

MSc Project - Reflective Essay

Project Title:	Deep Learning for Causal Discovery
Student Name:	Danish Shakeel Mohammad
Student Number:	240830263
Supervisor Name:	Dr. Anthony Constantinou
Programme of Study:	MSc. Artificial Intelligence

Analysis of Strengths and Weaknesses

1. **Vec2ADMG mappings:** The primary theoretical contribution in the report is the introduction of two surjective mappings, which project a continuous vector on the discrete combinatorial spaces of bow-free and ancestral ADMGs. The results for surjectivity, scale invariance, and reachability, ensure that all valid ADMGs are reachable during the optimization process.
2. **Modularity:** The components of the proposed framework, including the scoring metric, the optimization algorithm, and the vec2admg mapping, are all decoupled from one-another. One can, in principle, use different scores like the Akaike Information Criterion (AIC), or use different optimizers like Bayesian Optimization, or if one can come up with a vector-2-graph mapping with the right set of identifiability and consistency constraints depending on the scoring metric, then that can also be swapped out.
3. **Improvement RICF algorithm implementation:** As highlighted in the report, my modification of the ananke-causal python library's implementation of the RICF algorithm was able to consistently achieve a better than 50% speedup over the library's base implementation.
4. **Absence of theoretical guarantees for bow-free algorithms:** The Linear Gaussian bow-free ADMGs are only almost-everywhere identifiable, in contrast to the ancestral ADMGs which are gloably identifiable. This means that for certain parameter configurations, the BIC score is no longer consistent and the minimum value of the BIC score might be achieved on a graph different from the ground truth graph, irrespective of the sample size.
5. **Linear Gaussian Assumption:** The framework is designed for only linear gaussian confounded data. However, many real world systems (like in biology, earth science, climatology, economics etc where systems are governed by sysems of differential equations) are non-linear or have non-gaussian noise. The framework requires that the functional class of causal relationships and error distributions be known a-priori, which is not commonly the case.
6. **High Time Complexity:** The RICF algorithm essentially acts as a bottleneck on the scalability of the proposed framework to graph sizes larger than 30 nodes. Several real-world systems, however are larger. For example biological systems like perturbational gene expression datasets could have thousands of variables.

Presentation of the possibilities for future work

1. **Alternatives to Proximal Policy Optimization:** The work done during the project and the extensive set of experiments performed, highlight the large sample in-efficiency of the Proximal Policy Optimization (PPO) algorithm. The contrast with the much better performance

and faster runtime of the Covariance Matrix Adaptation - Evolutionary Strategy (CMA-ES) approach indicates that the issue lies with the PPO algorithm rather than with the general framework. Thus, it would be instructive to evaluate the performance of other more sample efficient deep reinforcement learning based algorithms on this problem. The primary reason for using deep RL is that, as demonstrated in [1], they can be effective in really high dimensional spaces. While the work done in the thesis only looks at graphs that have at-most 30 nodes, the authors of [1], for the comparatively simpler problem of Directed Acyclic Graph (DAG) discovery, demonstrate the effectiveness of PPO even in searching through a 20,000 dimensional space (for 200 node graphs). The main limitation in translating that performance to Acyclic Directed Mixed Graph (ADMG) search is that the computation of the BIC score is much more expensive. Thus, a more sample efficient algorithm than PPO can make better use of each BIC score evaluation.

The main reason why the PPO algorithm is sample in-efficient is its on-policy nature. This means that it discard previously observed samples after a single update epoch. Thus, off policy algorithms, which cache previously observed samples for later re-use, might be an effective avenue to explore. The Soft Actor Critic (SAC) algorithm is the most prominent Deep Reinforcement Learning off-policy counterpart of the PPO algorithm.

2. More Efficient BIC computation through RICF decomposition

The log-likelihood for a DAG can be factorized into independent terms for each node conditioned on its parents. This factorization allows the authors of [1] to cache the partial sums required for computating the BIC score with the node and parents combination acting as the key into the cache. This allows for a constant time update per node modification. The log-likelihood term for the bow-free ADMGs does not permit the same per-node factorization, drastically increasing the time complexity of the BIC computation. The Residual Iterative Fitting (RICF) algorithm, which is used to compute the BIC scores for Linear Gaussian bow-free Acyclic Directed Mixed Graphs, is invoked 10's of thousands of times, every time that we want to find the candidate ground truth graph for a given dataset. Its time complexity is cubic in the number of nodes. Thus, any improvement to the RICF algorithm is likely to have a reasonably large impact on the runtime and scalability of the framework proposed in the thesis.

Although there is no per-node decomposition of the log-likelihood score, it does however, permits a decomposition based on the set of bidirected edge connected components (called districts) in the graph. All the nodes in each district are connected by bidirected paths and there are no bidirected edges between districts. Using this decomposition, we can achieve a time complexity for the RICF algorithm that is cubic in the sizes of the districts. This can be especially impactful for sparse graphs with a large number of small districts.

Work that you would have executed if you had more time

1. **Longer experiment runs:** Each run of PPO-based Relcadilac, for 10-node, degree 4, 2000 sample graphs would take around 30 minutes to run using the hyperparamers mentioned in the report. This was when using 80,000 BIC function evaluation or total timesteps as the threshold for termination. Assuming that this 30 minute runtime does not change with variation in graph parameters like number of nodes or sample size (which is definitely not the case), the production of the graphs in the paper, only for the Relcadilac algorithm, required over 100 hours (4 varying graph attributes, 5 values per attribute, 5 runs per attribute value, and 2 graph classes). The authors of [1] run their PPO algorithm

on the simpler problem of DAGs for 1,280,000 total timesteps, which is a 16 fold increase. Thus, I would have run all the algorithms (even the CMA-ES algorithm) for longer (although not 16 times longer) to get a better idea of their capabilities at convergence.

2. **Alternative dataset to Sachs for real world testing:** As mentioned in the paper, although the Sachs dataset is quite widely used in the causal discovery literature, there is still some debate concerning its ground truth graph. I came to know this fact regarding the Sachs dataset quite late in the thesis work, or I would have substituted it for a better understood real world dataset. I had found the Causal Chambers real-world dataset, which derives its data from a pair of well-understood computer controlled devices that allow them to obtain a large set of observational and interventional datasets with known ground truth causal graphs. In the event of greater time availability I would have liked to explore the performance of the proposed algorithms on this dataset.
3. **More graph variations and model mis-specification:** While the algorithms were evaluated on varying graph densities (in terms of average degrees of the graphs), I would also have liked to evaluate the performance of the algorithms on large sparse graphs (40 or 50 node graphs with average degree 2). I would also be interested in evaluating the performance of the algorithms in the case of model mis-specifications. How do the algorithms handle the case where the noise is sampled from the Uniform distribution or the Laplace distribution rather than the normal distribution. Or if instead of violating the Gaussianity, I kept the Gaussian noise, but instead of linear functions, the data was sampled from a non-linear gaussian SCM. Potential non-linearities could be simple polynomials like quadratic or cubic functions, or more sophisticated non-linearities like a randomly initialized neural network. These experiments might be useful in testing the robustness of the proposed algorithm: is the algorithm more susceptible to non-linearities in the data or to non-gaussian errors?
4. **Comparison against more algorithms:** Another potential improvement to the test suite would be to include more algorithms to compare against. The set of algorithms compared against was limited due to the large amount of time required to test the current set of algorithms. Even the DCD algorithm required more than an hour on the 30-node graphs. Due to limitations of space in the report, some recent algorithms for causal discovery were not mentioned since they weren't directly relevant to the approach taken in the report. One such algorithm is the N-ADMG or Neural ADMG algorithm [2]. The authors establish that under conditions of non-linear additive noise SCMs, bow-free ADMGs, and observable and unobservable variable non-modulation, the ADMG is identifiable. They transform the ADMG into a DAG through "magnification" where they introduce a latent variable for each confounded edge, parameterize the non-linear functions using neural networks, use variational inference to approximate the posterior over the graphs and the latent variables, and use a differentiable constraint to enforce acyclicity, similar to the constraint used in DCD. As a point of comparison, this algorithm could be interesting in that it is a more general algorithm for non-linear confounding being compared to algorithms specifically designed for linear gaussian confounded data, like the one we propose.

Critical analysis of the relationship between theory and practical work produced

1. **Consistency of the BIC score:** The theoretical basis of the proposed framework is the asymptotic consistency of the BIC score. Ancestral ADMGs are globally

identifiable and form smooth curved exponential families for which the BIC score is consistent in the limit of infinite data. In the case of bow-free ADMGs, they don't form smooth curved exponential families and are only almost-everywhere identifiable. Thus, BIC is not consistent for this class of graphs. However, in practice, we don't see an apparent difference in the performance of algorithms between the two graph classes.

2. **Finite sample reality:** The finite sample reality also introduces a disconnect between theory and practice. In some cases, the CMA-ES algorithm was able to find negative fraction BIC excess (fractional BIC excess is the ratio of the difference between the BIC scores of the predicted and ground truth ADMGs divided by the ground truth BIC score) values indicating that it was able to find a graph with a BIC score that was lower than the BIC score of the ground truth ADMG. This is possible, despite the consistency guarantees of the BIC score, because the BIC algorithm guarantees consistency only in the limit of infinite data. Although this indicates the effective search capability of the CMA-ES algorithm, in practice it means that using the BIC score as a goodness-of-fit measure is sometimes suboptimal.
3. **Unknown MEC of Bow-free ADMGs:** Another wrinkle is the unknown markov equivalence class of bow-free ADMGs, requiring that we output a graph that actually belongs to the equivalence class of ancestral graphs.
4. The construction of the margin maximization term which is used in the objective function of the CMA-ES algorithm and the restriction of the search space in the PPO algorithm to a hypercube of side length 2 are contingent on the theoretical results of scale invariance, surjectivity, and reachability proposed and proved in the report.

Awareness of the Legal, Social, Ethical and Sustainability Issues

1. **Reinforcement Learning:** Deep Learning, due the large number of simultaneous parameter updates required, is generally quite compute heavy and needs GPUs to be able to learn effectively. The Reinforcement Learning algorithm chosen, PPO, being an on-policy algorithm is quite sample in-efficient and requires a large number of samples before it converges and shows good results. The PPO algorithm required around 30 minutes for 10 node graphs with 2000 sample size, meaning that if the algorithm were scaled up to match the sizes of several real-world datasets, it might require significant amounts of energy and computation. Additionally, the use of larger and larger deep learning models in causal discovery means that greater and greater computational resources must be made accessible to the practitioners and researchers in the domain of causal discovery. This means that good tools and new techniques might only be concentrated in the hands of the few, creating and exacerbating existing social biases.
2. **Real world impact:** The current state of the proposed framework is not robust enough to be leveraged in practical scenarios where the cost of confidently declaring incorrect causal relations might be high. For example, its use in the fields of economics might cause potentially harmful budget allocations or in the field of medicine it might cause damaging medication to be prescribed.

References:

- [1] B. Duong, H. Le, B. Huang, and T. Nguyen, "Reinforcement learning for causal discovery without acyclicity constraints," *Transactions on Machine Learning Research*, 2025. [Online]. Available: <https://openreview.net/forum?id=sNzBi8rZTy>
- [2] M. Ashman, C. Ma, A. Hilmkil, J. Jennings, and C. Zhang, "Causal reasoning in the presence of latent confounders via neural ADMG learning," in *The Eleventh International Conference on Learning Representations*, 2023. [Online]. Available: <https://openreview.net/forum?id=dcN0CaXQhT>