CONSTRAINT-BASED CAUSAL DISCOVERY FROM A COLLECTION OF CONDITIONING SETS

Kenneth Lee*, Bruno Ribeiro+, Murat Kocaoglu*

Elmore Family School of Electrical and Computer Engineering, Purdue University*

Department of Computer Science, Purdue University+

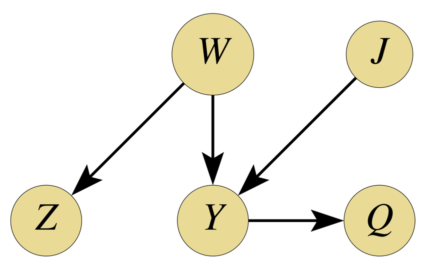
Paper → Code →

INTRODUCTION

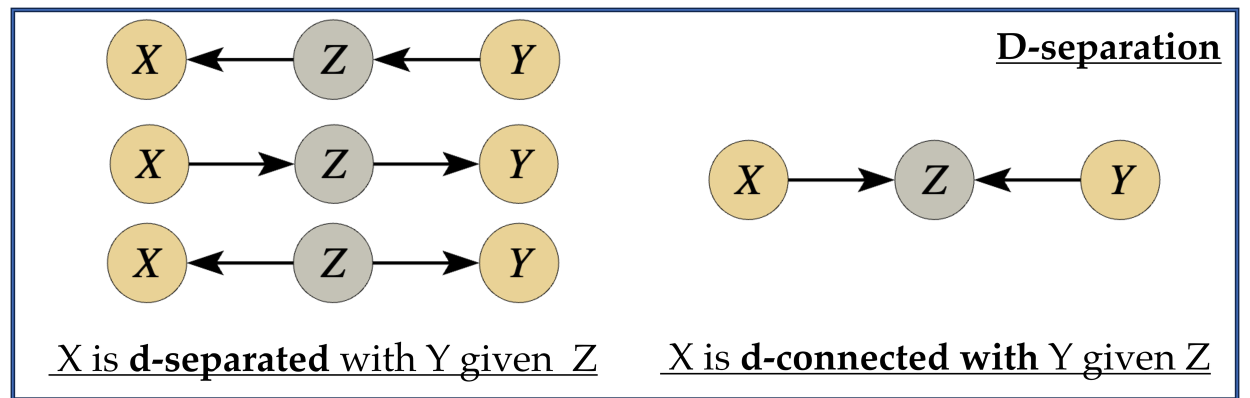
- **Objectives:** learn causal graphs using only CI tests restricted to a collection of conditioning sets.
- **Motivation:** it is not known the characterization of learning of causal graphs from a collection of conditioning sets. Not all the CI tests are equally reliable.
- **Contributions:** propose to learn causal graphs by using CI tests where the conditioning sets are restricted to a given set of conditioning sets including the empty set.

BACKGROUND

- *Pearlian framework* [Pearl' 09]: **Directed acyclic graphs (DAGs)** encode causal relation between variables.
- **Arrows:** Deterministic functional relations called *structural equations.*



$$X_i = f(Pa_{X_i}, E_{X_i}), E_i \perp\!\!\!\perp E_j$$



- There exists work that characterize and learn causal graphs from small conditioning set up to size k: **LOCI [Wienöbst et.al '20], kPC [Kocaoglu. '23].**

$$X \perp\!\!\!\perp Y | \mathbf{Z}, |\mathbf{Z}| \leq k$$

- We further relax the above by taking a more flexible class of conditioning sets called **conditionally closed sets**

CONDITIONALLY CLOSED SETS $\mathcal{C}$

- Let $\zeta = \{I_i\}$ be a set of CI statements of the form $I_i = (X, \mathbf{Z}, Y)$ *i.e.* $(X \perp\!\!\!\perp Y | \mathbf{Z})$ or $(X \not\perp\!\!\!\perp Y | \mathbf{Z})$ A set $\mathcal{C}$ is called **conditionally closed** if the following holds
  1. $\emptyset \in \mathcal{C}$
  2. $\exists X, Y \in \mathbf{V}, (X, \mathbf{C}, Y) \in \zeta \Rightarrow (A, \mathbf{C}, B) \in \zeta$ for all $A, B \in \mathbf{V}$ and for all $\mathbf{C} \in \mathcal{C}$

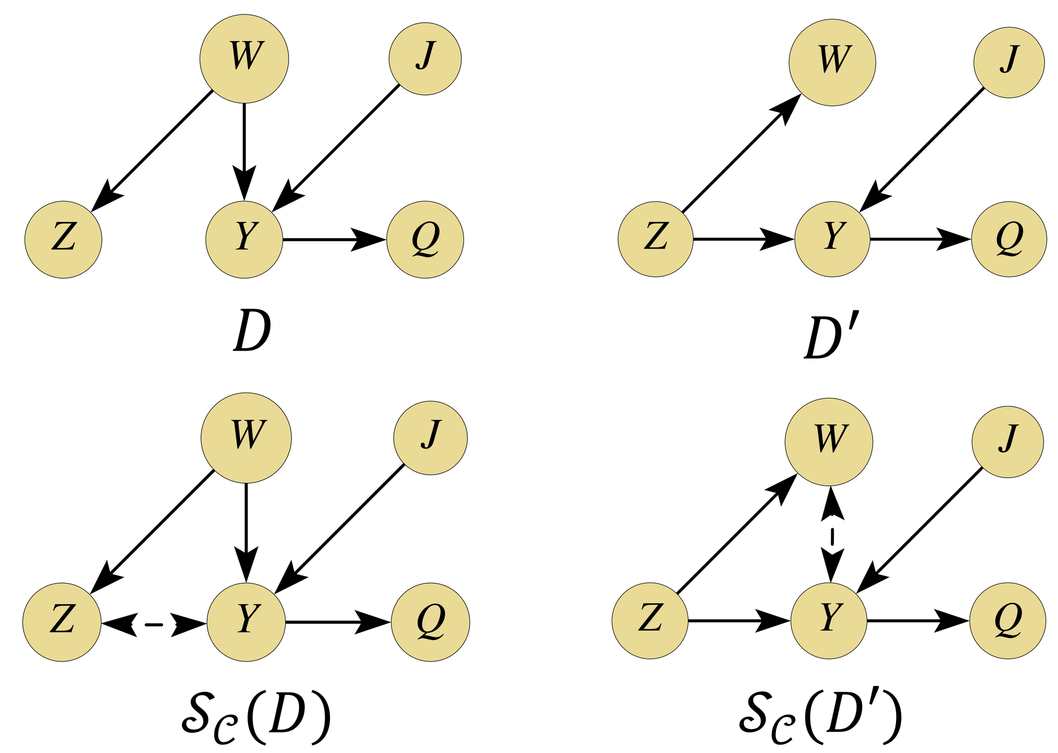
CONTACT INFORMATION

Web: muratkocaoglu.com/CausalML

Email: lee4094@purdue.edu

$\mathcal{C}$-COVERED and $\mathcal{C}$-CLOSURE GRAPHS

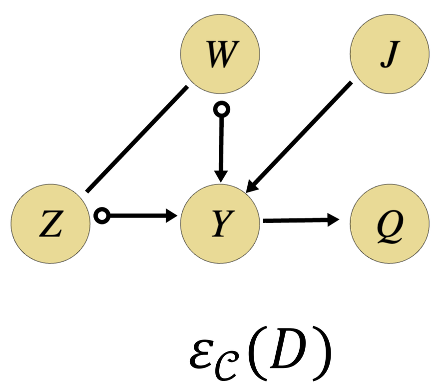
- **$\mathcal{C}$-covered:** In a DAG, X and Y are said to be **$\mathcal{C}$-covered** if there does not exist a member $\mathbf{C} \in \mathcal{C}$ s.t. $(X \perp\!\!\!\perp Y | \mathbf{Z})_D$
  - **Example**
    - Let $\mathcal{C} = \{\emptyset, \{Y\}\}$.
    - Z and Y are **$\mathcal{C}$-covered** in D.
    - Z and Q are **not $\mathcal{C}$-covered** in D.
- **$\mathcal{C}$-Closure graphs** of D, $\mathcal{S}_{\mathcal{C}}(D)$: **If** X and Y are **$\mathcal{C}$-covered:**
  (i) if $X \in An(Y)$ in D, then $X \to Y$ in $\mathcal{S}_{\mathcal{C}}(D)$
  (ii) if $X \notin An(Y)$ and $Y \notin An(X)$ in D, then $X \leftrightarrow Y$ in $\mathcal{S}_{\mathcal{C}}(D)$
  Else: X and Y **are not adjacent** in $\mathcal{S}_{\mathcal{C}}(D)$.



- **Lemma: $\mathcal{C}$-closure graphs** $\mathcal{S}_{\mathcal{C}}(D)$ **and DAG** D entail the same set of d-separation statements given any $\mathbf{C} \in \mathcal{C}$.
- **Theorem:** Two DAGs D, D' are **$\mathcal{C}$-Markov equivalent** if and only if $\mathcal{S}_{\mathcal{C}}(D), \mathcal{S}_{\mathcal{C}}(D')$ are Markov equivalent.

$\mathcal{C}$-ESSENTIAL GRAPHS

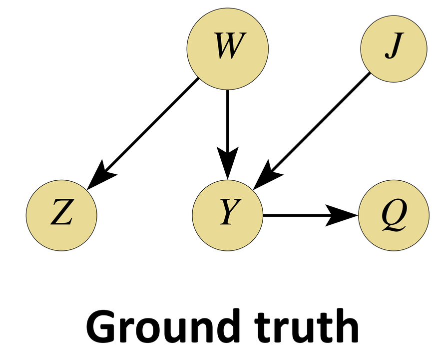
- **Characterizing the equivalence class of $\mathcal{C}$-closure graphs using edge union operation:**
  - $Xo - o\, Y := X \leftrightarrow Y \cup X \leftarrow Y \cup X \to Y$
  - $Xo \to Y := X \leftrightarrow Y \cup X \to Y$
  - $X - Y := X \leftarrow Y \cup X \to Y$
- **$\mathcal{C}$–essential graphs:** any DAG D, the edge union of all **$\mathcal{C}$-Closure graphs** that are Markov equivalent to $\mathcal{S}_{\mathcal{C}}(D)$ is called **$\mathcal{C}$–essential graph of** D, $\varepsilon_{\mathcal{C}}(D)$.

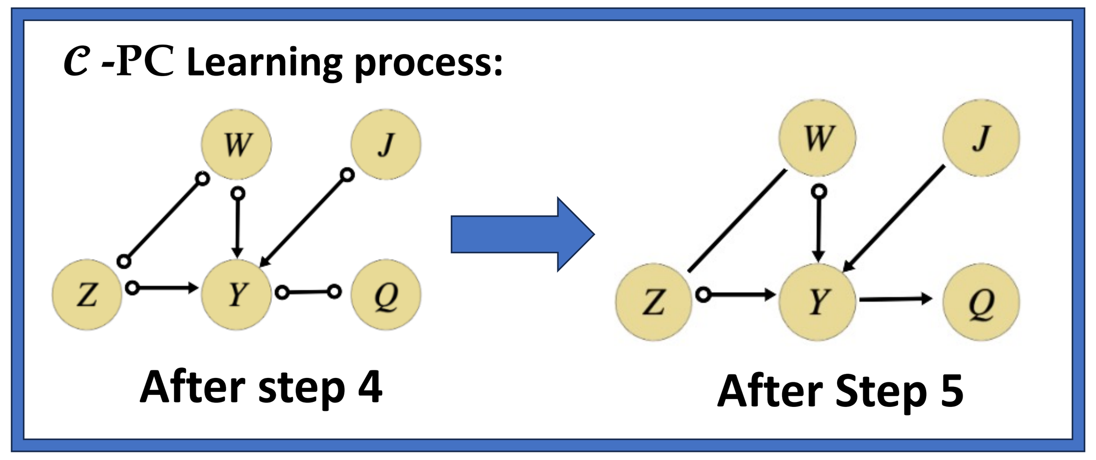


PROPOSED ALGORITHM: $\mathcal{C}$-PC

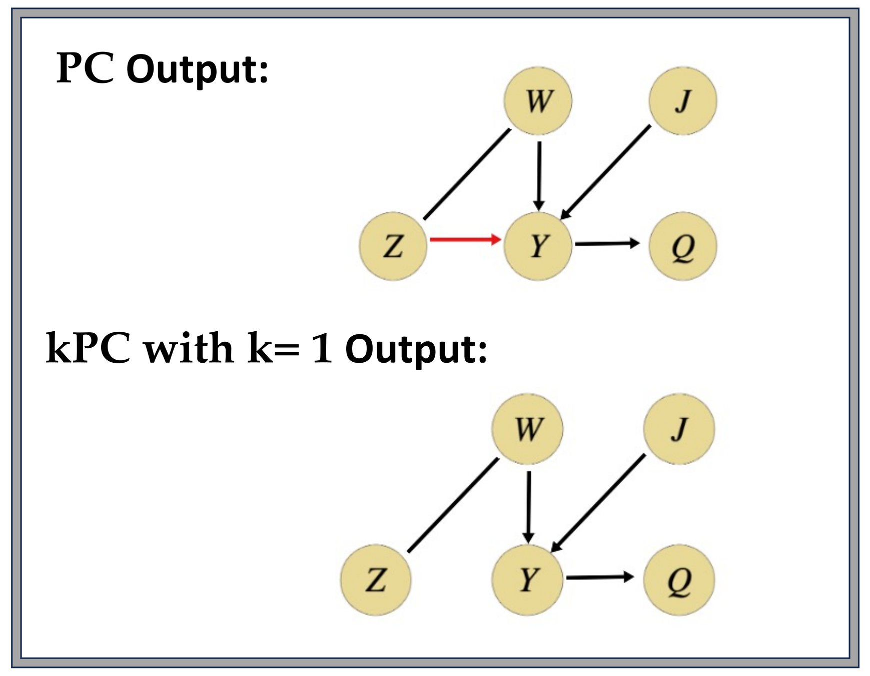
Input: observational data, a conditionally closed set $\mathcal{C}$, CI tester

1. Starts from a complete graph with circle edge o-o
2. Find separating sets $S_{X,Y}$ for every pair of variables X, Y by conditioning on $\mathbf{C} \in \mathcal{C}$.
3. Update M by removing the edges between pairs that are separable
4. Orient unshielded collider of M: For any induced subgraph $Xo - oZo - oY$, set $Xo \to Z \leftarrow oY$ for any non-adjacent pair X, Y where $S_{X,Y}$ does not contain Z
5. Apply FCI orientation rules **[Zhang '08]** and kPC orientation rules **[Kocaoglu '23]**





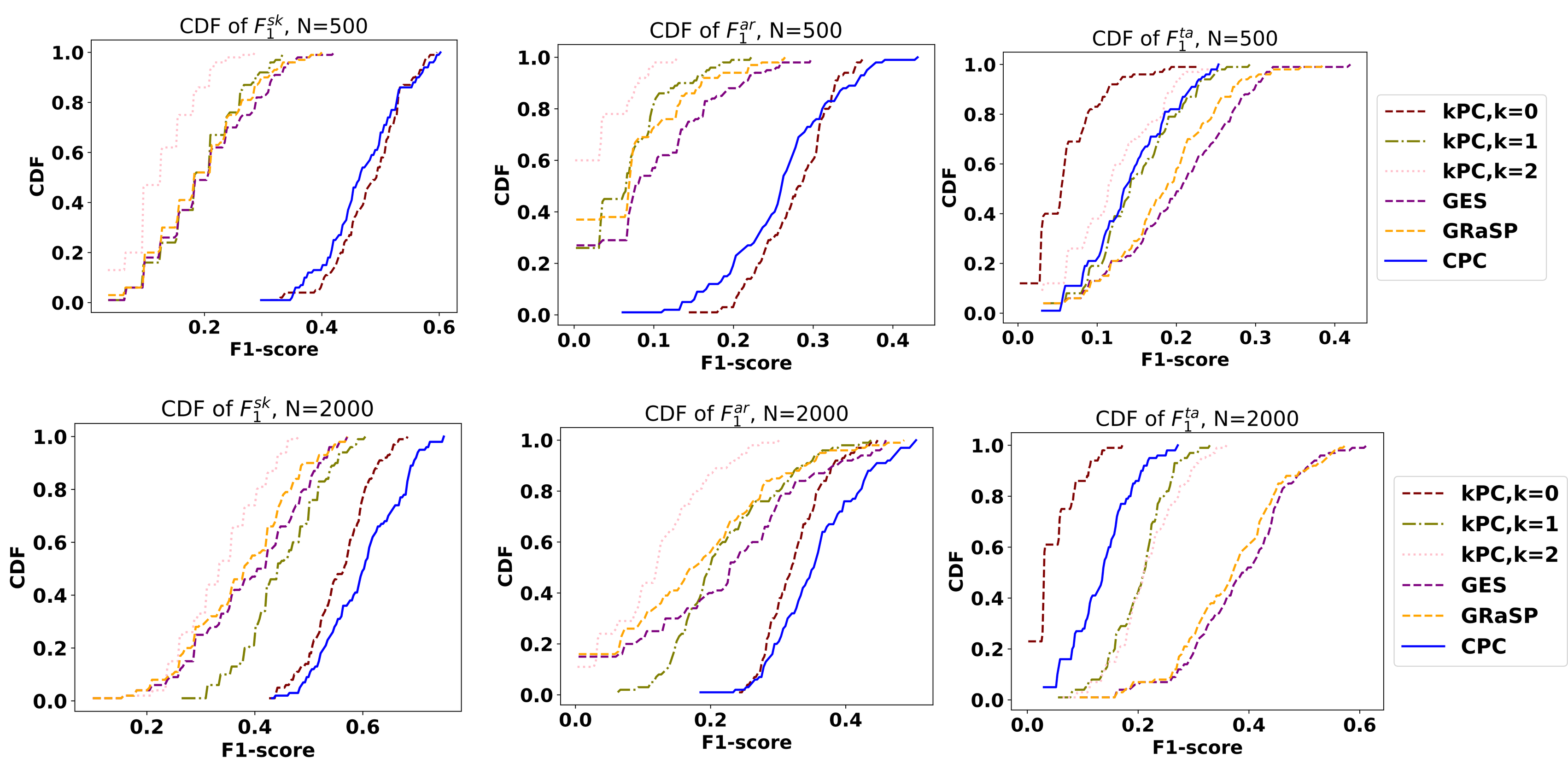
- **Theorem:** $\mathcal{C}$-PC algorithm is sound for learning $\mathcal{C}$-essential graph given a conditional independence oracle under causal Markov and $\mathcal{C}$-faithfulness assumptions



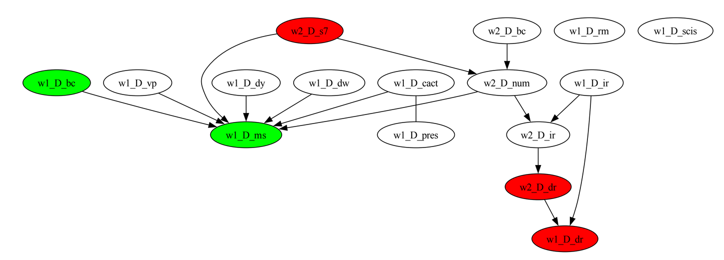
EXPERIMENTAL RESULTS

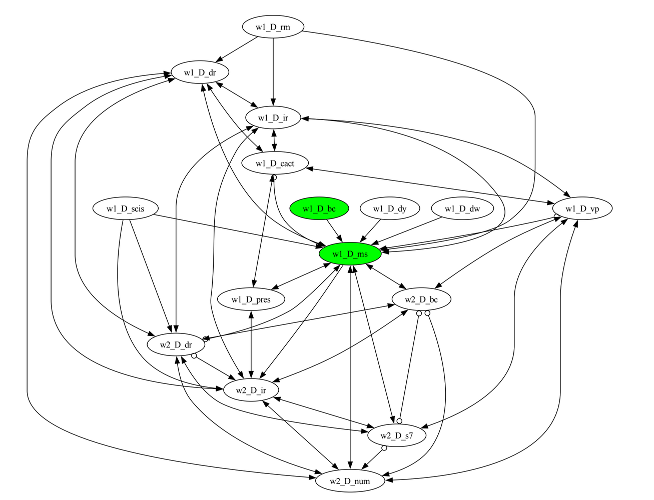
(**S:** Simulated , **R:** real-world)

- **S: Conditioning on high-dimensional variables (S)**
  - 100 random DAGs of size 30. Each has 2 or 30 states randomly assigned with 0.7 and 0.2 probabilities
  - Apply heuristic search to get conditioning sets that yield at least 5 samples per entry in the contingency table up to order 1.
  - F1-skeleton (left two); F1-arrowhead (right two); The closer to the bottom right the better



- **R: The Cognition and Aging USA (CogUSA) Study [McArdle et. al'07]**
  - 16 discrete variables 2 to 13 states), 8 variables have missing data. Incorrect **(red)** and correct (**green)** causal order.
  - $\mathcal{C}$-PC only condition on variables that do not have missing values.
  - Chisq tests with test-wise deletion.
  - $\mathcal{C}$-PC (left); MVPC **[Tu et.al '19]** (right)





CONCLUSION & FUTURE WORK

- We propose a sound algorithm called C-PC for learning causal graphs from a collection of conditioning sets known as conditionally closed sets. We extend an existing algorithm called k-PC that exhausts all CI tests of order up to some integer k to a setting where CI tests are restricted to a collection of conditioning sets.
- For future work, we want to further relax the restriction of a conditionally closed set and investigate whether one can systematically leverage arbitrary CI statements on top of all marginal independence relations for learning causal graphs.

REFERENCES

- J. Pearl, Causality: Models, Reasoning and Inference. Cambridge University Press, 2009.
- Tu, Ruibo, et al. "Causal discovery in the presence of missing data." The 22nd International Conference on Artificial Intelligence and Statistics. Pmlr, 2019.
- McArdle John, R. W. and Robert, W. Cognition and aging in the usa (cogusa), 2007-2009.
- Kocaoglu, M. Characterization and learning of causal graphs with small conditioning sets.415 Advances in Neural Information Processing Systems, 36, 2024
- Wienöbst, Marcel, and Maciej Liskiewicz. "Recovering causal structures from low-order conditional independencies." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. No. 06. 2020.
- Zhang, Jiji. "On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias." Artificial Intelligence 172.16-17 (2008): 1873-1896.