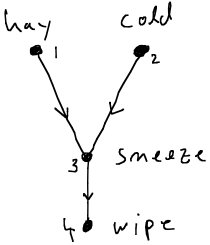
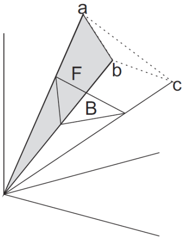
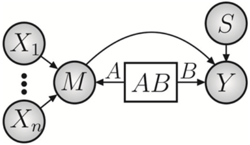
Marginal Entropies
for Causal Inference and Quantum Non-Locality







David Gross
University of Cologne

Telc
June 2015

Outline

- Background: Infering Causal Structures
- Entropic Marginals
- Quantum Causal Structures
- Relaxations of Causal Assumptions in Bell Scenarios

C. Majenz

R. Kueng

J.B. Brask

R. Chaves

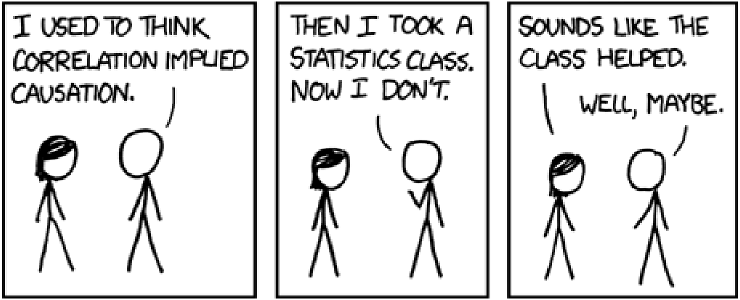
R. Chaves, L. Luft, D. Gross, New J. Phys. 16, 043001 (2014)

R. Chaves *et al.*, Proceedings of UAI 2014

R. Chaves, C. Majenz, D. Gross, Nature Communications 6, 5766 (2015)

R. Chaves, R. Kueng, J.B. Brask, D. Gross, Phys. Rev. Lett. 114, 140403 (2015)

Part 1: Infering causal structures



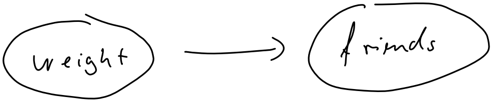
Is obesity contagious?

Empirical finding: People of similar weight more likely to be friends.

Is obesity contagious?

Empirical finding: People of similar weight more likely to be friends.

Various possible explanations:

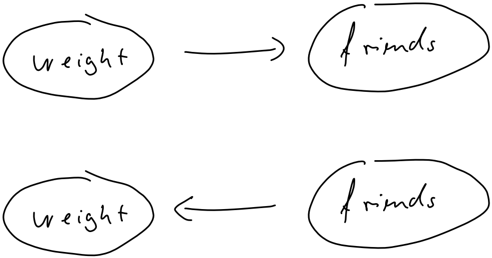


- Prefer friends with similar body constitution.

Is obesity contagious?

Empirical finding: People of similar weight more likely to be friends.

Various possible explanations:

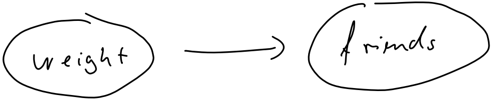


- Prefer friends with similar body constitution.
- Immitate eating habits of friends.
- "Obesity is contagious"

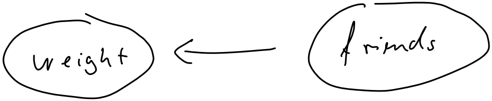
Is obesity contagious?

Empirical finding: People of similar weight more likely to be friends.

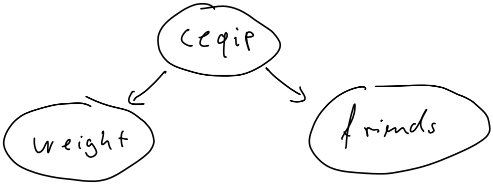
Various possible explanations:



- Prefer friends with similar body constitution.

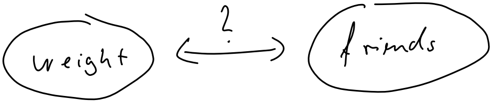


- Immitate eating habits of friends.
- "Obesity is contagious"



- Unobserved common cause.

Interventions



Causal relationships can be probed by *interventions*:

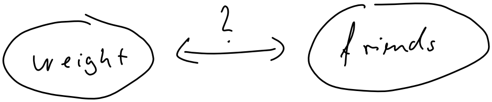
Compare

$$\Pr[\text{friends} \mid \text{same weight}]$$

$$\Pr[\text{friends} \mid \textit{do}(\text{same weight})]$$

$$\Pr[\textit{do}(\text{friends}) \mid \text{weight}].$$

Passive Causal Inference?



However:

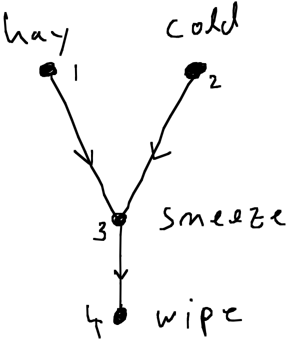
- Interventions often impractical / unethical

Natural Question:

Can one obtain information about causal relations from empirical observations?

Causal structures

To address problem, formalize notions:



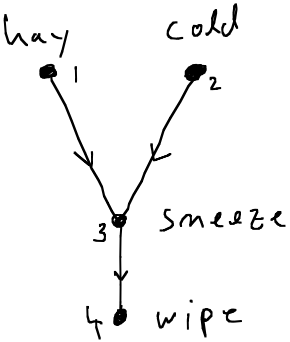
- For n variables $X_1, \ldots, X_n$,
- a *causal structure* or *Bayesian network* is directed acyclic graph,
- with ith variable deterministic function

$$X_i = f_i(\mathrm{pa}_i, u_i)$$

of its parents pa_i and "local randomness" u_i

Causal structures

To address problem, formalize notions:



- For n variables $X_1, \ldots, X_n$,
- a *causal structure* or *Bayesian network* is directed acyclic graph,
- with ith variable deterministic function

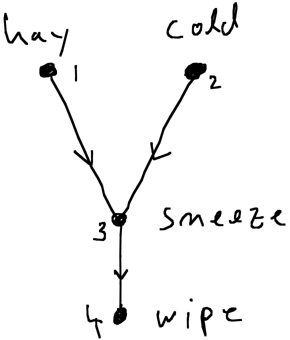
$$X_i = f_i(\mathrm{pa}_i, u_i)$$

of its parents pa_i and "local randomness" u_i

Chain rule of probability $\Rightarrow$ joint p.d.f. is

$$p(x_1, \ldots, x_i) = \Pi_{i=1}^{n} P(x_i | \mathrm{pa}_i, u_i).$$

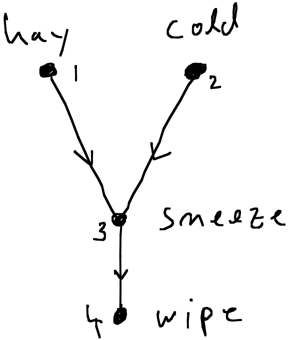
Local Markov Condition



A causal structure gives rise to distributions where,

- X_i is independent of its non-descendants, given its parents.

Local Markov Condition



A causal structure gives rise to distributions where,

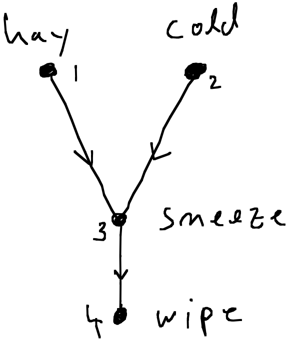
- X_i is independent of its non-descendants, given its parents.
- I.e. dependency on "past" only through parents.

$\Rightarrow$ criterion for *rejecting* causal structures. E.g.:

- if "wiping" found to be *not* independent of "cold" given "sneezing"

  $\Rightarrow$ above structure can't explain data.

Local Markov Condition



A causal structure gives rise to distributions where,

- X_i is independent of its non-descendants, given its parents.
- I.e. dependency on "past" only through parents.

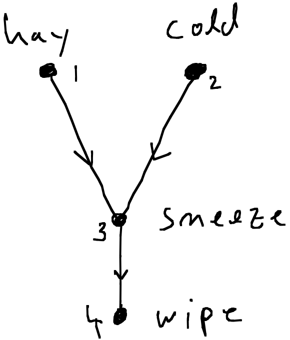
⇒ criterion for *rejecting* causal structures. E.g.:

- if "wiping" found to be *not* independent of "cold" given "sneezing"
  ⇒ above structure can't explain data.

Empirical independences hold clues about causation.

Local Markov Condition

Causal structure *does* imply testable conditions

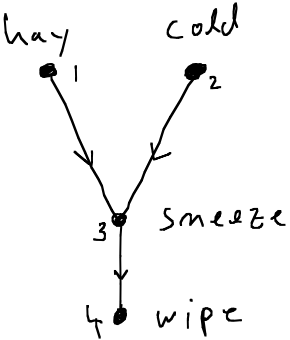


Ex.:

- "wiping" independent of "cold" *conditioned* on "sneazing".

Local Markov Condition

Causal structure *does* imply testable conditions



Ex.:

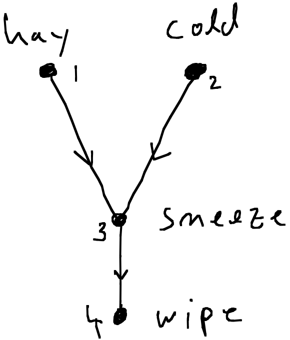
- "wiping" independent of "cold" *conditioned* on "sneazing".

More generally:

- X_i is independent of its non-descendants, given its parents.
- "Local Markov Condition".

Local Markov Condition

Causal structure *does* imply testable conditions



Ex.:

- "wiping" independent of "cold" *conditioned* on "sneazing".

More generally:

- X_i is independent of its non-descendants, given its parents.
- "Local Markov Condition".

Result:

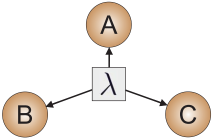
(1) *All* corollaries of causal structures follow from Local Markov Conditions.
(2) Recoverable aspects of causality graph well-understood.

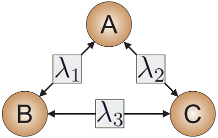
[Pearl, 2000]

Hidden variables (confounders / latent variables)

. . . however, analysis breaks down if only subset of variables accessible.

Ex.: “common ancestor” problem:

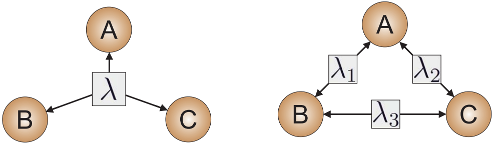




Hidden variables (confounders / latent variables)

. . . however, analysis breaks down if only subset of variables accessible.

Ex.: "common ancestor" problem:

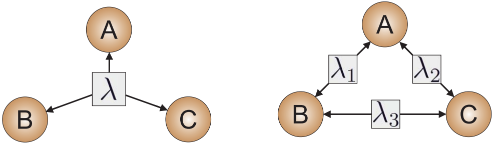


- Pair-wise structure implies no independences between A, B, C,
- but is not compatible, e.g. with 3 perfectly correlated coins.

Hidden variables (confounders / latent variables)

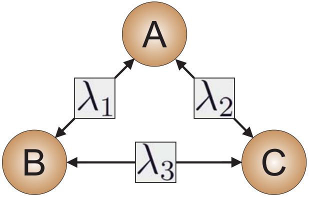
. . . however, analysis breaks down if only subset of variables accessible.

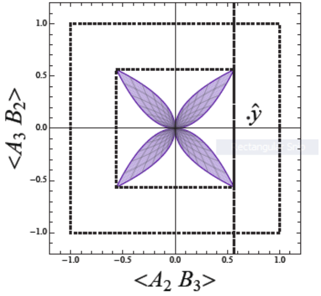
Ex.: "common ancestor" problem:



- Pair-wise structure implies no independences between A, B, C,
- but is not compatible, e.g. with 3 perfectly correlated coins.
- (amazingly, this example not yet fully characterized).

Algebraic Statistics



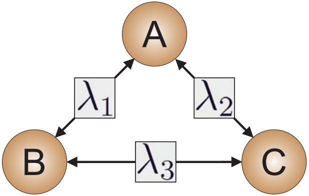


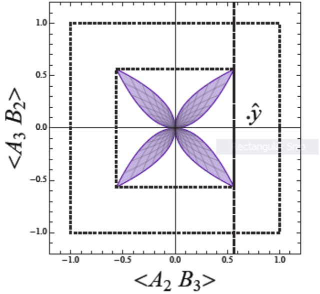
- Independences = algebraic constraints

$$p(x, y) = p(x)p(y)$$
$$\Leftrightarrow \quad \mathrm{rank}(p(x, y)) = 1$$

- Rank variety + Positivity
= real algebraic geometry

Algebraic Statistics



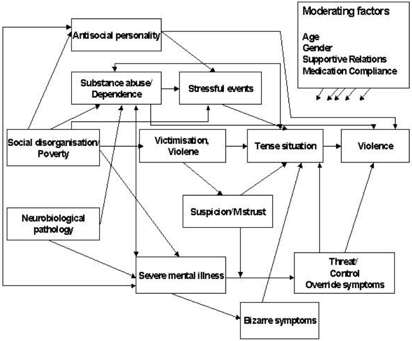


- Independences = algebraic constraints

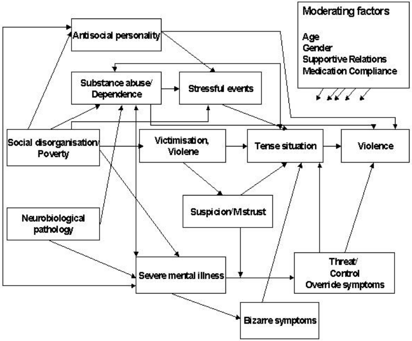
$$\begin{aligned} & p(x,y) = p(x)p(y) \\ \Leftrightarrow \quad & \text{rank}(p(x,y)) = 1 \end{aligned}$$

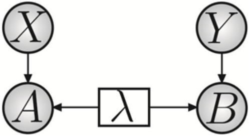
- Rank variety + Positivity = real algebraic geometry
- Nasty in theory and pratice...
- ...so new ideas needed.

Diverse Applications. . .

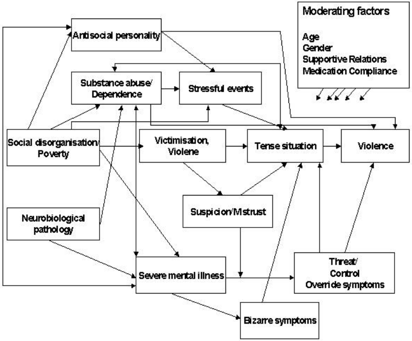


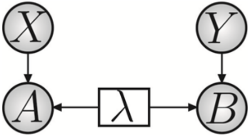
Diverse Applications. . .





Diverse Applications. . .





Bell inequalities for social networks 09jun11

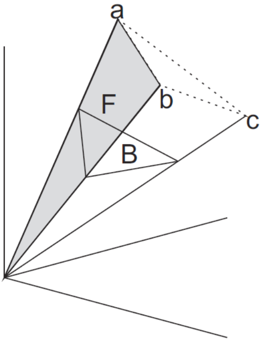
I'm happy to unveil a new paper, "A sequence of relaxations constraining hidden variable models".

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... ...t detecting hidden variables.

Part 2: Entropic Marginals

1. Entropy cone

Step 1/3: The unconstrained, global object.

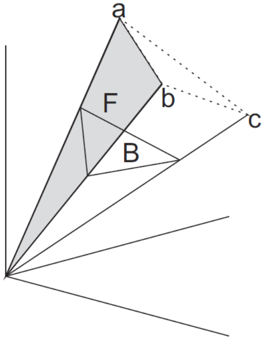


- Associate with $S \subset \{1, \ldots, n\}$ the joint entropy $S(X_S)$
- $\Rightarrow$ an *entropy vector* $v \in \mathbb{R}^{2^n}$, indexed by subsets

  Ex.: $\left(H(\emptyset), H(A), H(B), H(A,B)\right)$

1. Entropy cone

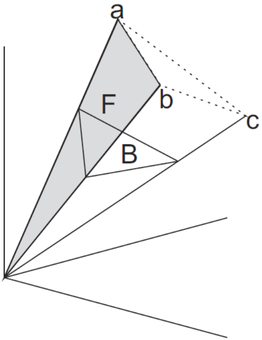
Step 1/3: The unconstrained, global object.



- Associate with $S \subset \{1, \ldots, n\}$ the joint entropy $S(X_S)$
- $\Rightarrow$ an *entropy vector* $v \in \mathbb{R}^{2^n}$, indexed by subsets
  Ex.: $(H(\emptyset), H(A), H(B), H(A,B))$
- Easy proof: Set of all such entropy vectors forms convex cone Γ_n^* (up to closure).

1. Entropy cone

Step 1/3: The unconstrained, global object.



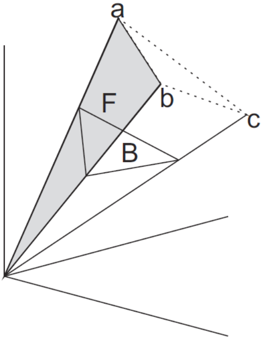
- Associate with $S \subset \{1, \ldots, n\}$ the joint entropy $S(X_S)$
- $\Rightarrow$ an *entropy vector* $v \in \mathbb{R}^{2^n}$, indexed by subsets
  Ex.: $(H(\emptyset), H(A), H(B), H(A, B))$
- Easy proof: Set of all such entropy vectors forms convex cone Γ_n^* (up to closure).

- Structure not fully understood, but...
- ...contained in *Shannon cone* cone Γ_n, defined by strong subadditivity and monotonicity.

$$H(A, B) \leq H(A, B, C), \quad H(A, B) \leq H(A)+H(B), \quad I(B : C|A) \geq 0.$$

1. Entropy cone

Step 1/3: The unconstrained, global object.



- Associate with $S \subset \{1, \dots, n\}$ the joint entropy $S(X_S)$
- $\Rightarrow$ an *entropy vector* $v \in \mathbb{R}^{2^n}$, indexed by subsets
  Ex.: $(H(\emptyset), H(A), H(B), H(A, B))$
- Easy proof: Set of all such entropy vectors forms convex cone Γ_n^* (up to closure).

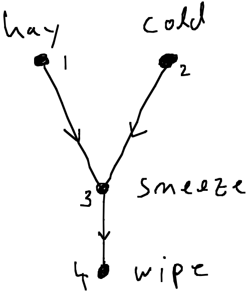
- Structure not fully understood, but...
- ...contained in *Shannon cone* cone Γ_n, defined by strong subadditivity and monotonicity.

$$H(A, B) \leq H(A, B, C), \quad H(A, B) \leq H(A)+H(B), \quad I(B : C|A) \geq 0.$$

- We will mostly work with Shannon relaxation.

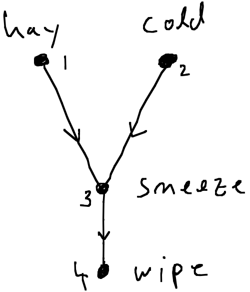
2. Causal constraints

Step 2/3: Now choose candidate structure and add causal constraints.



2. Causal constraints

Step 2/3: Now choose candidate structure and add causal constraints.



- That's easy:
  Conditional independences measured by mutual information:

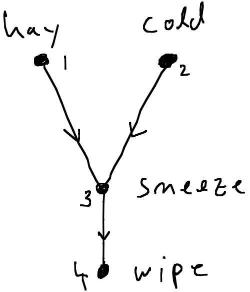
  $$I(\{\text{wipe}\} : \{\text{hay}, \text{cold}\} | \{\text{sneeze}\}) = 0.$$

- Can even relax:

  $$I(\{\text{wipe}\} : \{\text{hay}, \text{cold}\} | \{\text{sneeze}\}) \leq \epsilon.$$

2. Causal constraints

Step 2/3: Now choose candidate structure and add causal constraints.



- That's easy:
  Conditional independences measured by mutual information:

$$I(\{\mathrm{wipe}\} : \{\mathrm{hay}, \mathrm{cold}\} | \{\mathrm{sneeze}\}) = 0.$$

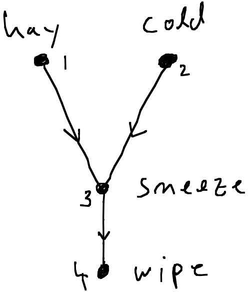
- Can even relax:

$$I(\{\mathrm{wipe}\} : \{\mathrm{hay}, \mathrm{cold}\} | \{\mathrm{sneeze}\}) \leq \epsilon.$$

- $\Rightarrow$ cone C of constraints.

2. Causal constraints

Step 2/3: Now choose candidate structure and add causal constraints.



- That's easy:
  Conditional independences measured by mutual information:

  $$I(\{\mathrm{wipe}\} : \{\mathrm{hay}, \mathrm{cold}\} | \{\mathrm{sneeze}\}) = 0.$$

- Can even relax:

  $$I(\{\mathrm{wipe}\} : \{\mathrm{hay}, \mathrm{cold}\} | \{\mathrm{sneeze}\}) \leq \epsilon.$$

- $\Rightarrow$ cone C of constraints.

$\Rightarrow$ new global cone $\Gamma_n \cap C$ of entropies subject to causal structure.

3. Marginalize

Step 3/3: Marginalize.

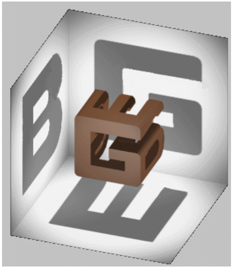
- Set $\mathcal{M} \subset 2^{\{1,\ldots,n\}}$ of jointly observable r.v.'s is *marginal scenario*.
- Classically: r.v.'s either observable or not
  QM: Some r.v.'s not jointly measureable.

3. Marginalize

Step 3/3: Marginalize.

- Set $\mathcal{M} \subset 2^{\{1,\dots,n\}}$ of jointly observable r.v.'s is *marginal scenario*.
- Classically: r.v.'s either observable or not
  QM: Some r.v.'s not jointly measureable.

Marginalize to $\mathcal{M}$:



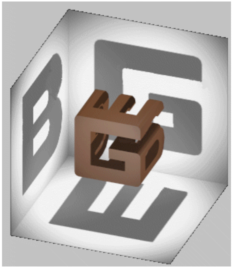
- Geometrically trivial:
  just restrict $\Gamma_n \cap C$ to observable coordinates.
- Algorithmically costly: $\Gamma_n \cap C$ represented in terms of inequalities (use, e.g. Fourier-Motzkin elimination)

3. Marginalize

Step 3/3: Marginalize.

- Set $\mathcal{M} \subset 2^{\{1,\ldots,n\}}$ of jointly observable r.v.'s is *marginal scenario*.
- Classically: r.v.'s either observable or not
  QM: Some r.v.'s not jointly measureable.

Marginalize to $\mathcal{M}$:

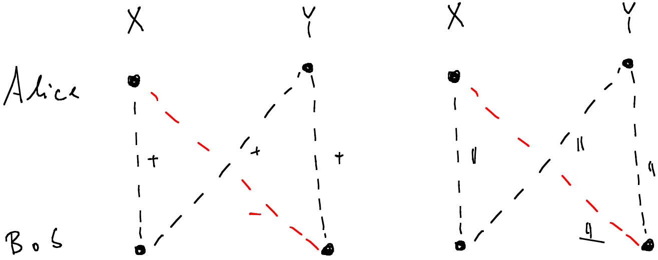


- Geometrically trivial:
  just restrict $\Gamma_n \cap C$ to observable coordinates.
- Algorithmically costly: $\Gamma_n \cap C$ represented in terms of inequalities (use, e.g. Fourier-Motzkin elimination)

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

1. Relation Entropy & Binary Bell Ineqs

Revisit "entropic CHSH" [Braunstein & Caves '88 (!)]

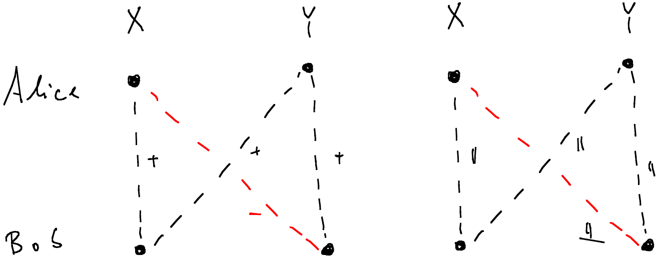


$$\begin{aligned} & \langle X_A X_B \rangle + \langle Y_A X_B \rangle + \langle Y_A Y_B \rangle - \langle X_A Y_B \rangle \leq \langle X_A \rangle + \langle X_B \rangle \\ - \quad & H(X_A X_B) - H(Y_A X_B) - H(Y_A Y_B) + H(X_A Y_B) \geq -H(X_A) - H(X_B) \end{aligned}$$

- Measures frustration in *degree* of correlation, rather than *sign*.
- Resembles "sign-reversed" CHSH. No coincidence...

1. Relation Entropy & Binary Bell Ineqs

Revisit "entropic CHSH" [Braunstein & Caves '88 (!)]

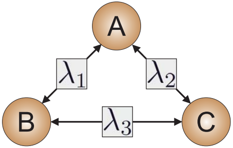


$$\langle X_A X_B \rangle + \langle Y_A X_B \rangle + \langle Y_A Y_B \rangle - \langle X_A Y_B \rangle \leq \langle X_A \rangle + \langle X_B \rangle$$

$$- \quad H(X_A X_B) - H(Y_A X_B) - H(Y_A Y_B) + H(X_A Y_B) \geq -H(X_A) - H(X_B)$$

- Measures frustration in *degree* of correlation, rather than *sign*.
- Resembles "sign-reversed" CHSH. No coincidence...
- Result: Negative of any multipartite entropic ineq also valid for probabilities. [NJP '13]
- Often, converse true $\Rightarrow$ Source of entropic Bell ineqs [NJP '13]

2. Common Ancestors & Strength of Causal Influence

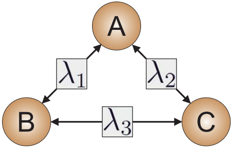


- Entropic constraints given by (perms of)

$$\mathcal{B} = I(A:B) + I(A:C) - H(A) \leq 0.$$

- Ex.: Perfectly correlated coins: $\mathcal{B} = 1$.

2. Common Ancestors & Strength of Causal Influence

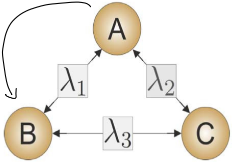


- Entropic constraints given by (perms of)

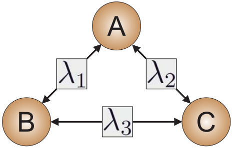
$$\mathcal{B} = I(A:B) + I(A:C) - H(A) \leq 0.$$

- Ex.: Perfectly correlated coins: $\mathcal{B} = 1$.

- Violation $\mathcal{B}$ interpretable as *causal strength* of direct influence $A \rightarrow B$ required to explain data [UAI '14]



2. Common Ancestors & Strength of Causal Influence

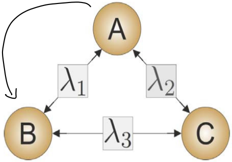


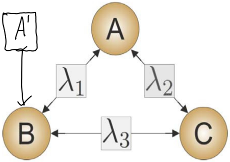
- Entropic constraints given by (perms of)

$$\mathcal{B} = I(A:B) + I(A:C) - H(A) \leq 0.$$

- Ex.: Perfectly correlated coins: $\mathcal{B} = 1$.

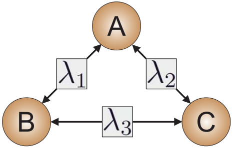
- Violation $\mathcal{B}$ interpretable as *causal strength* of direct influence $A \to B$ required to explain data [UAI '14]





- Def. causal strength $\mathcal{C}_{A \to B}$ as relative entropy distance incurred by cutting link.

2. Common Ancestors & Strength of Causal Influence

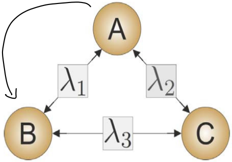


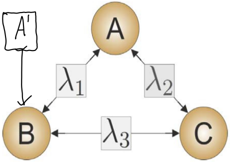
- Entropic constraints given by (perms of)

$$\mathcal{B} = I(A : B) + I(A : C) - H(A) \leq 0.$$

- Ex.: Perfectly correlated coins: $\mathcal{B} = 1$.

- Violation $\mathcal{B}$ interpretable as *causal strength* of direct influence $A \to B$ required to explain data [UAI '14]



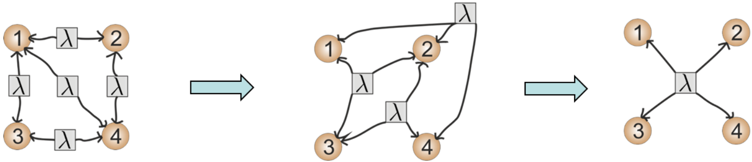


- Def. causal strength $\mathcal{C}_{A \to B}$ as relative entropy distance incurred by cutting link.
- Then $\mathcal{C}_{A \to B} \geq \mathcal{B}$. [UAI '14]

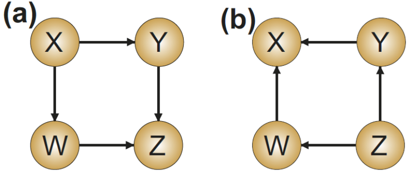
3. Many more. . .

Can treat. . .

- Scenarios of n observables with independent common ancestors influencing at most M each



- Direction of causation from pairwise marginals

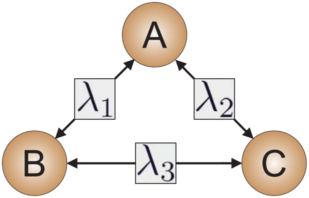


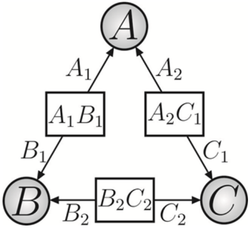
. . . and more.

[UAI '14]

Part 3: Quantum Causal Stuctures

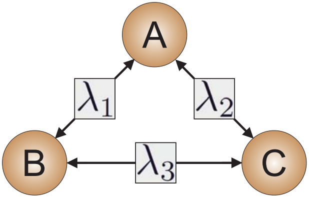
Quantum Causal Structures 1

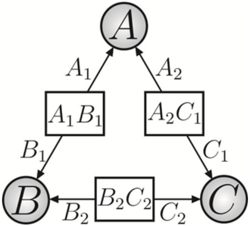




- With minor modifications, causal diagrams make sense for quantum systems.

Quantum Causal Structures 1





- With minor modifications, causal diagrams make sense for quantum systems.
- Nodes are states. Labels desginate systems.
- If node has incoming edges, state results from CP map applied to incoming systems.
- Sample diagram says

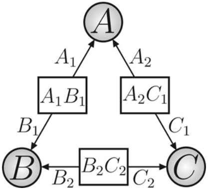
$$\rho_{ABC} = \left[\Phi_{A_1A_2\to A}\otimes\Phi_{B_1B_2\to B}\otimes\Phi_{C_1C_2\to C}\right]\left(\rho_{A_1B_2}\otimes\rho_{A_2C_2}\otimes\rho_{B_2C2_C}\right).$$

Quantum Causal Structures 2

How to build entropic constraints for quantum causal structures:

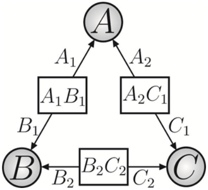
1. Use von Neumann entropy

   $\Rightarrow$ drop monotonicity ineq. $H(A, B) \geq H(A)$



Quantum Causal Structures 2

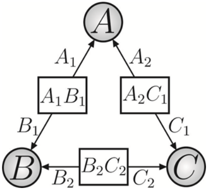
How to build entropic constraints for quantum causal structures:



1. Use von Neumann entropy
   $\Rightarrow$ drop monotonicity ineq. $H(A, B) \geq H(A)$
2. QM does not assign joint state to input & output of operation! No "$H(A_1, A)$"! Consider only *coexisting variables*!

Quantum Causal Structures 2

How to build entropic constraints for quantum causal structures:

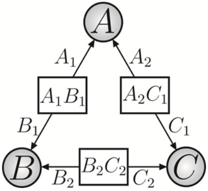


1. Use von Neumann entropy
   $\Rightarrow$ drop monotonicity ineq. $H(A,B) \geq H(A)$
2. QM does not assign joint state to input & output of operation! No "$H(A_1, A)$"! Consider only *coexisting variables*!
3. Use data processing inequality to relate non-coexisting variables. Ex.:

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

Quantum Causal Structures 2

How to build entropic constraints for quantum causal structures:

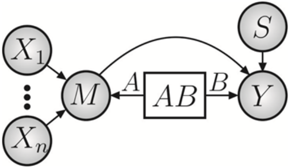


1. Use von Neumann entropy
   $\Rightarrow$ drop monotonicity ineq. $H(A, B) \geq H(A)$
2. QM does not assign joint state to input & output of operation! No "$H(A_1, A)$"! Consider only *coexisting variables*!
3. Use data processing inequality to relate non-coexisting variables. Ex.:

$$I(A : B) \leq I(A_1A_2 : B_1B_2).$$

. . . gives rich theory [Nat. Comm. '14].

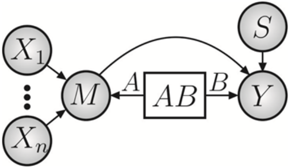
Quantum Causal Structures Ex.: Information Causality



Recall inf. caus. game: [Pawlowski *et al.*, Nature '09]

- Alices receives bits $X_1, \ldots, X_n$, sends message M to Bob
- Bob recives M and challenge $S \rightarrow$ outputs guess Y for X_S
- Aided by joint quantum state ρ_{AB}

Quantum Causal Structures Ex.: Information Causality



Recall inf. caus. game: [Pawlowski *et al.*, Nature '09]

- Alices receives bits $X_1, \ldots, X_n$, sends message M to Bob
- Bob recives M and challenge $S \to$ outputs guess Y for X_S
- Aided by joint quantum state ρ_{AB}

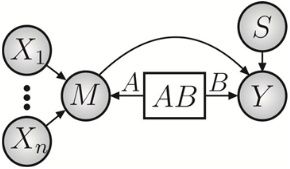
Original inequality:

$$\sum_s I(X_s : Y | S = s) \leq H(M)$$

Strengthening using systematic "quantum causal structures" prot.:

$$I(X_1 : Y_1, M) + I(X_2 : Y_2, M) + I(X_1 : X_2 | Y_2, M) \leq H(M) + I(X_1 : X_2).$$

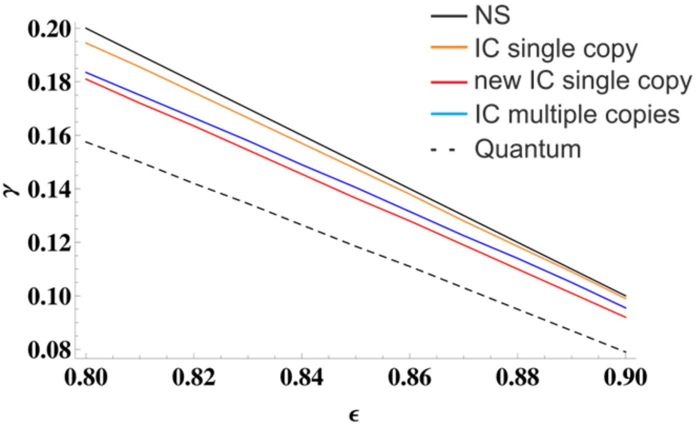
Quantum Causal Structures Ex.: Information Causality



Two consequences of strenghtened ineq.:

1. Violation measures "direct causal influence" $\mathcal{C}_{X \to Y}$

Quantum Causal Structures Ex.: Information Causality



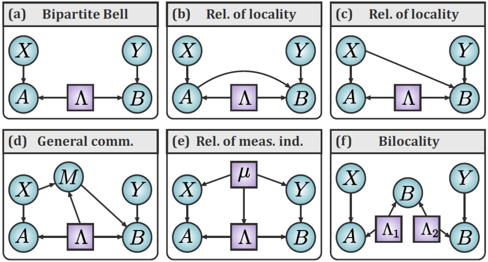
$$p(a,b|x,y) = \gamma P_{PR} + \epsilon P_{det} + (1-\gamma-\epsilon) P_{white}$$

Two consequences of strenghtened ineq.:

1. Violation measures "direct causal influence" $\mathcal{C}_{X\to Y}$
2. Detects more post-quantum correlations:

Part 4: Relaxations of causal assumptions in Bell scenarios

Relaxations of causal assumptions in Bell scenarios



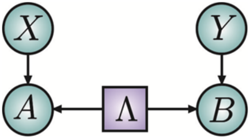
In this part:

- Do *not* work with entropies.

But show how...

- ...graphical notation of causality make it easy to reason about relaxations of causal assumptions.
- ...the idea of quantifying "causal influence" is fruitful for Bell scenarios.

Relaxations of causal assumptions in Bell scenarios



Constraints encoded by Bell causal structure have names:

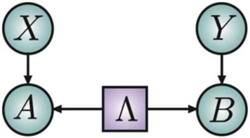
- Locality

$$p(b|x, y, \lambda) = p(b|y, \lambda).$$

- Measurement independence

$$p(x, y, \lambda) = p(x)p(y)p(\lambda).$$

Relaxations of causal assumptions in Bell scenarios



Constraints encoded by Bell causal structure have names:

- Locality

$$p(b|x, y, \lambda) = p(b|y, \lambda).$$

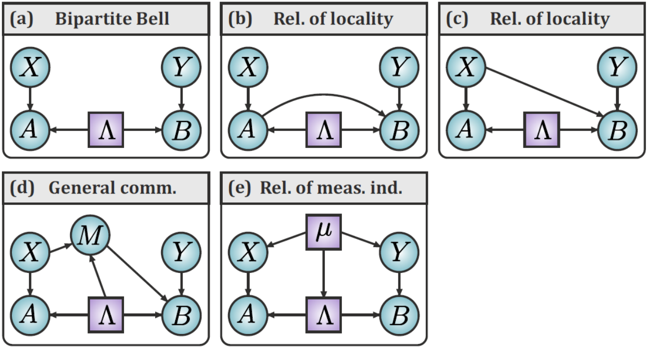
- Measurement independence

$$p(x, y, \lambda) = p(x)p(y)p(\lambda).$$

How much do we need to relax the causal assumptions entering in Bell's theorem to explain "non-local correlations" classically?

Relaxations

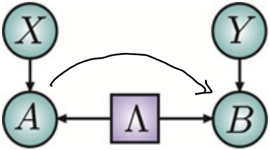
- Ingredient 1: More general *causal structures*

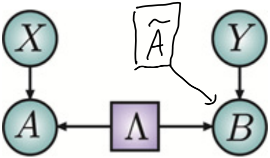


- Ingredient 2: Quantitative measures of *causal strength*

Relaxations

- Ingredient 1: More general *causal structures*
- Ingredient 2: Quantitative measures of *causal strength*





Meas. $\mathcal{C}_{A\to B}$ used here: Maximal change in total variational distance incurred by manually changing A:

$$\mathcal{C}_{A\to B} = \sup_{a,a'} \sum_{\lambda} p(\lambda) \Big| p(b|do(a),\lambda) - p(b|do(a'),\lambda) \Big|$$

Results

Main Result

Quantitative minimum of relaxation necessary to classically explain observed data can often be cast as a linear program. Moreover, closed-form results can often be obtained using duality theory.

Results

Main Result

Quantitative minimum of relaxation necessary to classically explain observed data can often be cast as a linear program. Moreover, closed-form results can often be obtained using duality theory.

- Causal interpretation of numerical CHSH violation:

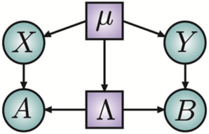
$$\min \mathcal{C}_{A\to B} = \min \mathcal{C}_{X\to B} = \max\{0, CHSH\}$$

Results

Main Result

Quantitative minimum of relaxation necessary to classically explain observed data can often be cast as a linear program. Moreover, closed-form results can often be obtained using duality theory.

- Quantitative bound on measurement dependence



$$\min \mathcal{M} = \max\{0, I_d/4\},$$

where

$$\mathcal{M} = \|p(\lambda, x, y) - p(\lambda)p(x, y)\|_{TV}$$

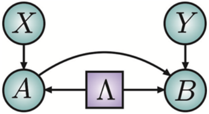
and I_d violation of CGLMP-inequality.

Results

Main Result

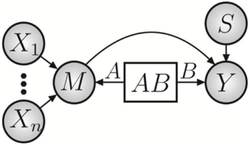
Quantitative minimum of relaxation necessary to classically explain observed data can often be cast as a linear program. Moreover, closed-form results can often be obtained using duality theory.

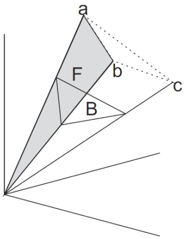
- Quantum violations even for classical models that allow for communication of measurement outcomes!



Summary

- Causal structures and Bell nonlocality go well together
- Independences linear constraints on entropies...
- ...fits the theory well.





Thank you
for your
attention

David Gross
University of Cologne

Telc
June 2015