A Machine Learning Approach to Causal Discovery in Quantum Mechanics

Omid Charrakh

Munich Center for Mathematical Philosophy

LMU Munich

July 2021

Motivation (1)

- Causal discovery algorithms (CDAs) are computational techniques for inferring causal facts from statistical data.

- CDAs check the pattern of dependences and independences in statistical data, and under some assumptions, return a graph indicating the causal relationships among the variables in data.

- In quantum mechanics (QM), there are weird correlations whose causal explanation has been debated since the early days of QM.

- It has been argued that CDAs yield unsatisfactory results if applied to quantum entanglement correlations.

Motivation (2)

- To resolve this challenge, some physicists have started to radically extend the whole framework of causal modeling to make it compatible with quantum scenarios.

- This talk paves an alternative way: I will introduce a novel approach to the causal problem of quantum entanglement using recent developments in machine learning (ML).

- I will explain how an ML-based CDA can be used to "learn" the causal structure of a problematic quantum scenario.

- More generally, I will discuss how ML tools can help us address some debates in quantum foundations.

Contents

What is the Problem? (1)

- **The causal Markov condition (CMC)** Every statistical dependence implies a causal dependence:

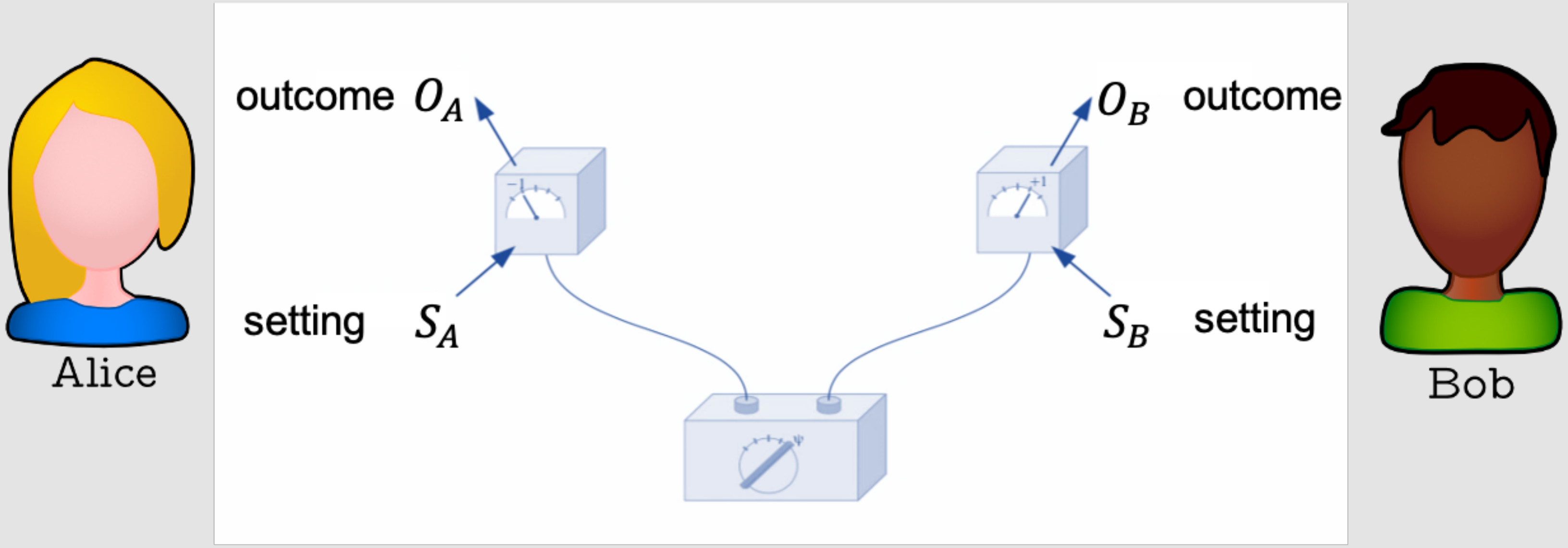
$$(X \not\perp\!\!\!\perp Y \mid Z) \implies (X \not\perp\!\!\!\perp Y \mid Z)_{\mathrm{c}}$$

- **The causal faithfulness condition (CFC)** Every statistical independence implies a causal independence:

$$(X \perp\!\!\!\perp Y \mid Z) \implies (X \perp\!\!\!\perp Y \mid Z)_{\mathrm{c}}$$

- CDAs usually rely on conditional independence (CI) tests to identify the causal structure of a scenario.

What is the Problem? (2)



CHSH scenario

- **The Causal Problem of Entanglement** CI-based discovery algorithms (that respect the CMC and the CFC assumptions) cannot provide a causal graph that is compatible with the statistics of the CHSH scenario.

What is the Problem? (3)

- Several solutions have been proposed to deal with the causal problem of entanglement.

- The present work is mainly inspired by the following proposal:

... CI-based causal discovery algorithms do not do justice to Bell's theorem. [Conditional] independences simply do not provide enough information. ***One needs a causal discovery algorithm that looks at the strength of correlations*** *to reproduce Bell's conclusion. (Wood and Spekkens; 2015, p.17)*

Causal Discovery Algorithms (1)

- CDAs are traditionally divided into three main families:

1. **Constraint-based** (e.g., the PC algorithm): identify CI relations between variables via statistical tests => link the variables that are not found to be independent.

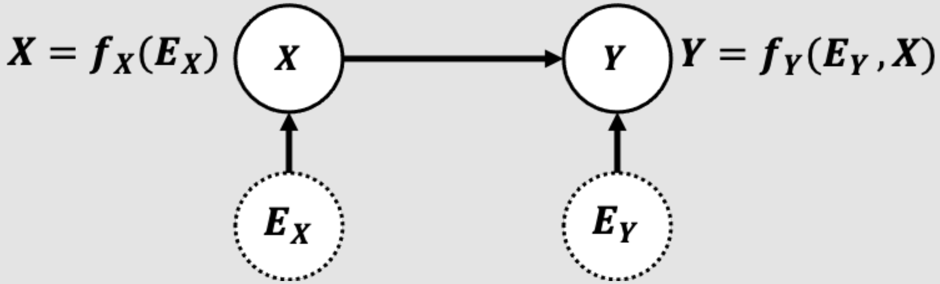
2. **Score-based** (e.g., the greedy equivalent search algorithm): define a global score function such as likelihood and BIC score => search for an optimal DAG that maximizes the score function.

3. **Hybrid** (e.g., Max-Min Hill-Climbing algorithm): combine ideas from constraint-based and score-based algorithms => make the searching process computationally efficient.

- There is, however, another family of CDAs that is restricted to scenarios with only two variables:

4. **Bivariate/pairwise** (e.g., the ANM algorithm): focus on the functional forms of the causal mechanisms => extract asymmetries induced by causal directions to orient the causal edge between two variables.

Causal Discovery Algorithms (2)



- **Two groups of bivariate CDAs:**

1. **Discriminative Models:** construct the most discriminant features available in the observational data => use these features to infer the causal direction

2. **Generative Models:** generate data in both directions via estimating the mechanisms => The causal direction is the one for which the generated data matches more with the observational data (or the mechanism is less complex)

Causal Discovery Algorithms (3)

- A few ML approaches combine the best of both worlds (i.e., traditional CDAs and bivariate CDAs).

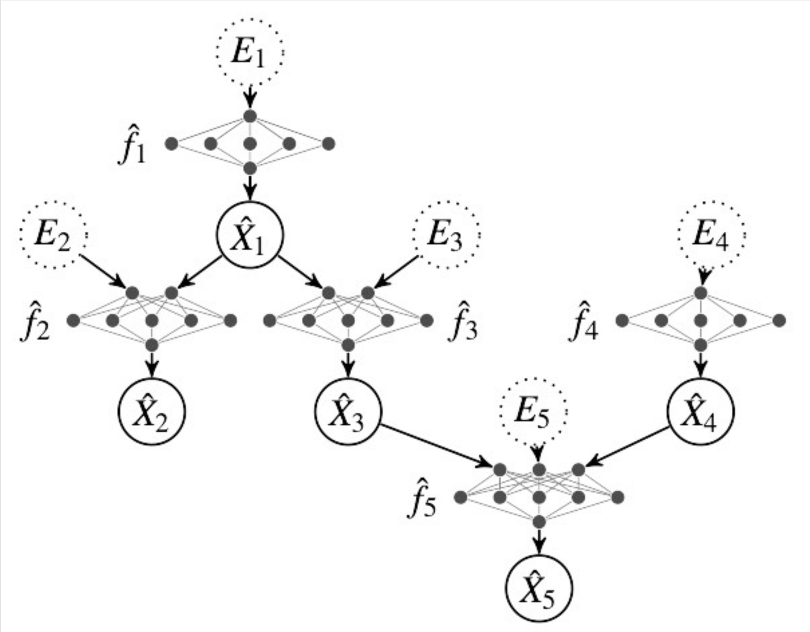
- **The main idea is as follows:**

1. **Search** for all plausible causal graphs.
2. **Learn** (or estimate) the functional form of the mechanisms for each of the graphs.
3. **Generate** synthetic data for each of the graphs.
4. **Compare** the real observational data with the returned synthetic data sets.

=> choose the graph that returns a *better* data set.

CGNN (1)

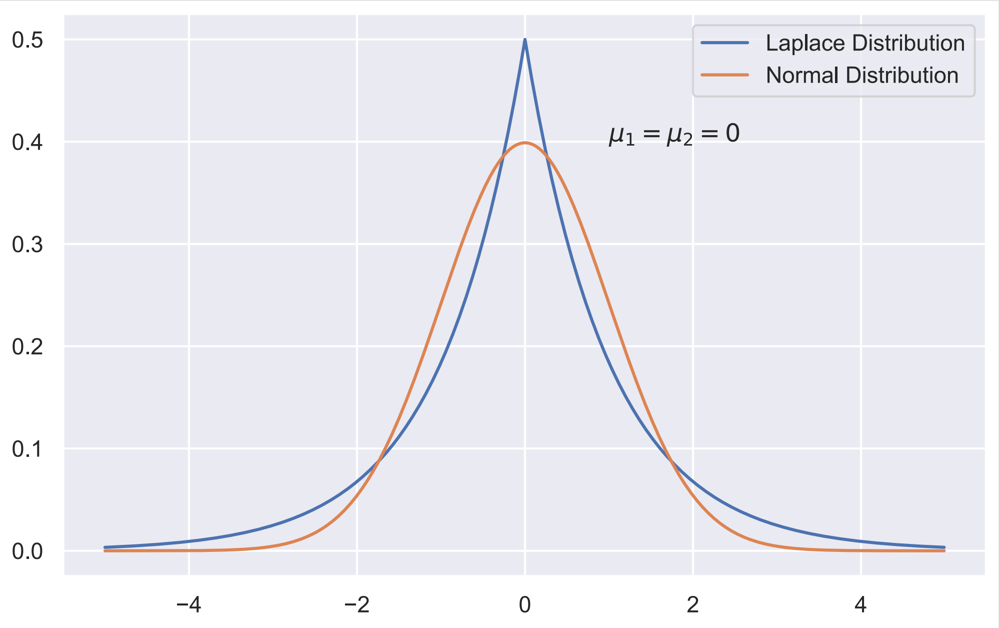
- The **Causal Generative Neural Networks (CGNN)** is an ML-based CDA that combines lots of advanced techniques from other successful discovery algorithms.
- To learn the causal mechanisms of a scenario (the learn step), the CGNN uses neural networks (NNs).
- To compare the real data with the synthetic data sets (the compare step), the CGNN uses a statistical criterion called maximum mean discrepancy (MMD).



CGNN (2)

➢ **What is so special about the CGNN?**

1. In contrast to traditional CDAs, it checks all possible causal graphs.
2. NNs are known to have a high capacity for learning very complex patterns.
3. By choosing the MMD as loss, the CGNN generates synthetic data whose entire distribution is close to the real data (not just a few features of it.)



CGNN (3)

➢ **What is problematic about the CGNN?**

1. The number of plausible graphs exponentially grows.
2. It is computationally super expensive => requires GPU power
3. It works only with continuous variables => cannot be applied to the standard CHSH data

A Simulated EPR Scenario

- Collaborating with Jan Dziewior from the Max Planck Institute of Quantum Optics, we simulated a novel quantum scenario.

- From a physical perspective, this is an extension of the EPR scenario where the settings and the outcomes are continuous. The joint distribution of our data:

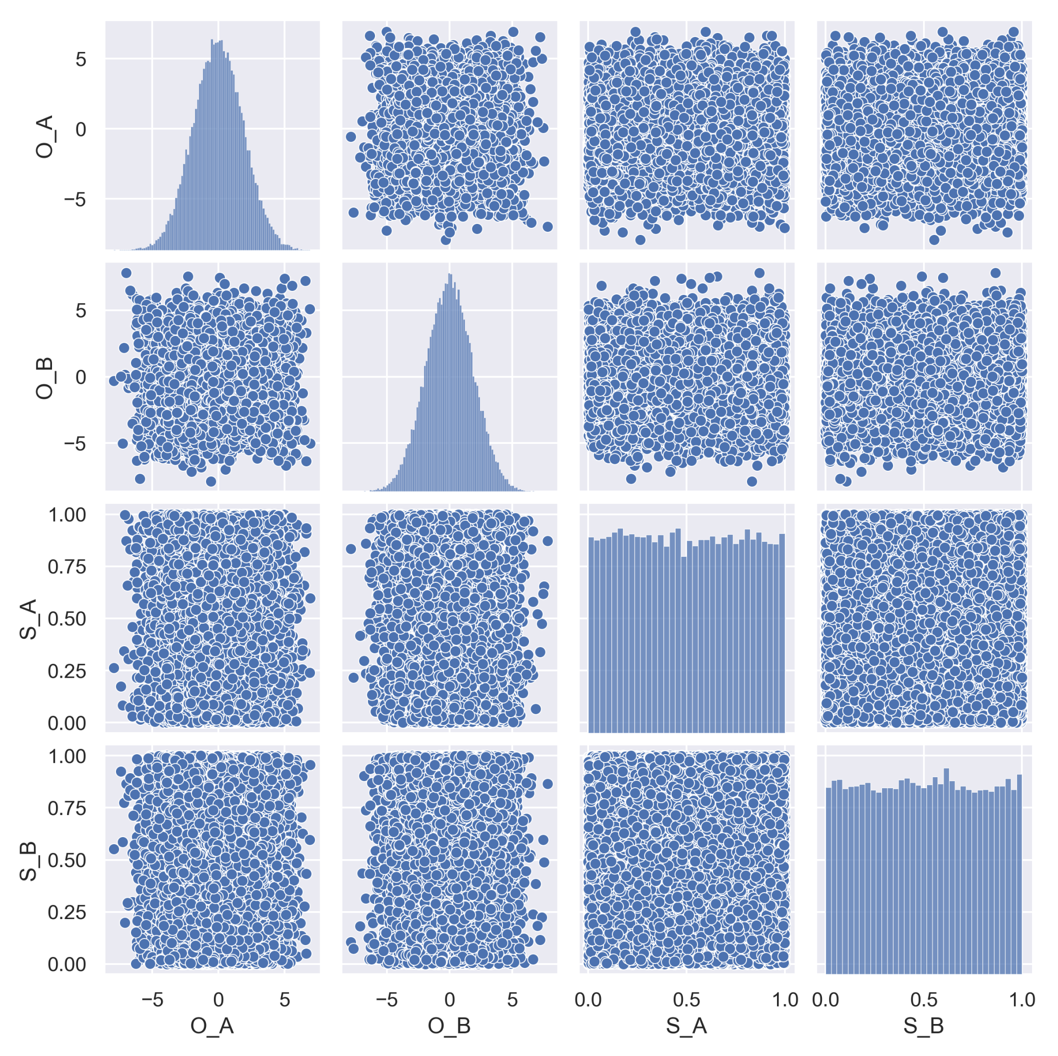
$$p(O_A, O_B, S_A, S_B) = p(O_A, O_B \mid S_A, S_B)\, p(S_A)\, p(S_B)$$

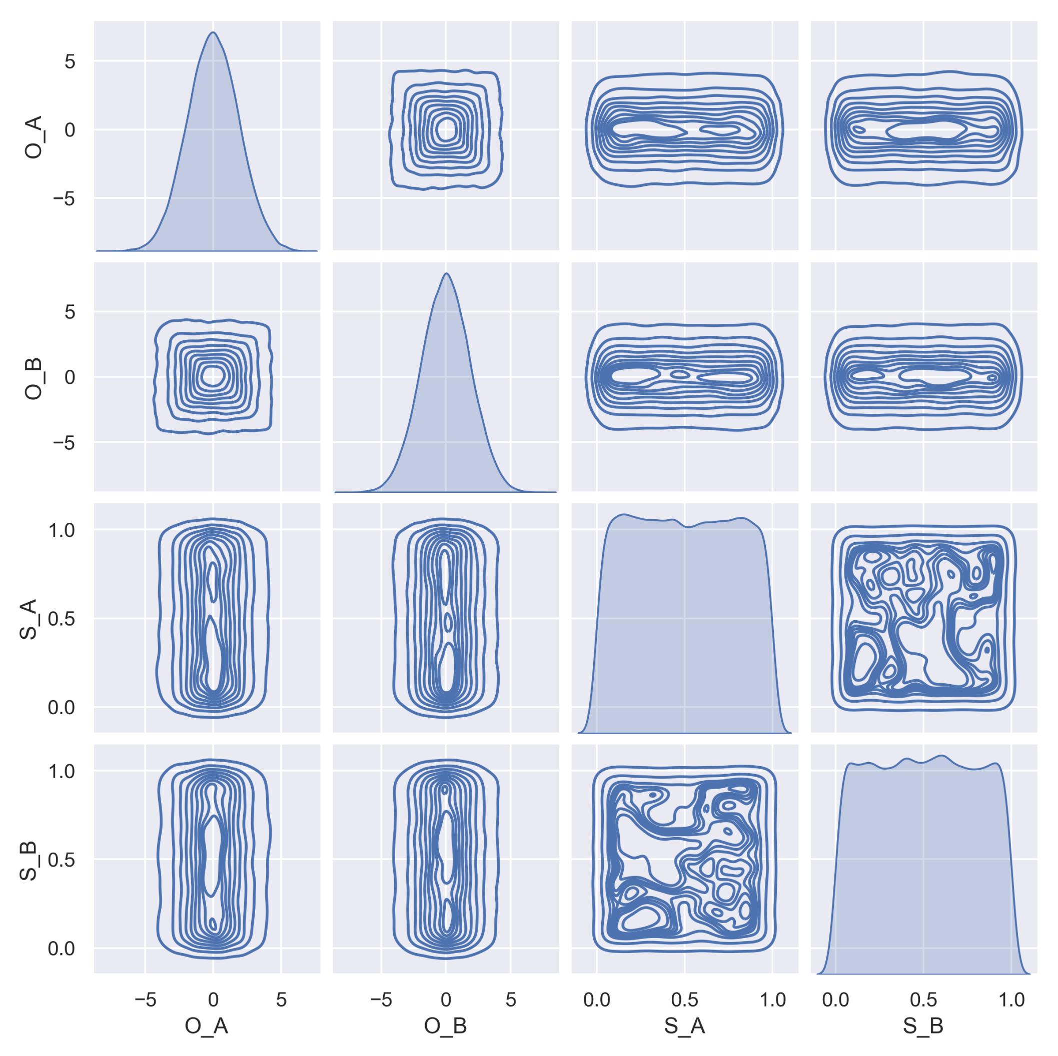
$$\text{where } p(O_A, O_B \mid S_A, S_B) = k_1 e^{-(S_A^2 + S_B^2 - 2\, k_2\, S_A S_B)}$$

- The two settings are marginally independent; the outcomes are conditionally dependent.

◉ **Disclaimer:** we have not had GPU yet, so not all the coming plots are reliable!

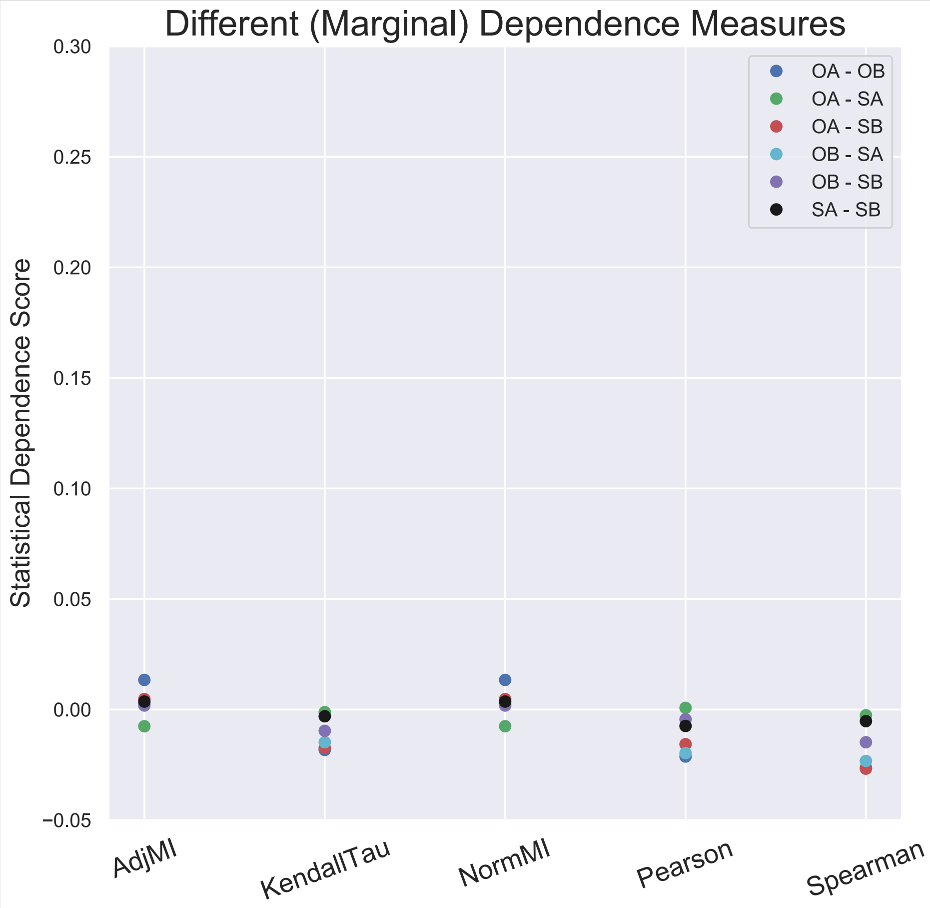
Distributions of Variables





- Our two settings have uniform distributions, while our outcomes are normally distributed.

Marginal Independeces in EPR

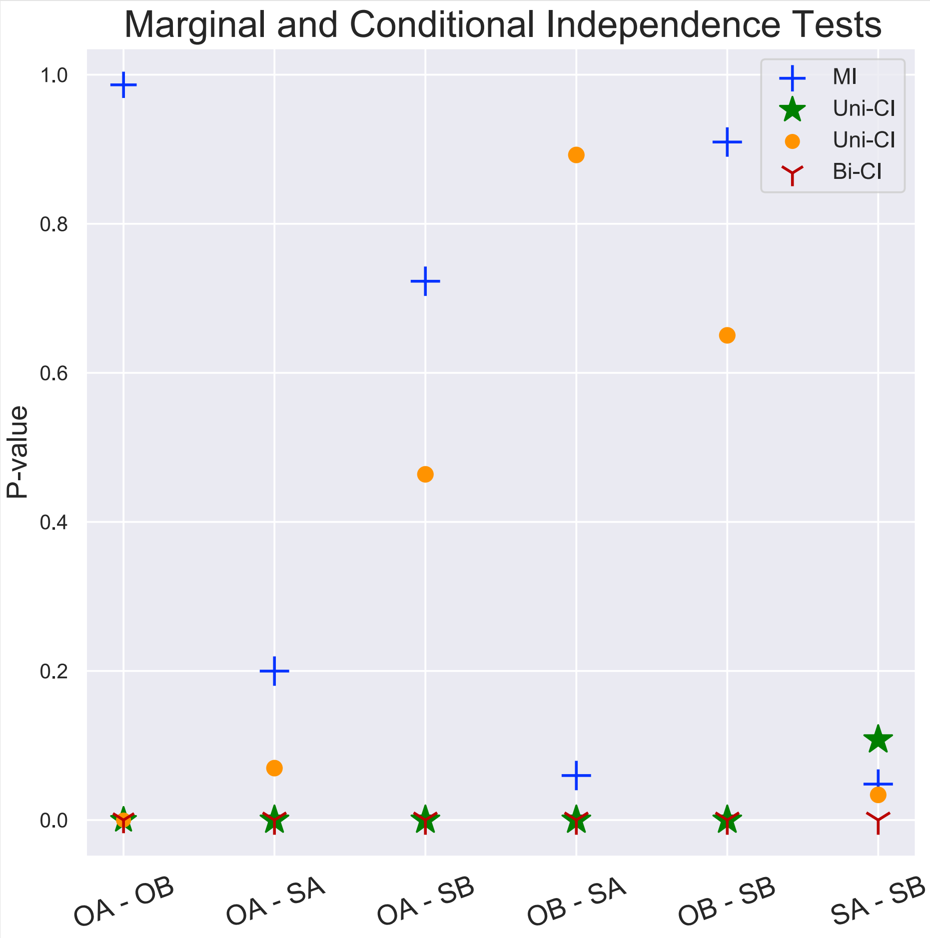


- All of our variables are marginally independent.
- The first non-trivial result is that the CFC is violated by our scenario as soon as we incorporate a causal edge between any two nodes.

Conditional Independences in EPR

$$\text{p-value} \sim 0 \implies X \not\perp\!\!\!\perp Y$$

$$\text{p-value} \sim \mathcal{U}(0,1) \implies X \perp\!\!\!\perp Y$$



- Although our variables are marginally independent, they are conditionally dependent. In particular:

$$O_A \perp\!\!\!\perp O_B \text{ but } O_A \not\perp\!\!\!\perp O_B \mid (S_A, S_B)$$

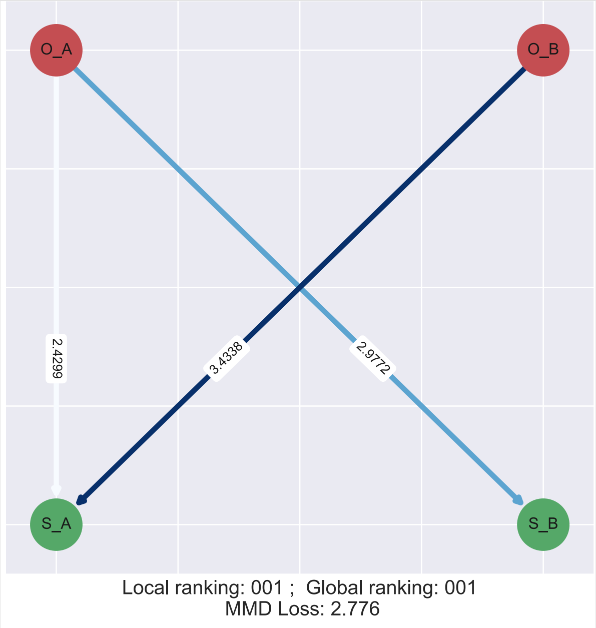
Traditional CDAs on EPR

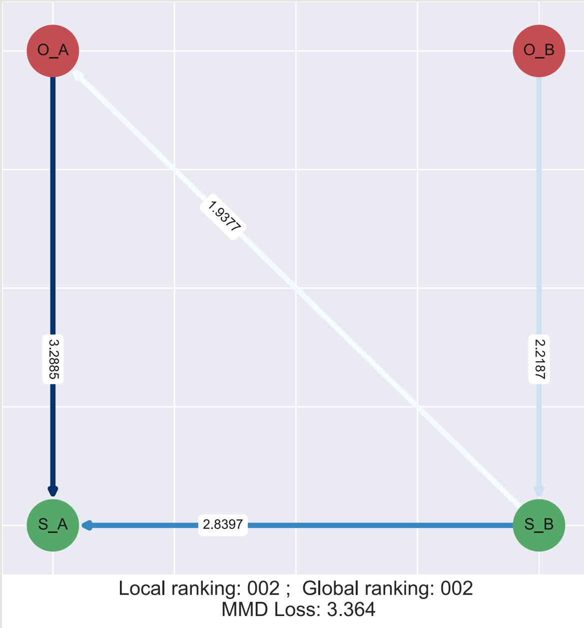
- LiNGAM, a bivariate CDA, is the only algorithm that returns a different (but still unareasonable) graph.

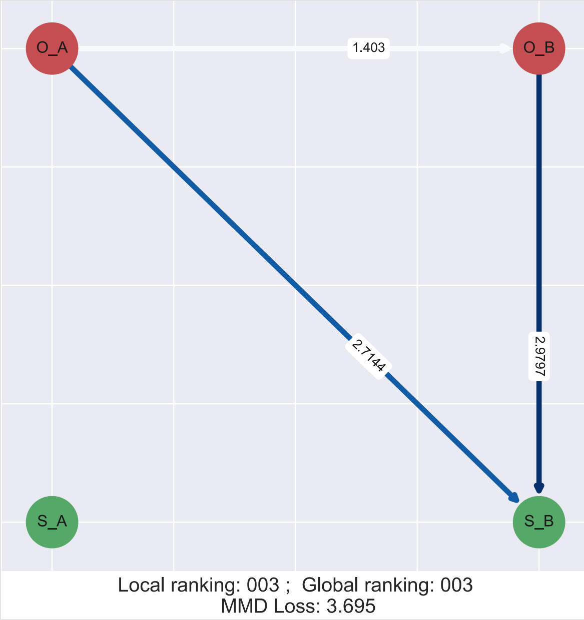
Best CGNN Graphs (1)

- In its standard format, the CGNN returns the best graph only. But, I modified the original codes so we can take a look inside the black box!

- We get 543 graphs; the first three one are:





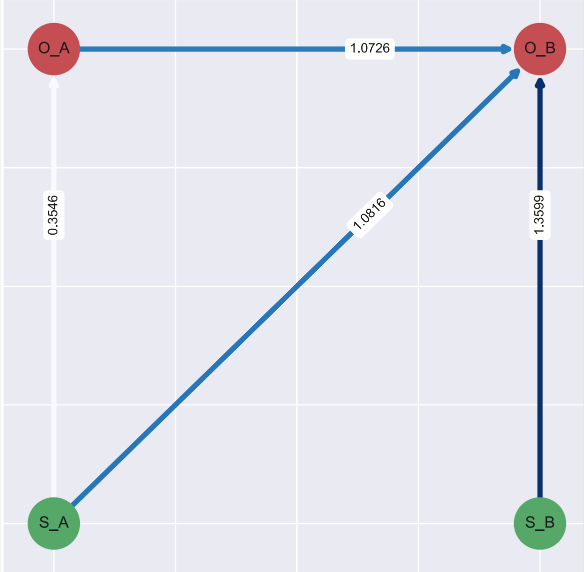


- The color of (and the number written on) each edge represents the importance of that edge for that graph.

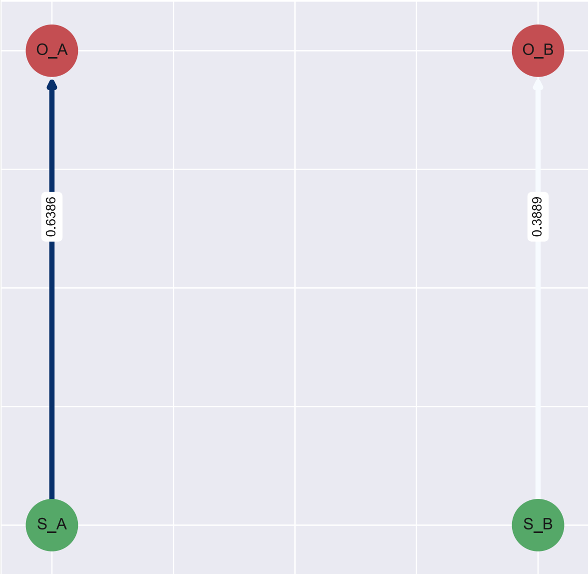
- The MMD Loss represents the quality of the graph: the higher the MMD Loss, the worse the graph.

Best CGNN Graphs (2)

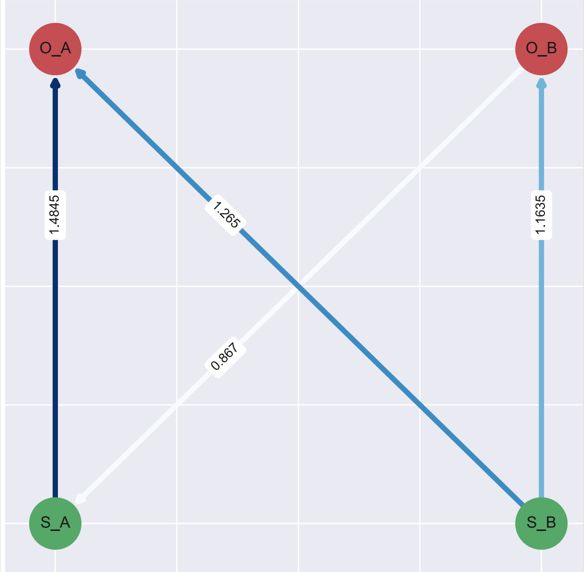
- In a restricted search space (where the settings are independent, and the outcomes are connected to their settings), we have 20 graphs; the first three graphs are:



Local ranking: 001 ; Global ranking: 025
MMD Loss: 4.725



Local ranking: 002 ; Global ranking: 048
MMD Loss: 4.94

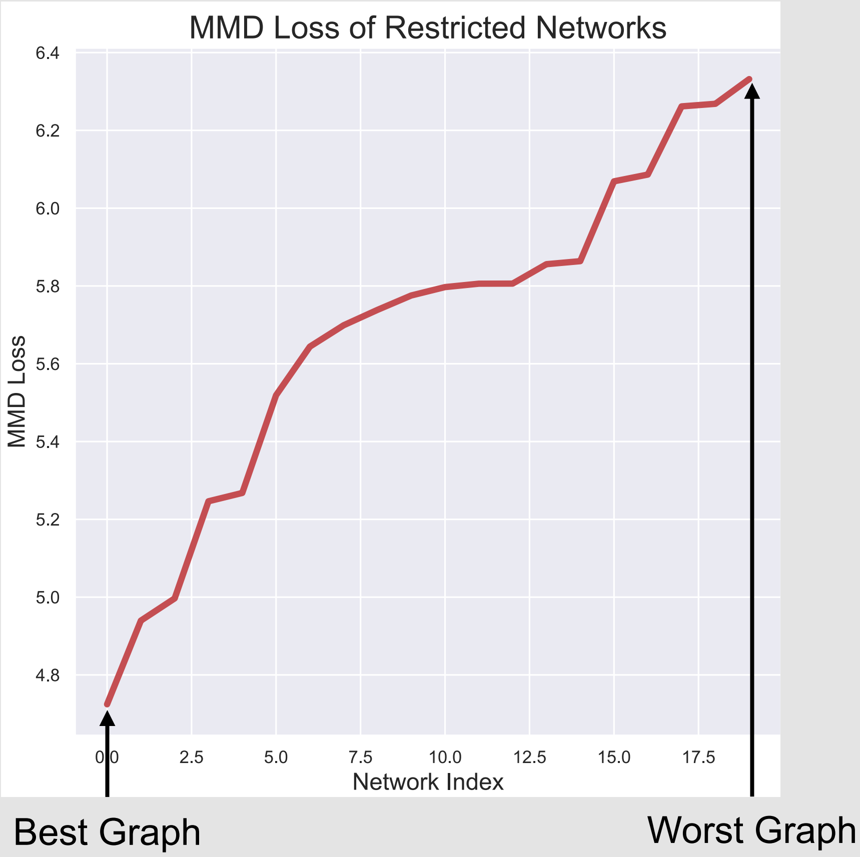


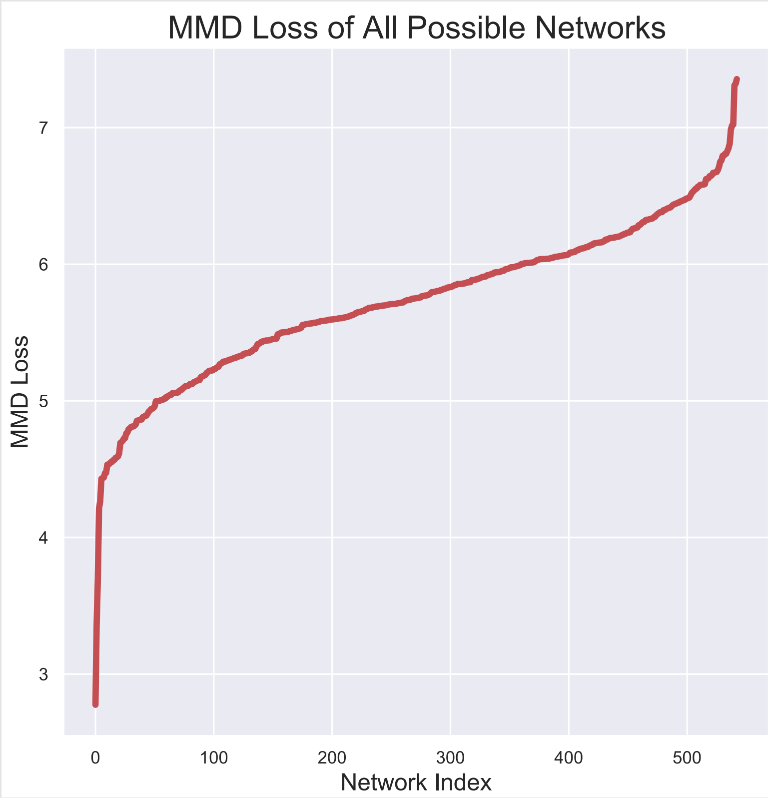
Local ranking: 003 ; Global ranking: 052
MMD Loss: 4.997

- The "local" and "global" rankings indicate the graph ranking in the restricted and unrestricted spaces, respectively.

Distribution of the MMD Loss

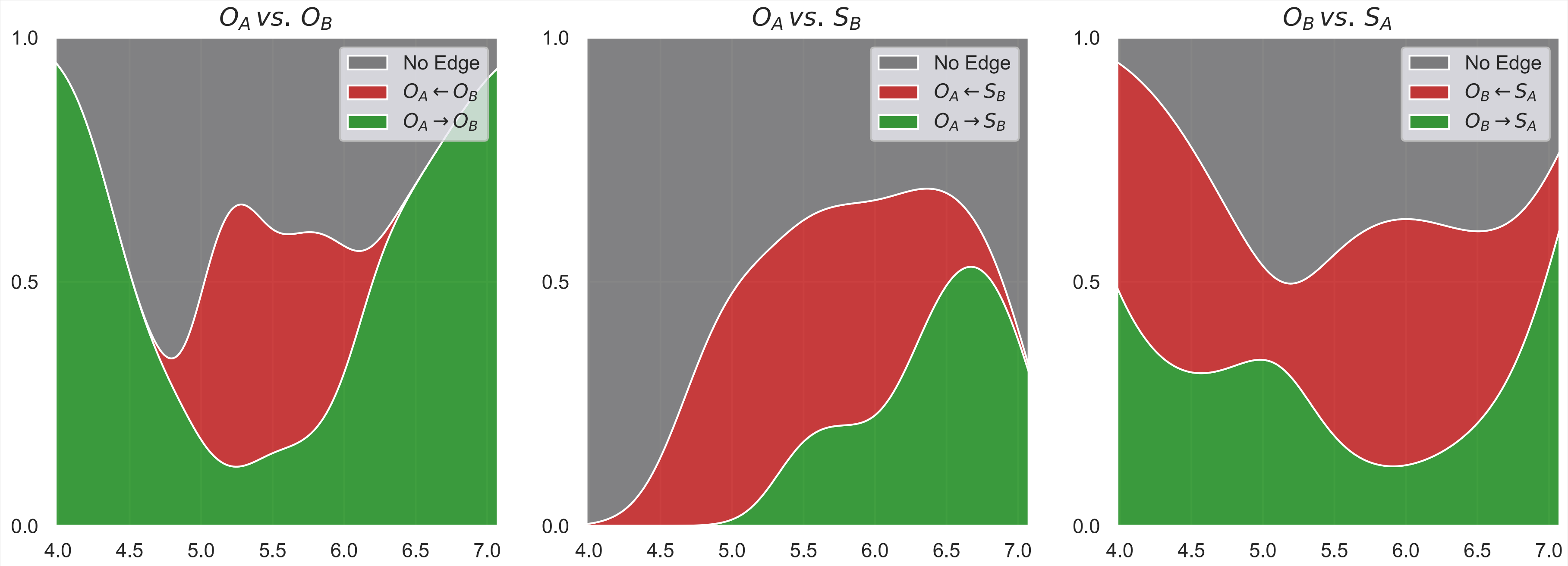
- To compare different returned graphs, let us visualize their MMD loss.





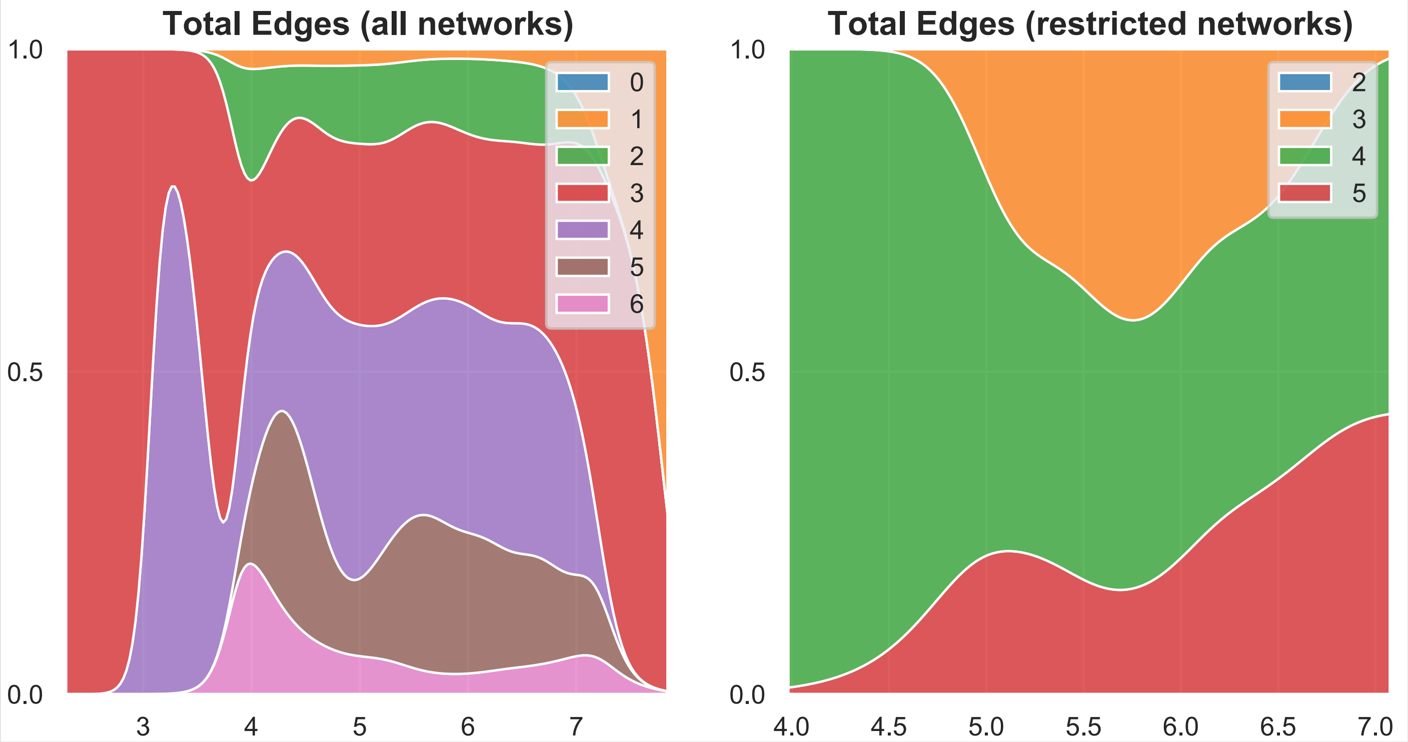
Orientations of Edges (1)

- Let us check how the orientations of edges vary from the best to the worst returned graphs.
- In the restricted space:



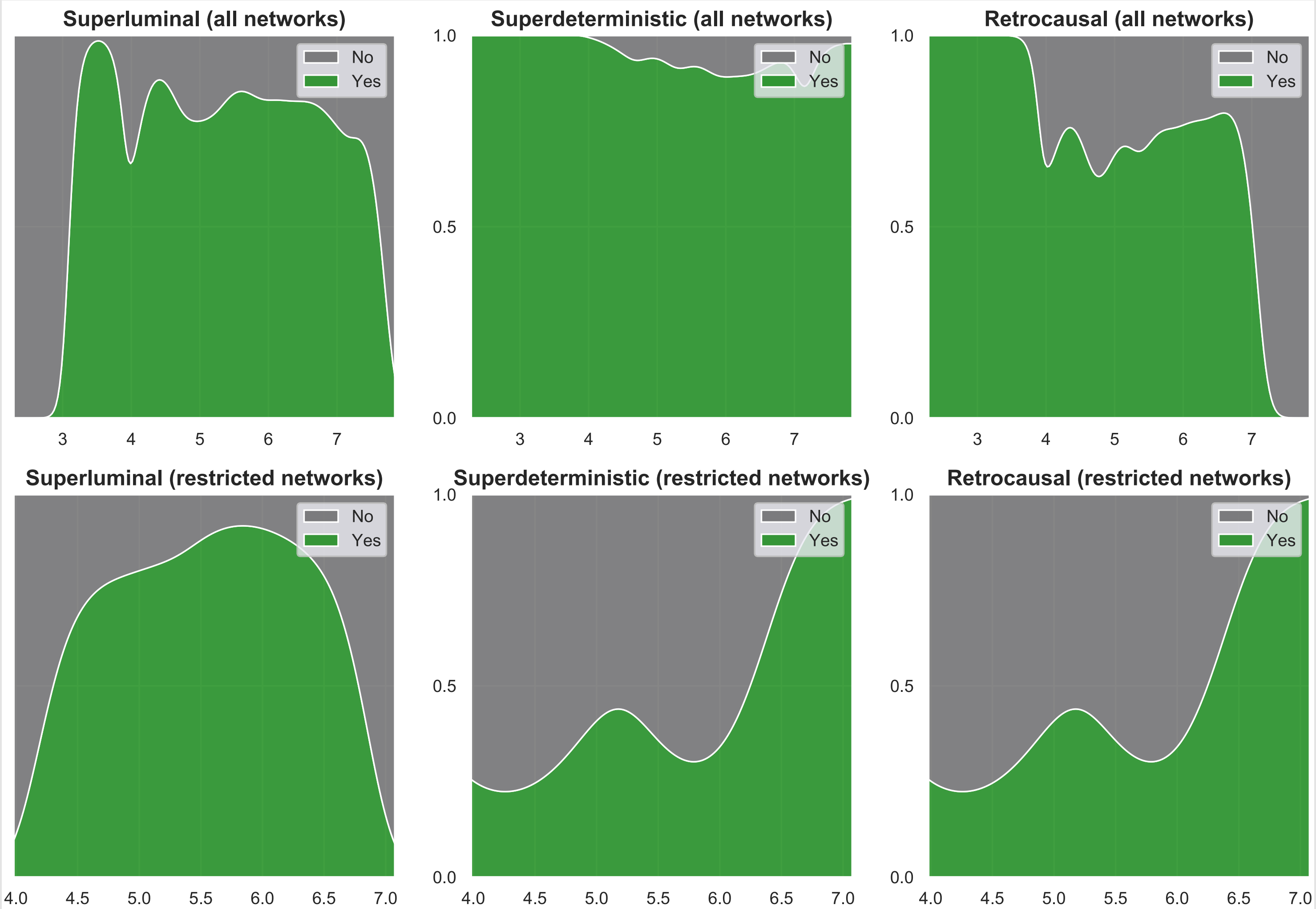
Total Number of Edges

- In ML algorithms, it is important to avoid overfitting and underfitting.
- In causal graphs, these two notions can be tracked via counting the total number of edges.
- Let us check how the total number of edges varies from the best to the worst graphs.



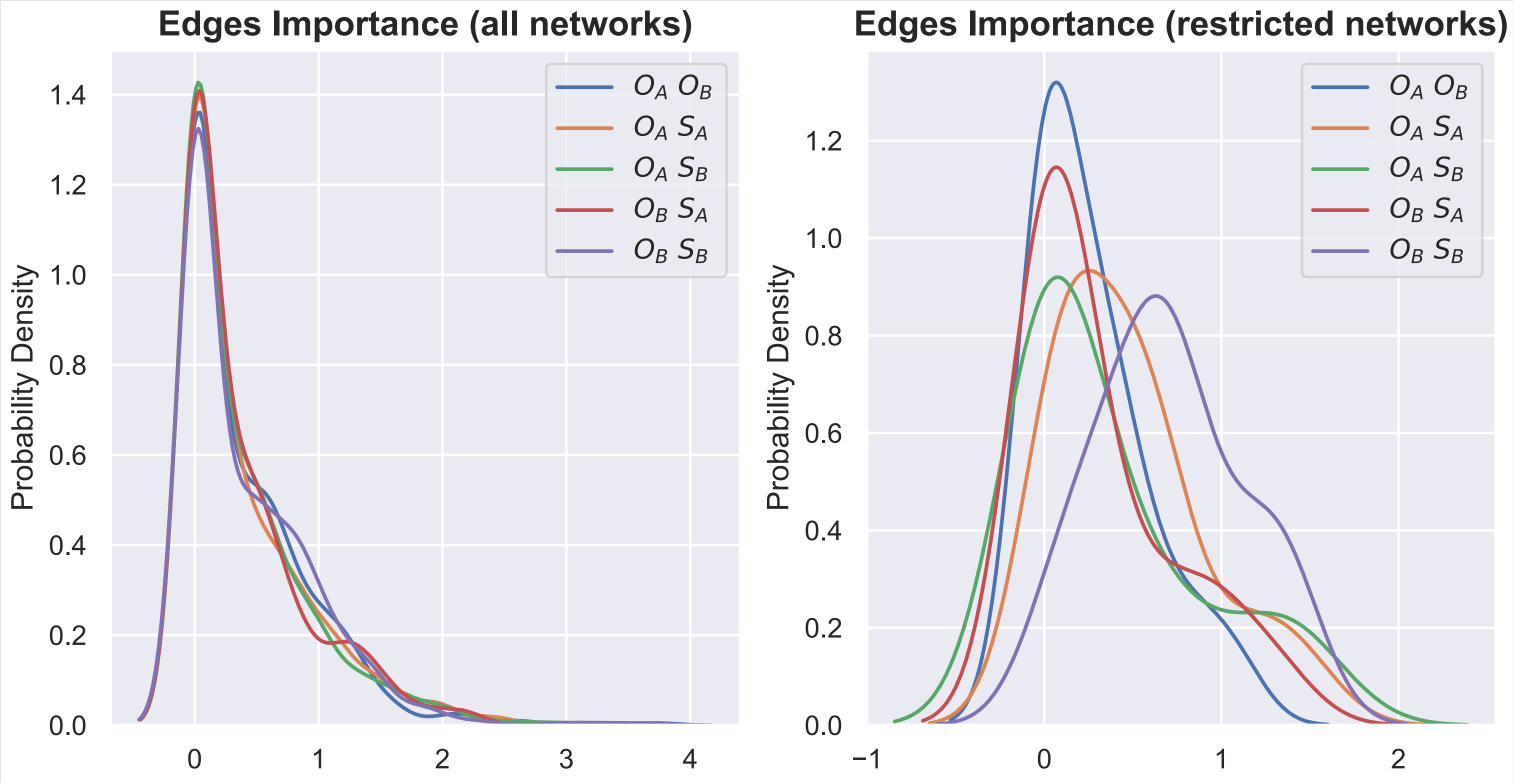
Interpretations of QM

- Superluminal, superdeterminism, and retrocausal interpretations of QM can be identified from the causal graph of a scenario => we can check how these interpretations are distributed from the best to the worst graphs.



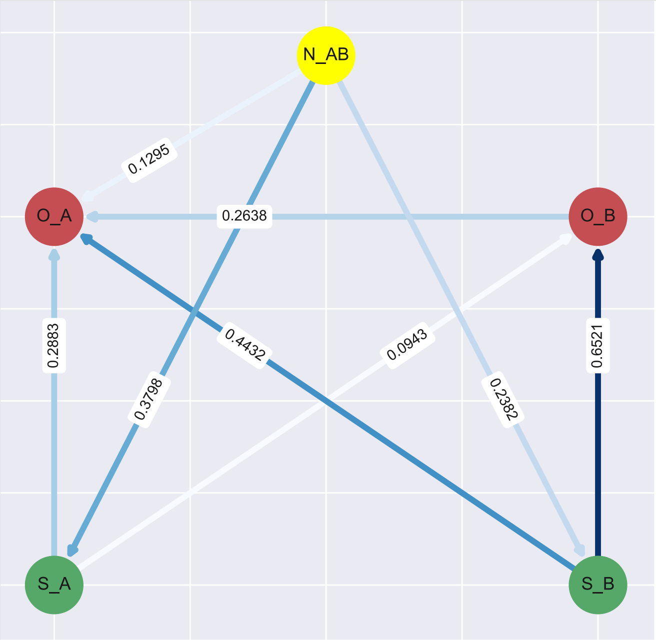
Importance of Edges

- The importance of each edge (i.e., the number that was written on each edge) is another parameter to be monitored:

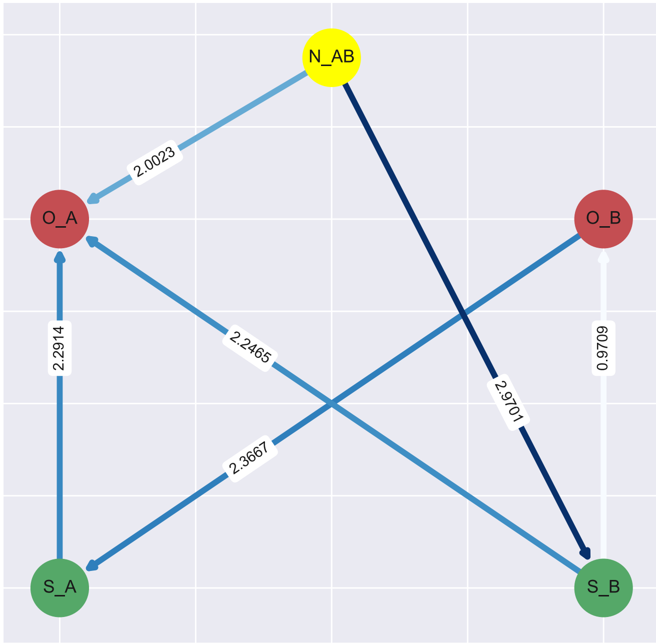


Hidden Variables (1)

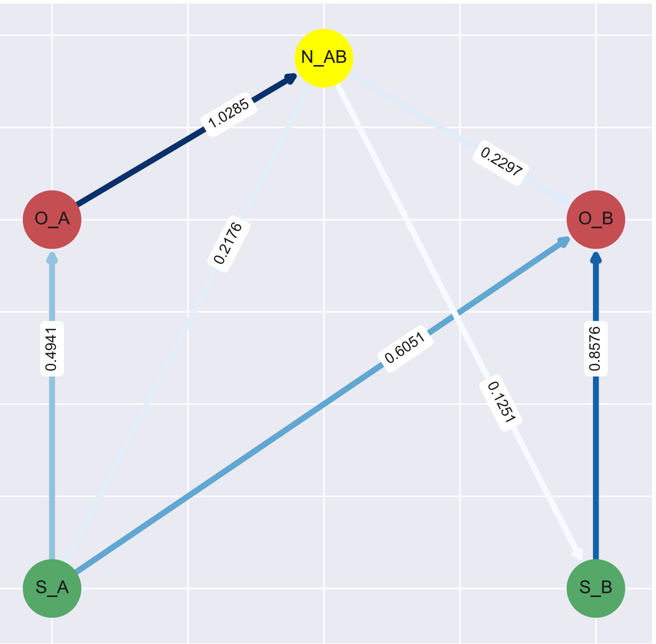
- ➢ CGNN can also handle hidden variables. Considering only one hidden variable, the first three graphs are:



Local ranking: 002 ; Global ranking: 041
MMD Loss: 4.256



Local ranking: 001 ; Global ranking: 022
MMD Loss: 3.978



Local ranking: 003 ; Global ranking: 071
MMD Loss: 4.442

Lessons (1)

- Besides solving highly complex problems in different disciplines, ML can provide us with new tools for approaching foundational questions.

- Of particular interest might be generative ML. Generative algorithms estimate the entire distribution of a given data set => they mimic the actual generative processes.

- These algorithms contain lots of internal parameters which cannot be interpreted directly.

- New insights into a given physical process can be gained if one examines the relations between the internal parameters of the algorithm and the actual parameters of the physical process.

Lessons (2)

- **A concrete example of utilizing this strategy for a foundational debate in QM and causal modeling:**

- Some supporters of the retrocausal interpretation of QM have argued that the CFC violation does not pose a serious challenge for retrocausation because this violation arises from internal hidden symmetries in retrocausal models.

- By looking at the internal parameters of the generative model sketched in this talk, we now can check whether the said conjecture (i.e., the presence of internal symmetries in retrocausal graphs) is promising or not.

- To this end, we can examine whether there exists a meaningful relationship between 1) the symmetries of our models' internal parameters (e.g., the importance of edges or the NN's weights) and 2) whether a graph is retrocausal or not.

Lessons (3)

- In other words, *the original question can be converted into a new causal question about the relationship between the internal parameters of ML and the external parameters of the given physical process.*

- Other foundational questions that might be addressed in this manner are:
  - the status of quantum contextuality,
  - debates surrounding parameter independence vs. outcome independence,
  - the status of the principle of independent causal mechanisms.

◉ **Last Words:** ML-based CDAs can be even more useful if we already know the true causal structure of the scenario. In such situations, we can examine what the machine has learned. This is the subject of my other project, which is going to be completed soon!

References

Bell, J.S., 1964. On the Einstein Podolsky Rosen paradox. Physics Physique Fizika, 1(3), p.195.

Einstein, Albert, Boris Podolsky, and Nathan Rosen. "Can quantum-mechanical description of physical reality be considered complete?." Physical review 47.10 (1935): 777.

Clauser, John F., et al. "Proposed experiment to test local hidden-variable theories." Physical review letters 23.15 (1969): 880.

Wood, C. J. & R. W. Spekkens (2012). The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning.

Näger, P.M., 2016. The causal problem of entanglement. Synthese, 193(4), pp.1127-1155.

Pearl, J. (2000). Causality: Models, reasoning, and inference. Cambridge: Cambridge University Press.

Guyon, Isabelle, Alexander Statnikov, and Berna Bakir Batu, eds. Cause Effect Pairs in Machine Learning. Springer, 2019.

Peters, Jonas, Dominik Janzing, and Bernhard Schölkopf. Elements of causal inference: foundations and learning algorithms. The MIT Press, 2017.

Goudet, Olivier, et al. "Learning functional causal models with generative neural networks." Explainable and interpretable models in computer vision and machine learning. Springer, Cham, 2018. 39-80.

Lopez-Paz, David, et al. "Towards a learning theory of cause-effect inference." International Conference on Machine Learning. PMLR, 2015.

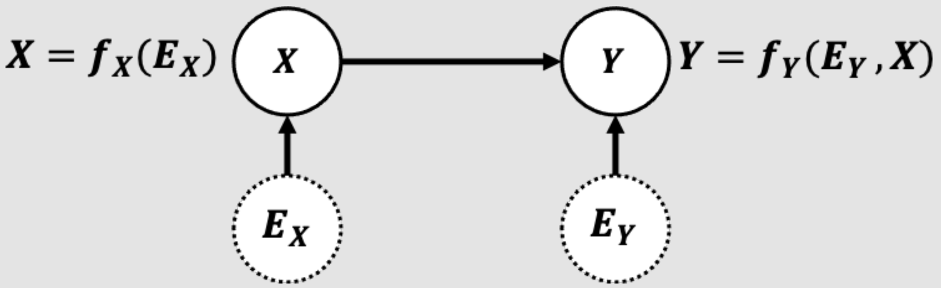
Almada, D., et al. "Are retrocausal accounts of entanglement unnaturally fine-tuned?." arXiv preprint arXiv:1510.03706 (2015).

Evans, Peter W. "Quantum causal models, faithfulness, and retrocausality." The British Journal for the Philosophy of Science 69.3 (2018): 745-774.

Weinstein, Steven. "Neural networks as" hidden" variable models for quantum systems." arXiv preprint arXiv:1807.03910 (2018).

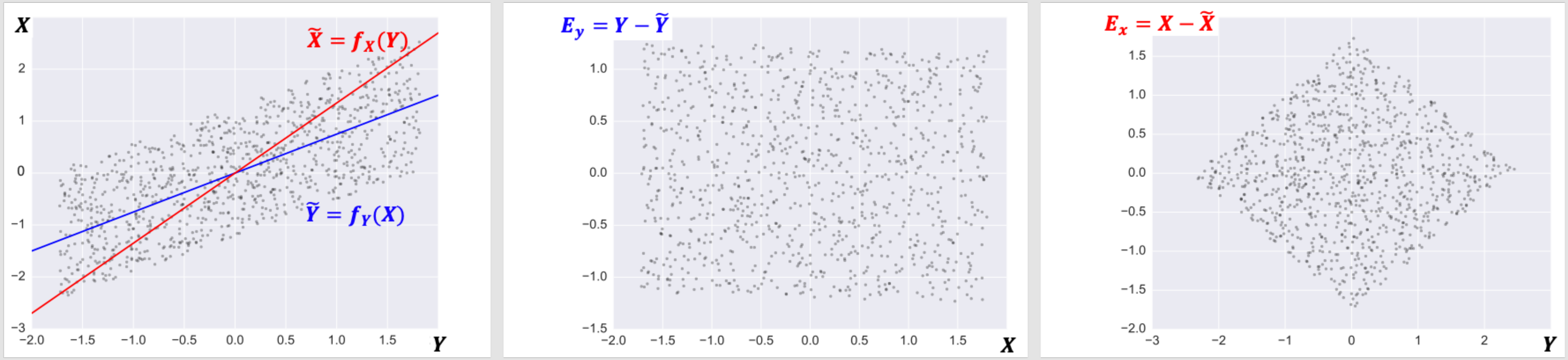
Supplementary Slides

Causal Discovery Algorithms (4)



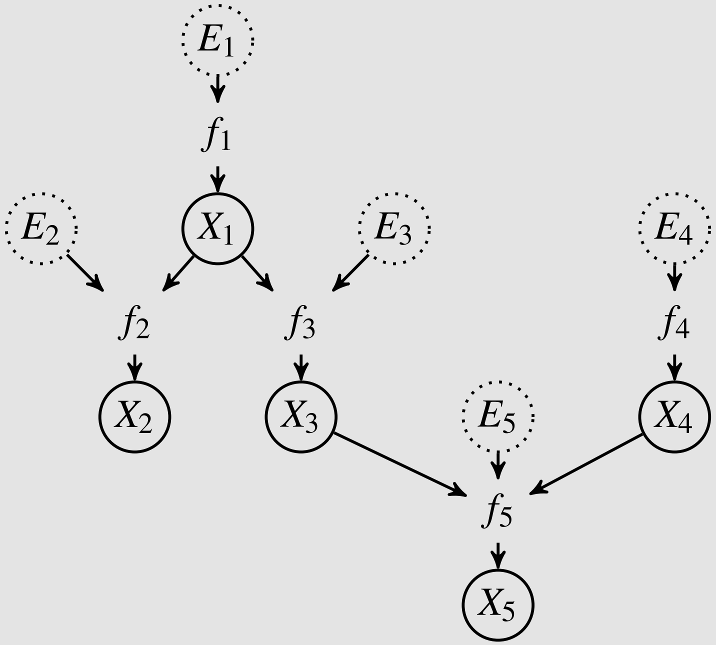
- Observational data contains several types of asymmetries that can be exploited to infer the causal direction between two variables.

- Below is an example of such asymmetries used by a bivariate CDA to infer the direction of causation ($X \rightarrow Y$):



Causal Discovery Algorithms (5)

- Bivariate CDAs can uncover the causal structure of multi-variable scenarios via independently orienting individual edges.

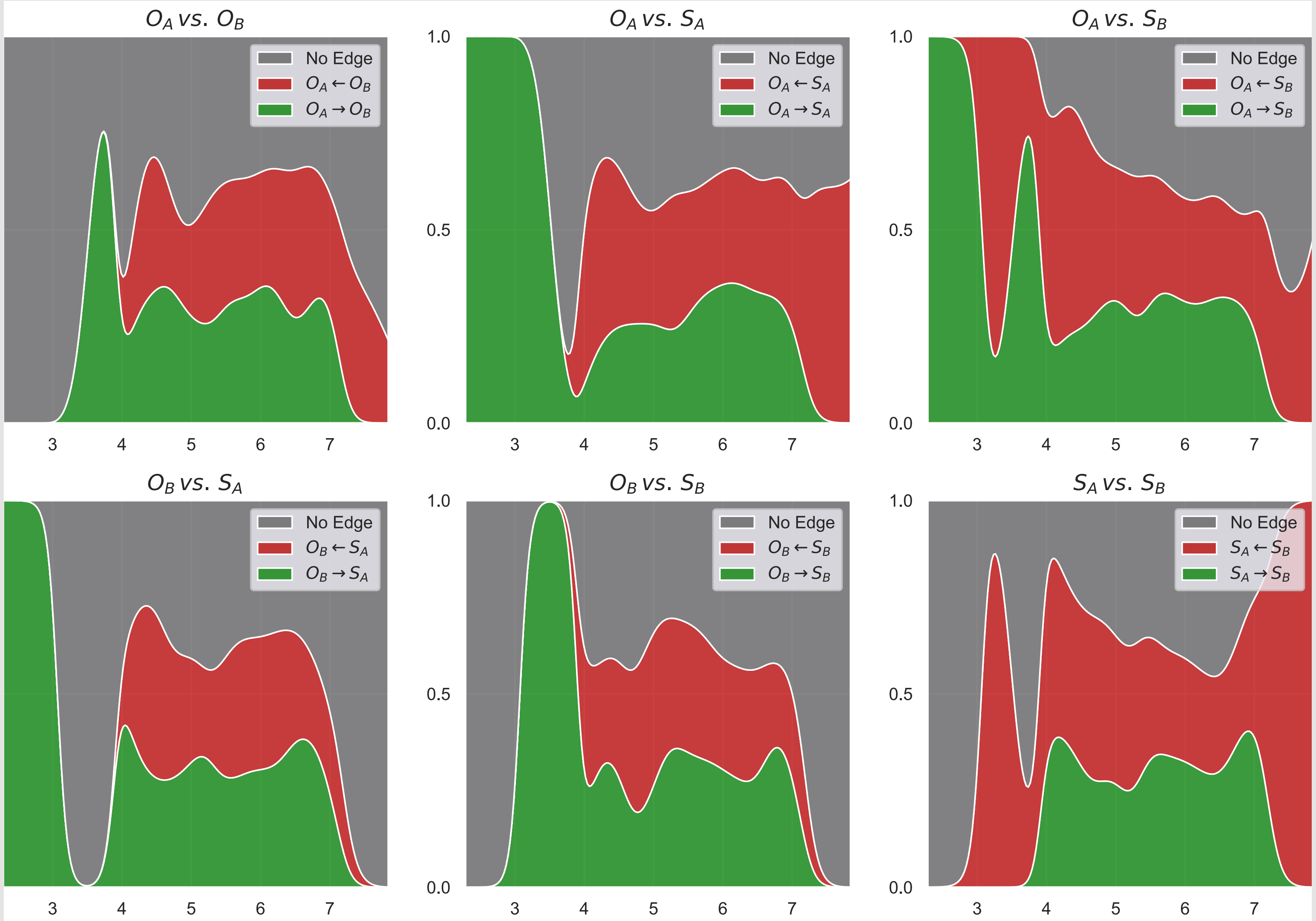


$$
\begin{cases}
X_1 = f_1(E_1) \\
X_2 = f_2(X_1, E_2) \\
X_3 = f_3(X_1, E_3) \\
X_4 = f_4(E_4) \\
X_5 = f_5(X_3, X_4, E_5)
\end{cases}
$$

- But, they would be better complemented with the other CDA families because of relying on local information only => they cannot efficiently see interactions between all variables.

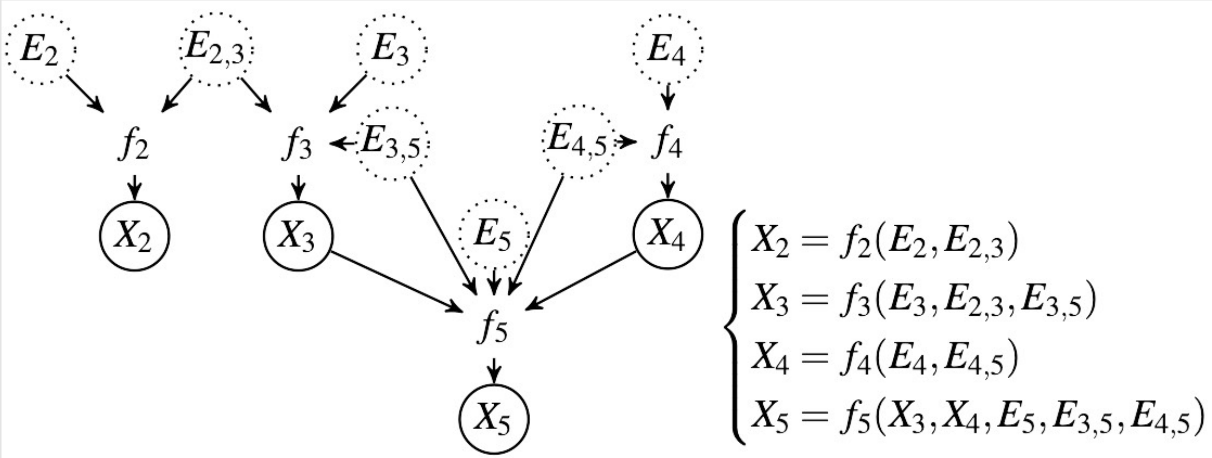
Orientations of Edges (2)

- Another idea is to check how the orientations of edges vary from the best to the worst returned graphs.
- For the unrestricted space:



Hidden Variables (2)

- The original CGNN algorithm provides a solution to handle the hidden variables. This solution is not directly applicable to our scenario as we need the hidden variables to take more complex structures.

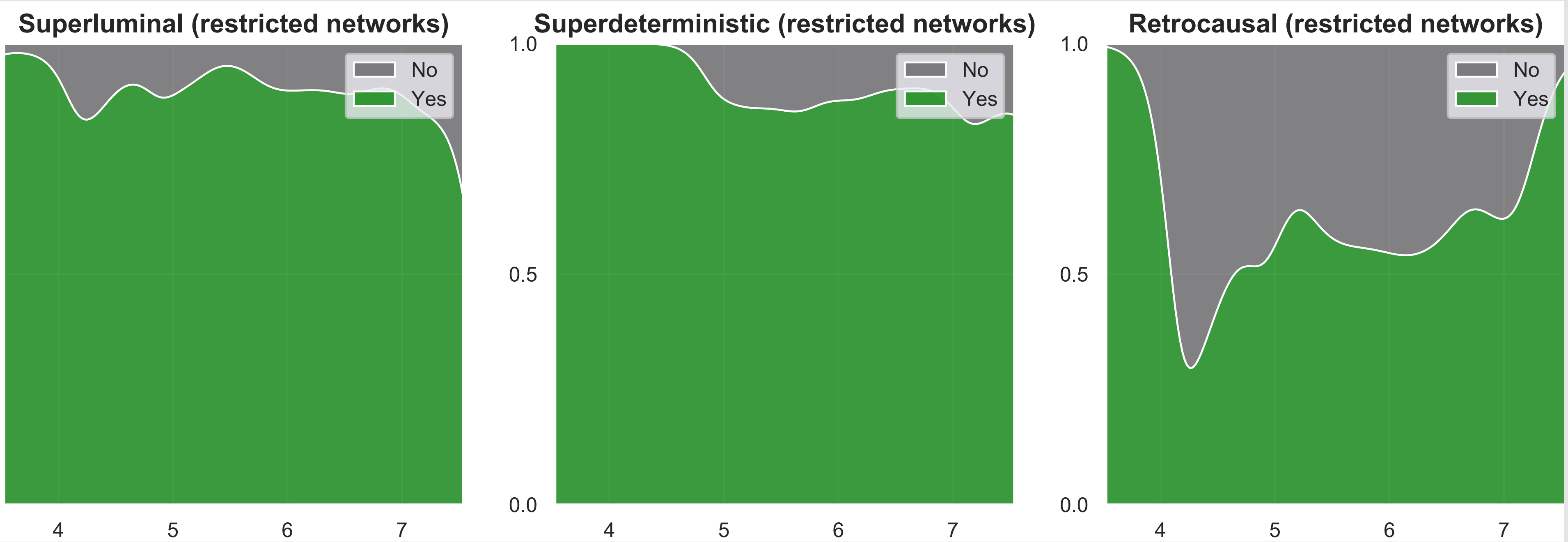


- I thus developed an alternative solution: treat each hidden variable as an observed variable whose values are randomly distributed and then restrict the search space.

- The only problem is again the computational expense!

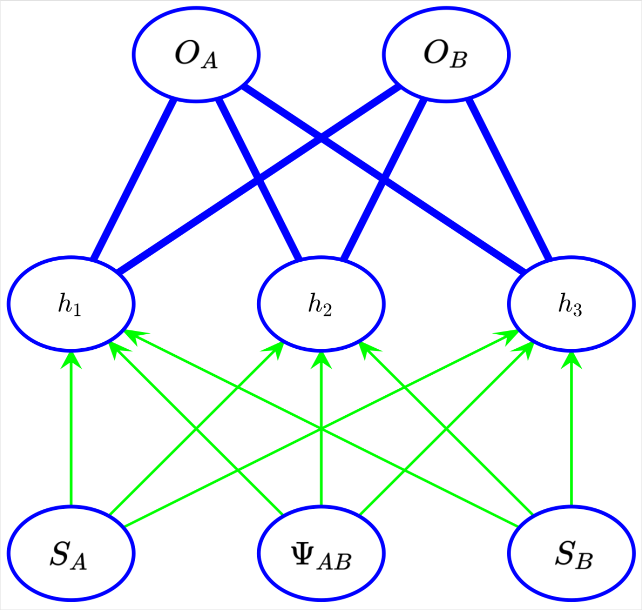
Hidden Variables (3)

➢ All the previous visualizations can be repeated for the case of hidden variables. For instance:



Connections to Other Works

- To the best of my knowledge, our idea is the first work in quantum foundations that employs ML to uncover causal relationships. Weinstein's paper is probably the closest to ours:



- There are also some works exploiting reinforcement learning to determine the maximum violation of various Bell inequalities. Some others have used supervised learning methods to classify classical, quantum, and post-quantum distributions.