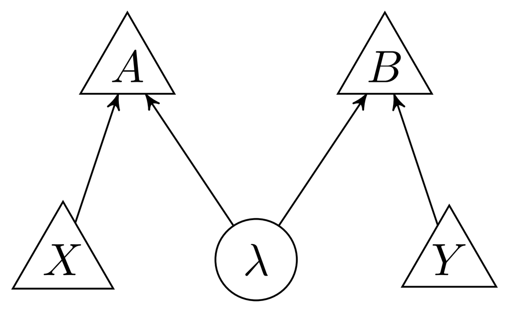
Which Causal Scenarios Might Support "Non-Local" Correlations

Shashaank Khanna, Matthew Pusey
Mathematical Physics Group, Department of Mathematics, University of York, YO10 5DD, UK
shashaank.khanna@york.ac.uk, matthew.pusey@york.ac.uk

What's an "interesting" scenario

A causal relationship can be described using the formalism of Generalised Bayesian Networks. This framework allows the depiction of cause and effect relations (causal scenarios) effectively using generalised directed acyclic graphs (GDAGs). A GDAG is considered **"not interesting"** [1] if the classical correlations existing in it are just constrained by the observable conditional independencies in it. **This implies that no non-classical correlations can ever be achieved in such "non-interesting" scenarios.** A standard "interesting" causal scenario which possess non-classical correlations is the Bell scenario shown below.



The problem of characterizing **"interesting" causal scenarios or GDAs of 7 nodes is an open one** and it is this problem we consider here. Characterizing "interesting" scenarios is not only important from the perspective of quantum foundations but also in developing device-independent quantum information protocols.

GDAGs and Markovianity

A generalized directed acyclic graph G (GDAG) is a pair (V, E), where V is a set of nodes and $E \subseteq V \times V$ is a set of directed edges, and which has no directed cycles. In our work, generalized directed acyclic graphs will represent causal structures. An edge $X \to Y$ shows a possibility of a direct causal influence of X on Y. The nodes that directly influence a given node are its parents. The set of parents of a node Y are denoted as $PA(Y)$. Observed nodes which depict classical random variables are represented as triangles and unobserved nodes which represent latent variables or other general resources like the quantum state are drawn as circles.

Markov Condition

A probability distribution P is Markov relative to a GDAG (generalized directed acyclic graph) G, if P satisfies:

$$P(X^1......X^n) = \prod_i P(X^i|PA(X^i)) \tag{1}$$

$\mathcal{C}$ vs $\mathcal{I}$ gap

A causal scenario has 2 basic components-: 1) Observed Nodes (which we observe and can manipulate) and 2) Unobserved Nodes (which we posit to explain the observed correlations but cannot manipulate).

$\mathcal{I}$-: is the set of all probability distributions over observed nodes that follow the observed conditional independencies (and dependencies or correlations).

$\mathcal{C}$-: is the set of all probability distributions over observed nodes which follow our hypothesis about the existence of certain "hidden or non-observed nodes" giving rise to the observed correlations. In the language of Bayesian Networks, $\mathcal{C}$ is just the set of all marginals over observed nodes that follow from the Markov conditions (Equation 1).

For eample for the Bell Scenario we have,

$$\mathcal{C} = \{P(A,B,X,Y) : \sum_\lambda P(A|X,\lambda)P(B|Y,\lambda)P(X)P(Y)P(\lambda)\} \tag{2}$$

$$\mathcal{I} = \{P(A,B,X,Y) : X \perp_p BY, Y \perp_p AX\} \tag{3}$$

where, where $X \perp_p BY$ means X is conditionally independent of BY (and similarly for the other condition).

If and only if $\mathcal{C} \neq \mathcal{I}$, it might be possible that the causal scenario possess "Non-Classical" correlations. Here we find which causal scenarios of 7 nodes in total might possess "Non-Classical" correlations. In the process, we needed to resort to heavy computation and as a result, **also found certain simple techniques to accelerate Fourier-Motzkin elimination and to remove redundant inequalities more efficiently.**

Selected References

[1] Joe Henson *et al* 2014 *New J. Phys.* **16** 113043
[2] RJ Evans *Scandinavian Journal of Statistics* **43** (3), 625-648

Correct E-separation theorem for "interestingness"

$(X \perp Y|Z)_{del_W}$ means X is d-separated of Y given Z after deletion of W in a GDAG where d-separation is a particular graphical criterion in the theory of Bayesian -Networks.

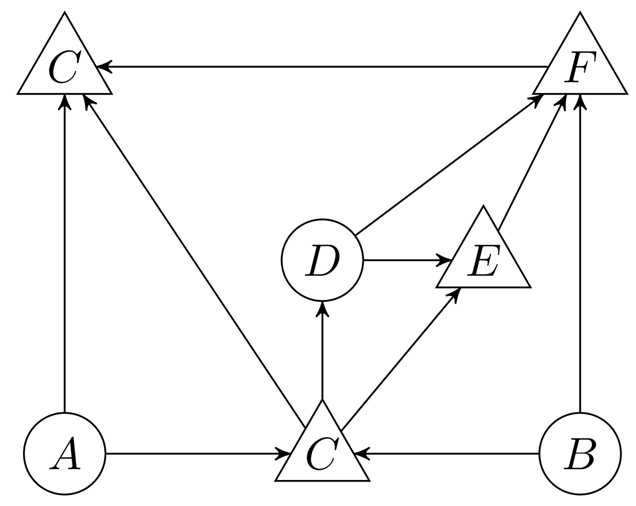
If $(X \perp Y|Z)_{del_W}$ (which implies $(x_i \perp y_j|Z)_{del_W}$ $\forall$ $x_i \subseteq X, y_j \subseteq Y$) holds in the GDAG and no member of Z is a descendant from any nodes in W then for any fixed value $W = w$ the conditional probability [2]

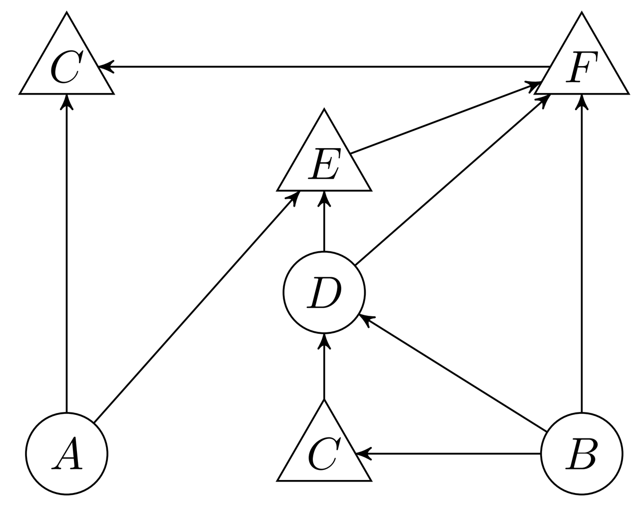
$$P(X,Y,W|Z) = P^*(X,Y,W|Z) \tag{4}$$

for any fixed value $W = w$ and where,

$$P^*(X,Y,W=w|Z) = P^*(W=w|X,Y,Z)P^*(X|Z)P^*(Y|Z) \tag{5}$$

Theorem: Let X, Y, Z, W be disjoint sets of observable nodes in a GDAG and let OBS be the set of all observable nodes in the GDAG. If $(X \perp Y|Z)_{del_W}$ holds true for the GDAG and no member of Z is a descendant from any nodes in W then Equations 4 and 5 do not follow from the conditional independences in $\mathcal{I}$ if and only if $\mathcal{I}$ excludes all relations of the form $x_i \perp y_j|S$ $\forall$ $x_i \subseteq X, y_j \subseteq Y, S \subseteq OBS - X - Y$.





For example, in the first GDAG above for $X = G$, $Y = E$, $Z = C$, $W = F$, we have $(G \perp E|C)_{del_F}$ and no member in $Z = C$ is a descendant from any member in $W = F$. Also, $(G \perp E|CF)$ does not hold true in $\mathcal{I}$. Hence our theorem concludes it as interesting. In the 2nd the GDAG above for $X = G$, $Y = C$, $Z = \phi$, $W = F$, we have $(G \perp C|\phi)_{del_{EF}}$ and no member in $Z = \phi$ is a descendant from any member in $W = EF$. Also, $(G \perp C|EF)$ does not hold in $\mathcal{I}$ for this GDAG. Therefore our theorem concludes it is interesting as well.

Our theorem is quite significant because it **reduces the unclassified GDAGs of 7 nodes to just 20**. We later see that these 20 GDAGs are in fact interesting.

Accelerating Fourier-Motzkin Elimination

We try to find Shannon-type entropic inequalities for the remaining 20 GDAGs to show that they are "interesting". But doing this computationally was infeasible because of the large number of redundancies generated while Fourier-Motzkin elimination of unobserved variables. We accelerate the Fourier-Motzkin elimination process to generate the inequalities computationally in the following 2 ways-:

1) At each step eliminate that variable whose elimination produces the least number of redundant inequalities.
2) Or at each step eliminate all variables in parallel over different cores but **proceed with the elimination of the variable which produces the least number of non-redundant inequalities.**

Results - Possible Interesting Scenarios of 7

Some out of the 20 GDAGs of 7 nodes for which we could find Shannon-type entropic inequalities and which turned out to be **"interesting"** are shown below. These **might possess "Non-Local" correlations**.