causal structures**
- The information-theoretic approach to classical causal inference
- The generalization to quantum causal structures
- Where to go from here?

Classical Causal Structures

- For n variables $X_1, \ldots, X_n$, the causal relationships are encoded in a causal structure, represented by a directed acyclic graph (**DAG**)
- *i*th variable being a deterministic function

$$x_i = f_i(pa_i, u_i)$$

of its parents $\mathbf{pa_i}$ and „local randomness“ $\mathbf{u_i}$



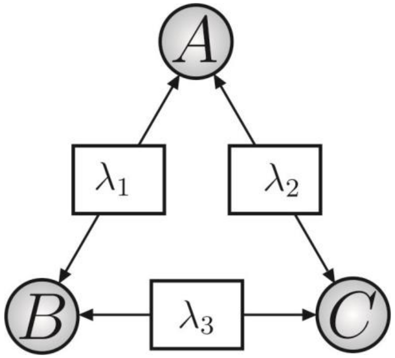
[See J. Pearl, *Causality*]

Classical Causal Structures

- For n variables X_1, ... ,X_n, the causal relationships are encoded in a causal structure, represented by a directed acyclic graph (**DAG**)
- *i*th variable being a deterministic function

$$\mathbf{x_i = f_i(pa_i, u_i)}$$

of its parents $\mathbf{pa_i}$ and „local randomness" $\mathbf{u_i}$



- Causal relationships are encoded in the conditional independencies (CIs) implied by the DAG

$$p(\lambda_1, \lambda_2) = p(\lambda_1)p(\lambda_2)$$
$$p(A, B|\lambda_1) = p(A|\lambda_1)p(B|\lambda_1)$$

...

[See J. Pearl, *Causality*]

Is a given probability distribution compatible with a presumed *causal structure*?

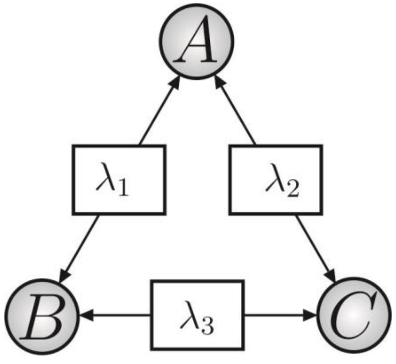
Is a given probability distribution compatible with a presumed *causal structure*?

⇨ Iff the given probability distribution fullfils all the CIs implied by the DAG

Is a given probability distribution compatible with a presumed *causal structure*?

➡ Iff the given probability distribution fullfils all the CIs implied by the DAG

Example: Is a given $p(\lambda_1, \lambda_2, \lambda_3, A, B, C)$ compatible with

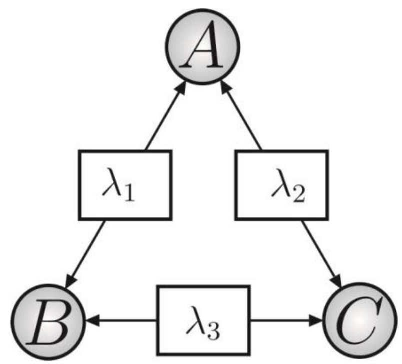


?

Is a given probability distribution compatible with a presumed *causal structure*?

Iff the given probability distribution fullfils all the CIs implied by the DAG

Example: Is a given $p(\lambda_1, \lambda_2, \lambda_3, A, B, C)$ compatible with



?

$$p(\lambda_1, \lambda_2) = p(\lambda_1)p(\lambda_2)$$

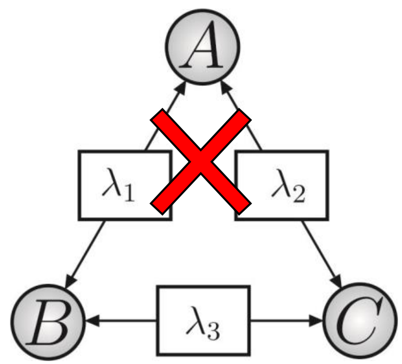
$$p(A, B|\lambda_1) = p(A|\lambda_1)p(B|\lambda_1)$$

...

Is a given probability distribution compatible with a presumed *causal structure*?

Iff the given probability distribution fullfils all the CIs implied by the DAG

Example: Is a given $p(\lambda_1, \lambda_2, \lambda_3, A, B, C)$ compatible with



?

$$p(\lambda_1, \lambda_2) = p(\lambda_1)p(\lambda_2)$$

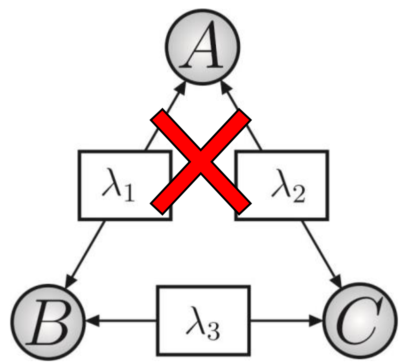
$$p(A, B|\lambda_1) \neq p(A|\lambda_1)p(B|\lambda_1)$$

...

Is a given probability distribution compatible with a presumed *causal structure*?

⇨ Iff the given probability distribution fullfils all the CIs implied by the DAG

Example: Is a given $p(\lambda_1, \lambda_2, \lambda_3, A, B, C)$ compatible with



?

$$p(\lambda_1, \lambda_2) = p(\lambda_1)p(\lambda_2)$$

$$p(A, B|\lambda_1) \neq p(A|\lambda_1)p(B|\lambda_1)$$

...

- If the full probability distribution (of all nodes in a DAG) is available, CIs hold all information required to solve the compatibility problem

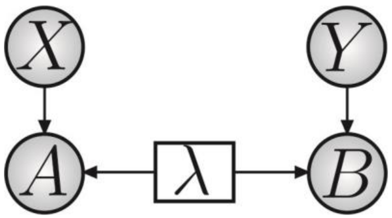
However…

Marginal Scenarios

- Usually and for a variety of reasons not all variables in a DAG are observable, i.e., not all CIs are available from empirical data

Marginal Scenarios

- Usually and for a variety of reasons not all variables in a DAG are observable, i.e., not all CIs are available from empirical data

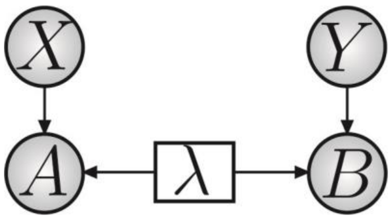


$$p(x, y, \lambda) = p(x)p(y)p(\lambda)$$
$$p(a|x, y, \lambda) = p(a|x, \lambda)$$

...

Marginal Scenarios

- Usually and for a variety of reasons not all variables in a DAG are observable, i.e., not all CIs are available from empirical data



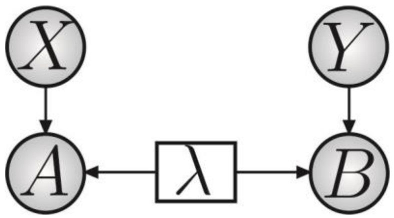
$$p(x, y, \lambda) = p(x)p(y)p(\lambda)$$
$$p(a|x, y, \lambda) = p(a|x, \lambda)$$

...

$$\begin{aligned} p(a, b|x, y) &= \sum_\lambda p(a, b, \lambda|x, y) \\ &= \sum_\lambda p(a, b|x, y, \lambda)p(\lambda|x, y) \\ &= \sum_\lambda p(a|x, \lambda)p(b|y, \lambda)p(\lambda) \end{aligned}$$

Marginal Scenarios

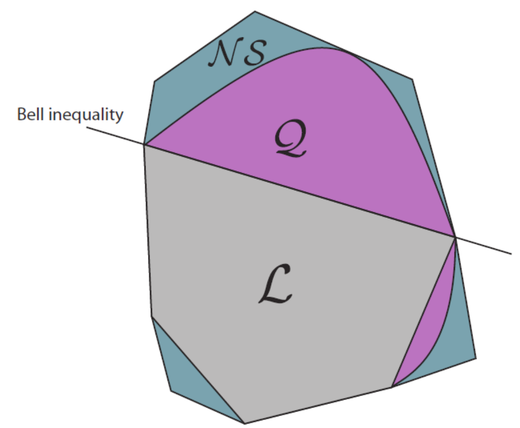
- Usually and for a variety of reasons not all variables in a DAG are observable, i.e., not all CIs are available from empirical data



$$p(x,y,\lambda) = p(x)p(y)p(\lambda)$$
$$p(a|x,y,\lambda) = p(a|x,\lambda)$$

...

$$\begin{aligned} p(a,b|x,y) &= \sum_{\lambda} p(a,b,\lambda|x,y) \\ &= \sum_{\lambda} p(a,b|x,y,\lambda)p(\lambda|x,y) \\ &= \sum_{\lambda} p(a|x,\lambda)p(b|y,\lambda)p(\lambda) \end{aligned}$$



Pic from **[Rev. Mod. Phys. 86, 419 (2014)]**

- CIs impose non-trivial constraints on the level of the observable variables, for example, Bell inequalities.

Challenge

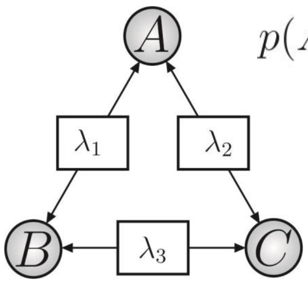
- Describe marginals compatible with DAGs

Challenge

- Describe marginals compatible with DAGs
- The observable probability contains the full causal information empirically available...

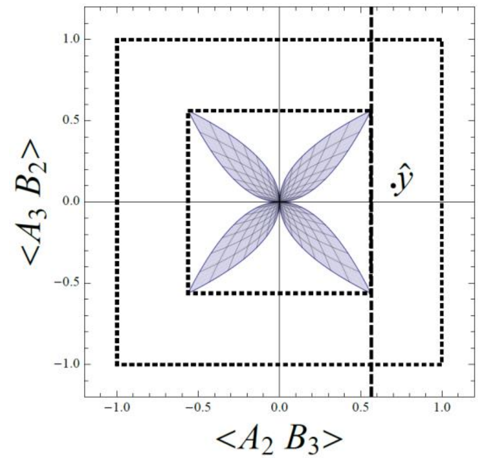
Challenge

- Describe marginals compatible with DAGs
- The observable probability contains the full causal information empirically available...
- ..very difficult, non-convex sets (algebraic geometry methods required, see for instance **[Geiger & Meek, UAI 1999]**)



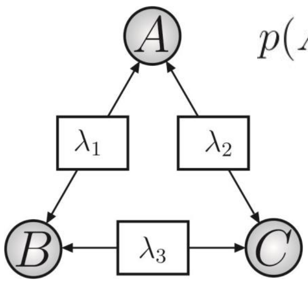
$$p(A,B,C) = \int d\lambda_1 d\lambda_2 d\lambda_3 p(\lambda_1)p(\lambda_2)p(\lambda_3) \, p(A|\lambda_1,\lambda_2)p(B|\lambda_1,\lambda_3)p(C|\lambda_2,\lambda_3)$$

Picture from **[Steeg & Galstyan, UAI 2011]**



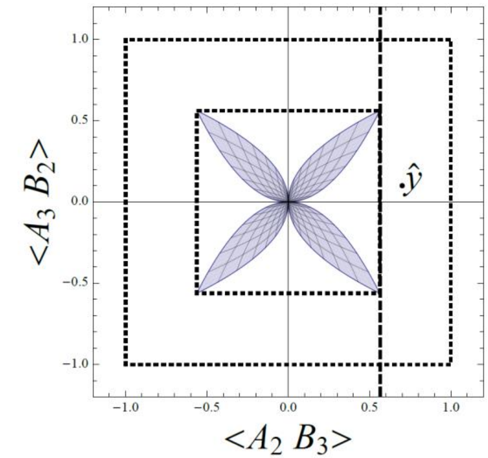
Challenge

- Describe marginals compatible with DAGs
- The observable probability contains the full causal information empirically available...
- ..very difficult, non-convex sets (algebraic geometry methods required, see for instance **[Geiger & Meek, UAI 1999]**)



$$p(A,B,C) = \int d\lambda_1 d\lambda_2 d\lambda_3 p(\lambda_1)p(\lambda_2)p(\lambda_3) \, p(A|\lambda_1,\lambda_2)p(B|\lambda_1,\lambda_3)p(C|\lambda_2,\lambda_3)$$

Picture from **[Steeg & Galstyan, UAI 2011]**



<u>Our idea</u>

Rely on entropic information!

- Concise characterization as a convex set
- Naturally encodes the causal constraints
- Quantitative and stable tool

- Causal structures
- The information-theoretic approach to classical causal inference
- The generalization to quantum causal structures
- Where to go from here?

[T. Fritz and RC, IEEE Trans. Inf. Th. 59, 803 (2013)]

[RC, L. Luft, D. Gross, NJP 16, 043001 (2014)]

[RC, L. Luft, T. Maciel, D. Gross, D. Janzing, B. Schölkopf, UAI 2014]

Causal Entropic cone

Step 1/3: *Unconstrained, global object*

- Entropic vector $V \in \mathbb{R}^{2^n}$:each entry is the Shannon entropy $H(X_S)$ indexed by $S \subset \{1, \ldots, n\}$

Example: *2 vars* $\rightarrow$ $(H(\emptyset), H(A), H(B), H(A, B))$

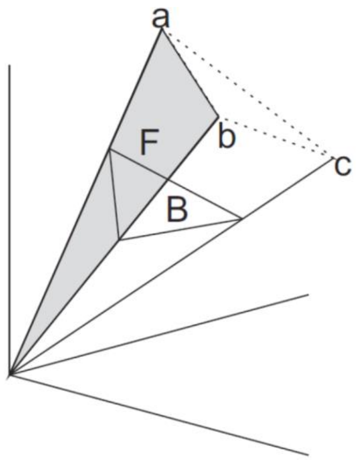
Causal Entropic cone

Step 1/3: *Unconstrained, global object*

- Entropic vector $v \in \mathbb{R}^{2^n}$:each entry is the Shannon entropy $H(X_S)$ indexed by $S \subset \{1, \ldots, n\}$

Example: *2 vars* $\rightarrow$ $(H(\emptyset), H(A), H(B), H(A,B))$

- Defines a convex cone. Structure not fully understood, but...



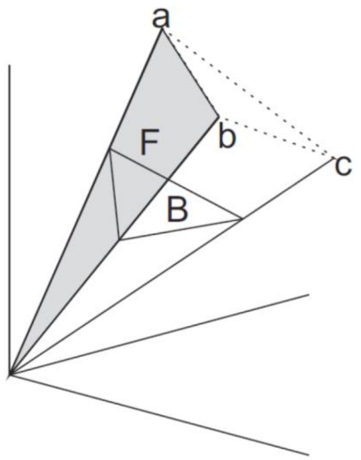
Causal Entropic cone

Step 1/3: *Unconstrained, global object*

- Entropic vector $v \in \mathbb{R}^{2^n}$:each entry is the Shannon entropy $H(X_S)$ indexed by $S \subset \{1, \ldots, n\}$

Example: *2 vars* → $(H(\emptyset), H(A), H(B), H(A, B))$

- Defines a convex cone. Structure not fully understood, but...



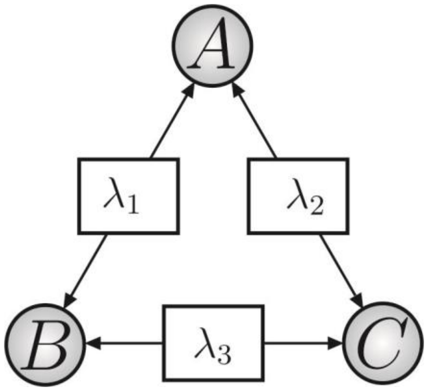
- ...contained in Shannon Cone Γ_n, defined by strong subadditivity and monotonicity

$$H(A, B, C) + H(A) \leq H(A, B) + H(A, C)$$

$$H(A, B) \leq H(A, B, C)$$

Causal Entropic cone

Step 2/3: *Choose candidate structure and add causal constraints*



- Piece of cake! Conditional independences are naturally embedded in mutual informations

$$p(\lambda_1, \lambda_2) = p(\lambda_1)p(\lambda_2)$$
$$p(A, B|\lambda_1) = p(A|\lambda_1)p(B|\lambda_1)$$
$$\longrightarrow$$
$$I(\lambda_1 : \lambda_2) = 0$$
$$I(A : B|\lambda_1) = 0$$

- We can even relax (stable!)

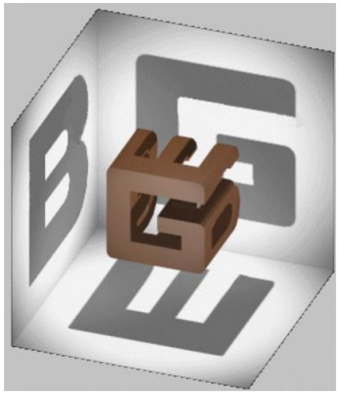
$$I(\lambda_1 : \lambda_2) \leq \epsilon_1$$
$$I(A : B|\lambda_1) \leq \epsilon_2$$

- *C: set of constraints*

- New global cone $\Gamma_n \cap C$ of entropies subject to causal structure

Causal Entropic cone

Step 3/3: *Marginalize to* $\mathcal{M}$



- $\mathcal{M} \subset 2^{\{1,\ldots,n\}}$: set of joint observables

- **Geometrically trivial**:

  just restrict $\Gamma_n \cap C$ to observable coordinates

- **Algorithmically costly**: $\Gamma_n \cap C$ represented in terms of inequalities (use, eg, Fourier-Motzkin elimination)

Final result: description of marginal, causal entropic cone $(\Gamma_n \cap C)_{|\mathcal{M}}$ in terms of „entropic Bell inequalities“

[T. Fritz and RC, IEEE Trans. Inf. Th. 59, 803 (2013)]

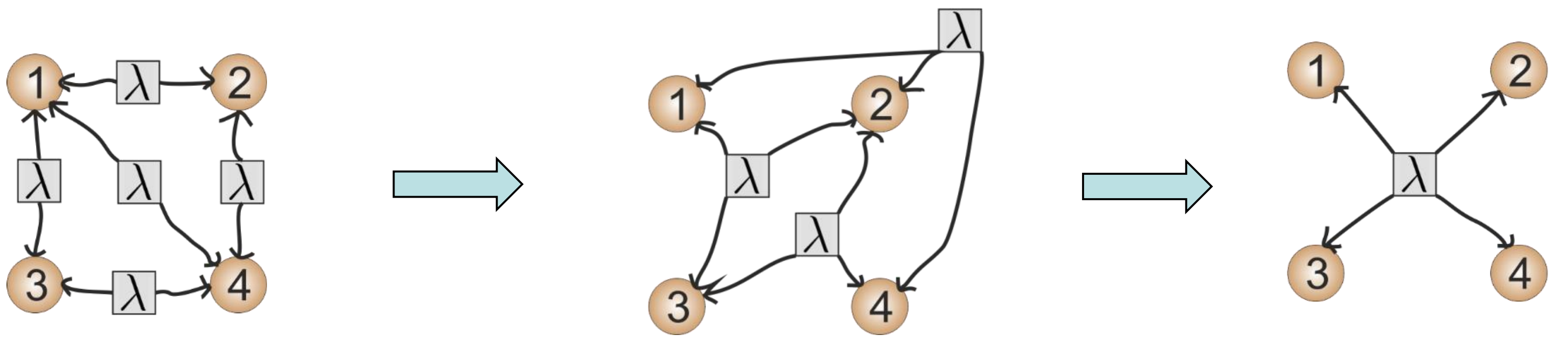
[RC, L. Luft, D. Gross, NJP 16, 043001 (2014)]

[RC, L. Luft, T. Maciel, D. Gross, D. Janzing, B. Schölkopf, UAI 2014]

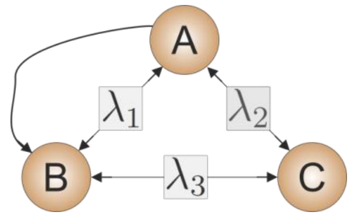
Applications

- <u>Entropic Bell inequalities</u> **[Braunstein & Caves PRL 61, 662 (1988)]**

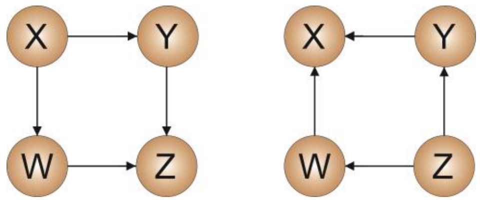
- <u>Common ancestors problem</u>: Can the correlations between n observable variables be explained by independent common ancestors connecting at most M of them? **[Steudel & Ay, arXiv:1010.5720]**



- <u>Quantifying Causal Influences</u> **[D. Janzing et al, Ann. of Stat. 41, 2324 (2013)]**



- <u>Witnessing direction of causation from pairwise information</u>



- Causal structures
- The information-theoretic approach to (classical) causal inference
- **The generalization to quantum causal structures**
- Where to go from here?

[RC, C. Majenz & D. Gross, Nat. Comm. 6, 5766 (2015)]

Quantum Causal Structures

- Different formulations have been proposed. For an incomplete list see:

  [R. Tucci, arXiv:quant-ph/0701201 (2007)]
  [M. S. Leifer & R. W. Spekkens, Phys. Rev. A 88, 052130 (2013)]
  [T. Fritz, arXiv:1404.4812]
  [J.Henson, R. Lal, M. F. Pusey New J. Phys. 16, 113043 (2014)]
  [J. Pienaar & C. Brukner, arXiv:1406.0430]

Quantum Causal Structures

- Different formulations have been proposed. For an incomplete list see:

  [R. Tucci, arXiv:quant-ph/0701201 (2007)]
  [M. S. Leifer & R. W. Spekkens, Phys. Rev. A 88, 052130 (2013)]
  [T. Fritz, arXiv:1404.4812]
  [J.Henson, R. Lal, M. F. Pusey New J. Phys. 16, 113043 (2014)]
  [J. Pienaar & C. Brukner, arXiv:1406.0430]

- We propose a graphical representation that allow us to generalize the information-theoretic approach

- Informally, a quantum causal structure specifies the functional dependency between a collection of quantum states and classical variables.

Quantum Causal Structures

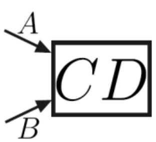
Building Blocks

- Classical variable

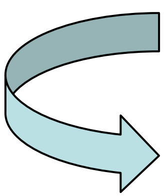
- Quantum State

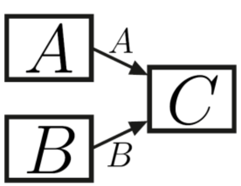
$$\boxed{AB} \triangleq \varrho_{AB}$$

- Quantum Operation (CPTP map)



$$\triangleq \Phi_{AB\to CD} : A \otimes B \to C \otimes D$$

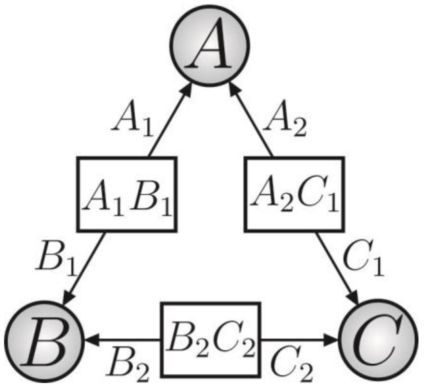




$$\triangleq \varrho_C = \Phi_{AB\to C}(\varrho_A \otimes \varrho_B)$$

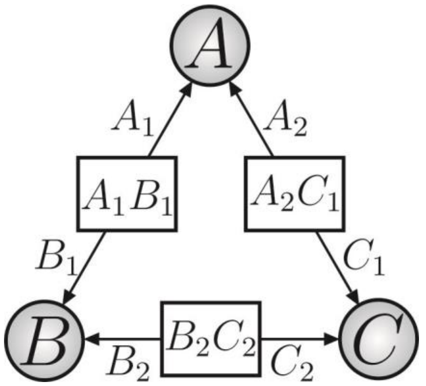
Entropic description of Quantum Causal Structures

[RC, C. Majenz & D. Gross, Nat. Comm. 6, 5766 (2015)]



Entropic description of Quantum Causal Structures

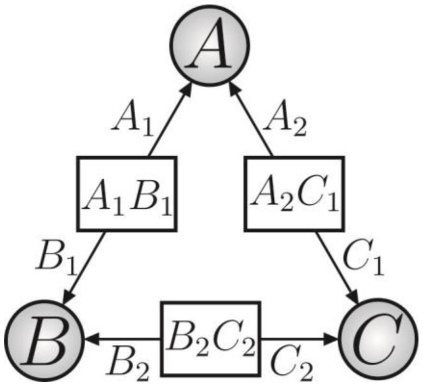
[RC, C. Majenz & D. Gross, Nat. Comm. 6, 5766 (2015)]



- Replace Shannon entropy with von Neumann entropy
  (For purely classical variables both coincide)

Entropic description of Quantum Causal Structures

[RC, C. Majenz & D. Gross, Nat. Comm. 6, 5766 (2015)]

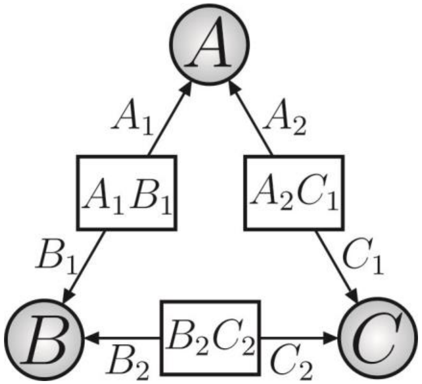


- Replace Shannon entropy with von Neumann entropy (For purely classical variables both coincide)

- von Neumann entropy respect strong subadditivity but not monotonicity (replaced by weak monotonicity)

Entropic description of Quantum Causal Structures

[RC, C. Majenz & D. Gross, Nat. Comm. 6, 5766 (2015)]

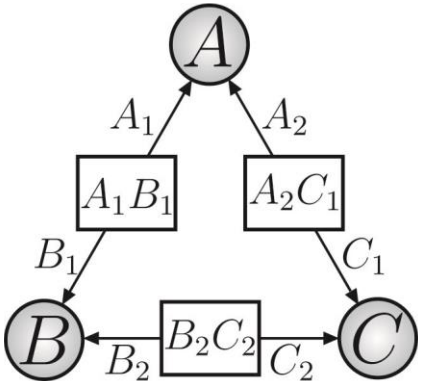


- Replace Shannon entropy with von Neumann entropy (For purely classical variables both coincide)
- von Neumann entropy respect strong subadditivity but not monotonicity (replaced by weak monotonicity)
- Measurements disturb/destroy the quantum system

$$\times \; H(A, A_1, A_2)$$

Entropic description of Quantum Causal Structures

[RC, C. Majenz & D. Gross, Nat. Comm. 6, 5766 (2015)]



- Replace Shannon entropy with von Neumann entropy (For purely classical variables both coincide)

- von Neumann entropy respect strong subadditivity but not monotonicity (replaced by weak monotonicity)

- Measurements disturb/destroy the quantum system

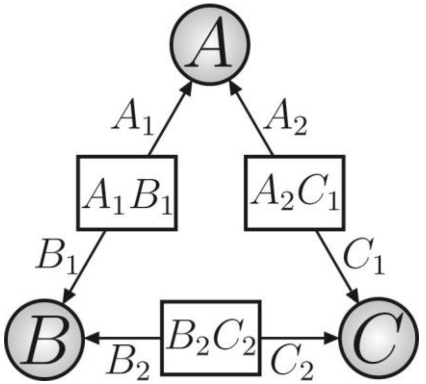
$$\times \; H(A, A_1, A_2)$$

Some CIs that are classically valid cannot be defined in the quantum case

$$\times \; I(A : B|A_1, B_1) = 0$$

Entropic description of Quantum Causal Structures

[RC, C. Majenz & D. Gross, Nat. Comm. 6, 5766 (2015)]



- Replace Shannon entropy with von Neumann entropy (For purely classical variables both coincide)

- von Neumann entropy respect strong subadditivity but not monotonicity (replaced by weak monotonicity)

- Measurements disturb/destroy the quantum system

✗ $$H(A, A_1, A_2)$$

Some CIs that are classically valid cannot be defined in the quantum case

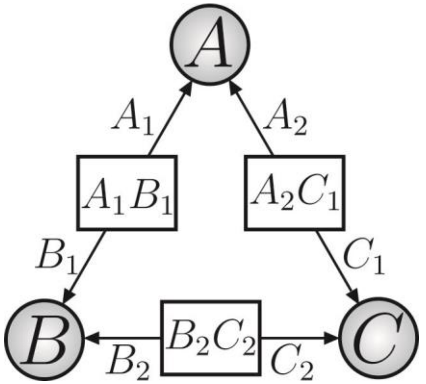
✗ $$I(A : B|A_1, B_1) = 0$$

Independecies still hold

✓ $$I(A_1B_1 : A_2C_1) = 0$$

Entropic description of Quantum Causal Structures

[RC, C. Majenz & D. Gross, Nat. Comm. 6, 5766 (2015)]



- Replace Shannon entropy with von Neumann entropy
  (For purely classical variables both coincide)

- von Neumann entropy respect strong subadditivity but
  not monotonicity (replaced by weak monotonicity)

- Measurements disturb/destroy the quantum system

$$\times \; H(A, A_1, A_2)$$

Some CIs that are classically valid cannot be defined in the quantum case

$$\times \; I(A : B | A_1, B_1) = 0$$

Independecies still hold

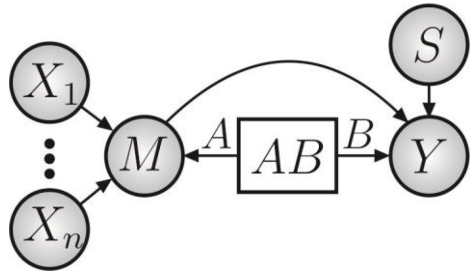
$$\checkmark \; I(A_1 B_1 : A_2 C_1) = 0$$

- We need a rule mapping the quantum states to classical variables (data processing)

$$\checkmark \; I(A : B) \leq I(A_1 A_2 : B_1 B_2)$$

Information causality

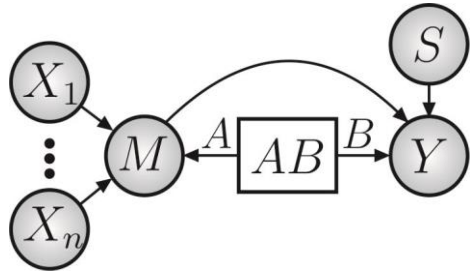
[Pawlowski et al, Nature 461, 1101 (2009)]



Task: Bob must output a guess Y_S about the s-th bit X_S of Alice

Information causality

[Pawlowski et al, Nature 461, 1101 (2009)]



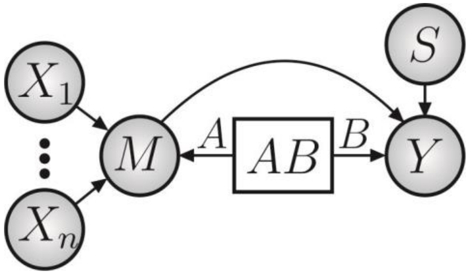
Task: Bob must output a guess Y_S about the s-th bit X_S of Alice

IC inequality ⇒

$$\sum_{i=1,\ldots,n} I(X_i : Y_i) \leq H(M)$$

Information causality

[Pawlowski et al, Nature 461, 1101 (2009)]



<u>Task</u>: Bob must output a guess Y_S about the s-th bit X_S of Alice

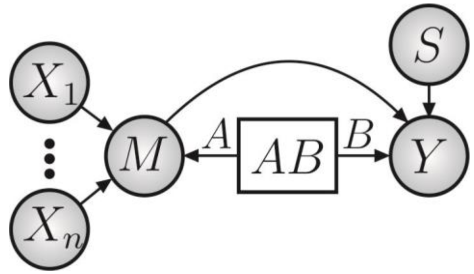
<u>IC inequality</u> ⇒

$$\sum_{i=1,\dots,n} I(X_i : Y_i) \leq H(M)$$

- Implicitly restricting to the marginal scenario: $\{X_i, Y_i\}$, $\{M\}$, with i=1,...,n

Information causality

[Pawlowski et al, Nature 461, 1101 (2009)]



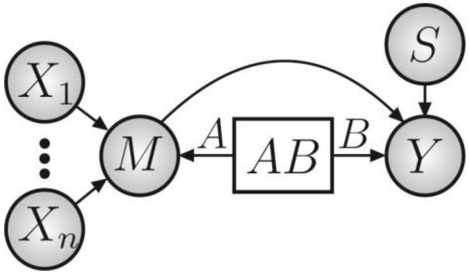
Task: Bob must output a guess Y_S about the s-th bit X_S of Alice

IC inequality ⇨ $$\sum_{i=1,\ldots,n} I(X_i : Y_i) \leq H(M)$$

- Implicitly restricting to the marginal scenario: $\{X_i, Y_i\}$, $\{M\}$, with i=1,...,n
- Most general marginal scenario: $\{X_1,\ldots,X_n,Y_s,M\}$ with s=1,...,n

Information causality

[Pawlowski et al, Nature 461, 1101 (2009)]



Task: Bob must output a guess Y_S about the s-th bit X_S of Alice

IC inequality ⇒ $$\sum_{i=1,\ldots,n} I(X_i : Y_i) \leq H(M)$$

- Implicitly restricting to the marginal scenario: $\{X_i, Y_i\}$, $\{M\}$, with i=1,...,n

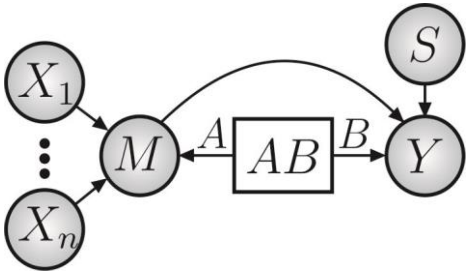
- Most general marginal scenario: $\{X_1,\ldots,X_n,Y_s,M\}$ with s=1,...,n

Tigher IC inequality

$$\sum_{i=1,\ldots,n} I(X_i : Y_i, M) + \sum_{i=2,\ldots,n} I(X_1 : X_i | Y_i, M) \leq H(M)$$

Information causality

[Pawlowski et al, Nature 461, 1101 (2009)]



Task: Bob must output a guess Y_S about the s-th bit X_S of Alice

IC inequality ⇨ $$\sum_{i=1,\ldots,n} I(X_i : Y_i) \leq H(M)$$

- Implicitly restricting to the marginal scenario: $\{X_i, Y_i\}$, $\{M\}$, with i=1,...,n
- Most general marginal scenario: $\{X_1,\ldots,X_n,Y_s,M\}$ with s=1,...,n

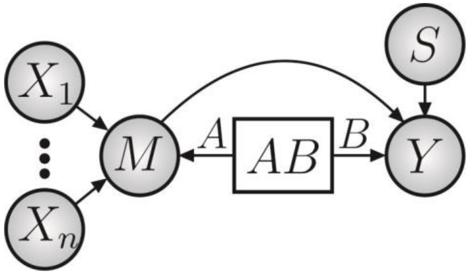
Tigher IC inequality

$$\sum_{i=1,\ldots,n} I(X_i : Y_i, M) + \sum_{i=2,\ldots,n} I(X_1 : X_i | Y_i, M) \leq H(M)$$

- The new inequality witness the non quantumness of distributions that could not be detected by the original one.

Information causality

[Pawlowski et al, Nature 461, 1101 (2009)]



Task: Bob must output a guess Y_S about the s-th bit X_S of Alice

IC inequality ⇒ $$\sum_{i=1,\ldots,n} I(X_i : Y_i) \leq H(M)$$

- Implicitly restricting to the marginal scenario: $\{X_i, Y_i\}$, $\{M\}$, with i=1,...,n
- Most general marginal scenario: $\{X_1,\ldots,X_n,Y_s,M\}$ with s=1,...,n

Tigher IC inequality

$$\sum_{i=1,\ldots,n} I(X_i : Y_i, M) + \sum_{i=2,\ldots,n} I(X_1 : X_i|Y_i, M) \leq H(M)$$

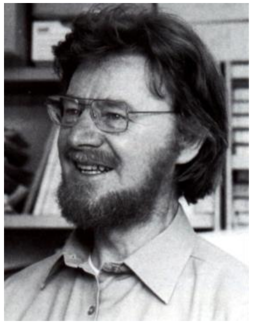
- The new inequality witness the non quantumness of distributions that are not detected by the original one.
- Can be easily generalized to the case of a quantum message (Dense Coding)

- Causal structures
- The information-theoretic approach to (classical) causal inference
- The generalization to quantum causal structures
- **Where to go from here?**

What we know…

- Entropies allow for a non-trivial, quantitative and operational discrimination between causal relationships...

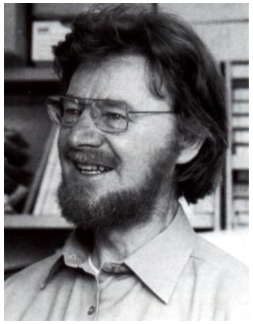
... both in classical and quantum problems



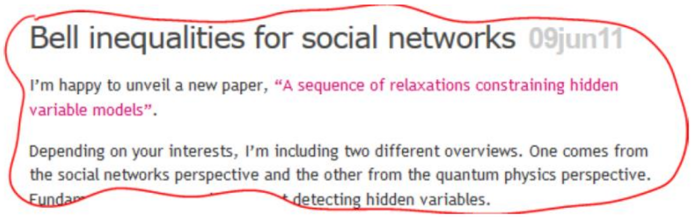
What we know…

- Entropies allow for a non-trivial, quantitative and operational discrimination between causal relationships...

... both in classical and quantum problems



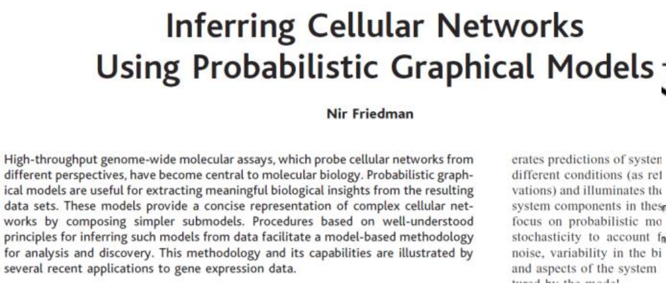
…and what we would like to know



Bell inequalities for social networks 09jun11

I'm happy to unveil a new paper, "A sequence of relaxations constraining hidden variable models".

Depending on your interests, I'm including two different overviews. One comes from the social networks perspective and the other from the quantum physics perspective. Fundam[...] detecting hidden variables.

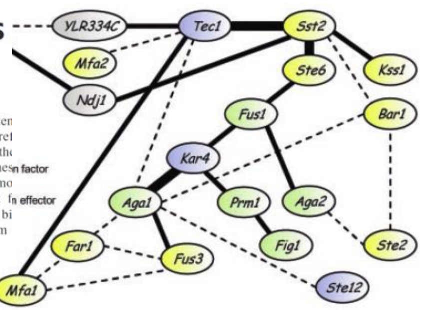


Inferring Cellular Networks Using Probabilistic Graphical Models

Nir Friedman

High-throughput genome-wide molecular assays, which probe cellular networks from different perspectives, have become central to molecular biology. Probabilistic graphical models are useful for extracting meaningful biological insights from the resulting data sets. These models provide a concise representation of complex cellular networks by composing simpler submodels. Procedures based on well-understood principles for inferring such models from data facilitate a model-based methodology for analysis and discovery. This methodology and its capabilities are illustrated by several recent applications to gene expression data.

erates predictions of systen different conditions (as re vations) and illuminates th system components in thes focus on probabilistic mo stochasticity to account f noise, variability in the bi and aspects of the system

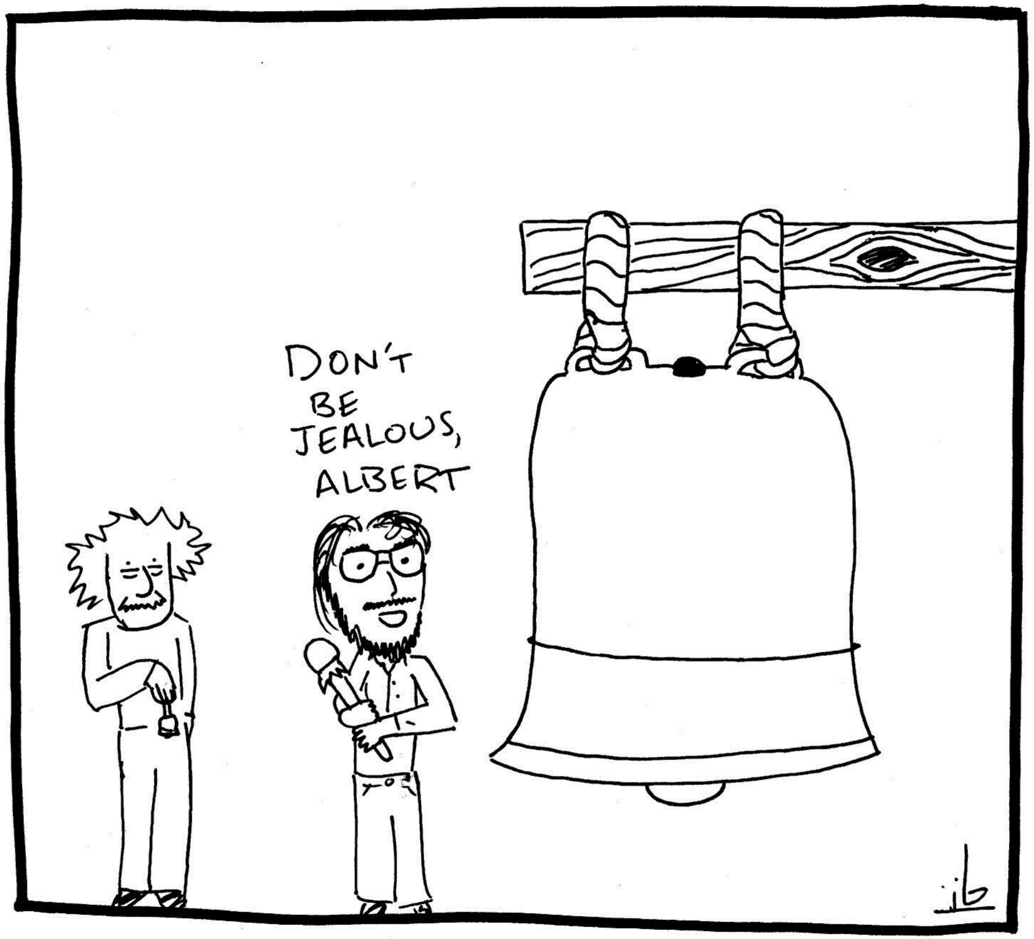


- Beyond Bell's theorem? Nonlocality in quantum networks... see **[T. Fritz, NJP 14, 103001 (2012)]**

- New information-theoretical principles? Multipartite Information Causality?

Thanks!

- **RC, C. Majenz & D. Gross, "*Information-Theoretic Implications of Quantum Causal Structures*", Nature Communications 6, 5766 (2015)**
- **RC, L. Luft, T. Maciel, D. Gross, D. Janzing, B. Schölkopf, *"Inferring latent structures via information inequalities"*, Proceedings of Uncertainty in Artificial Intelligence (2014)**



BELL'S INEQUALITY