Localised Natural Causal Learning Algorithms under Weak Conditions

Kai Teh, Kayvan Sadeghi, Terry Soo
Department of Statistical Science, University College London

UCL

Overview

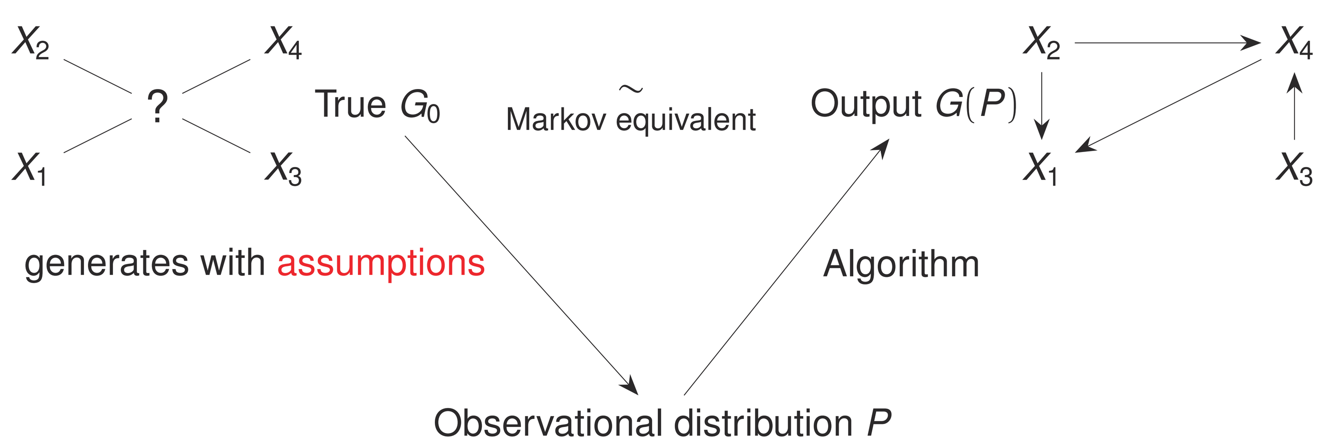
Overview Provide an algorithm (Me-LoNS) to perform Causal Discovery/Learning under reasonably weak assumptions.

Keywords Causal Discovery, Graphical Models, Faithfulness Assumption

Problem Setting

Our setting for Causal Discovery/Learning:

1. Purely *observational*, not accounting for interventional data.
2. Constraint-based, assume access to a *conditional independence oracle*.



Under some assumptions, the output $G(P)$ is the 'same' as the true causal graph. The most common is the *faithfulness* assumption.

Graphical separations in G_0 $\iff$ Conditional independencies in P

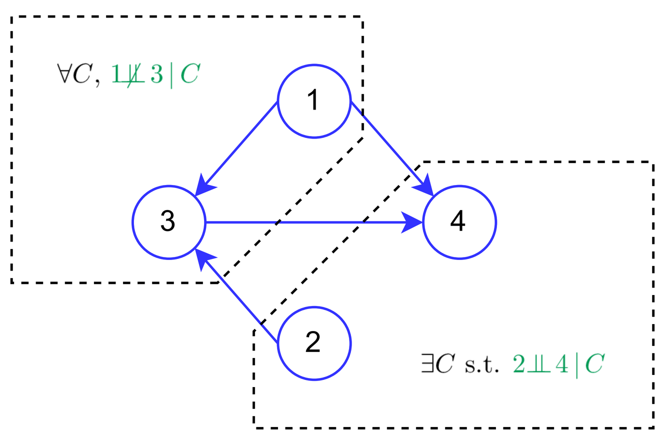
Too strong! May sometimes not hold empirically :(

Our assumptions

Assumption relates observational distribution P and true causal graph G_0.

1. P is *adjacency faithful* to G_0.

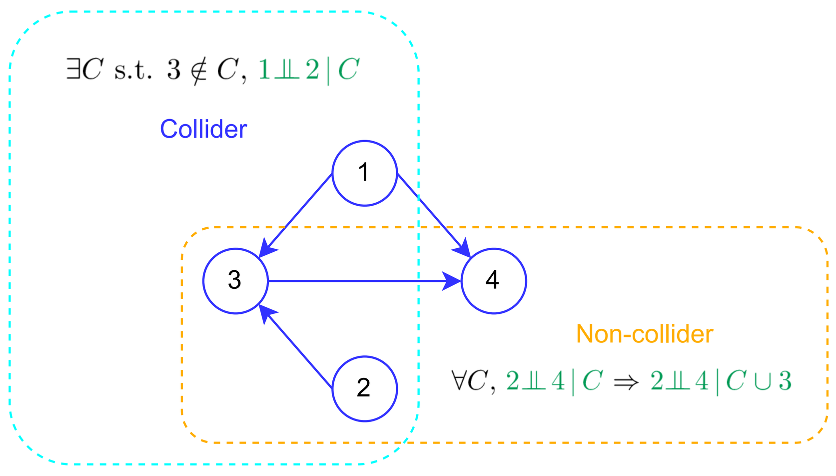
   $i - j \in G_0 \iff i \perp\!\!\!\perp j \mid C$ for some C



2. P is v-ordered upward stable (V-OUS) and collider-stable to G_0.

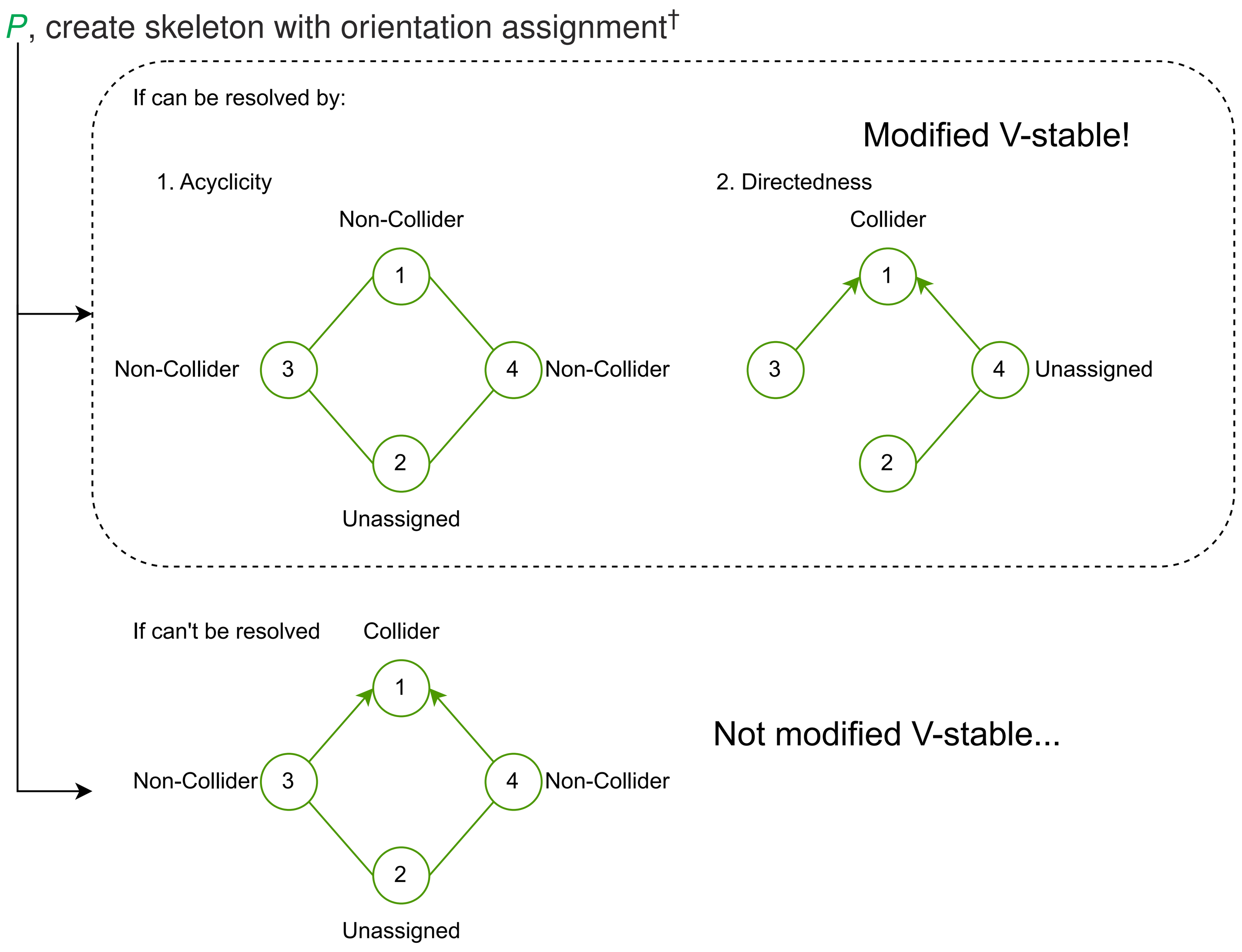
   $i \to k \leftarrow j \in G_0 \Rightarrow i \perp\!\!\!\perp j \mid C$ for some C s.t. $k \notin C$

   non-collider $i \sim k \sim j \in G_0 \Rightarrow (i \perp\!\!\!\perp j \mid C \Rightarrow i \perp\!\!\!\perp j \mid C \cup k$ for all $C)$



3. P is modified V-stable.

   From P, create skeleton with orientation assignment†



Modified V-stable Localised Natural Structure Learning (Me-LoNS)

A modification of the (C)PC algorithm via replacing the oreintation and propagation step with the following.

- Orientation step replaced with†:

$$i - k - j \;\begin{cases} \xrightarrow{\exists C \quad (i \perp\!\!\!\perp j \mid C,\ i \not\perp\!\!\!\perp j \mid C \cup k)} & \text{collider} \\ \xrightarrow{\forall C \quad (i \perp\!\!\!\perp j \mid C \Rightarrow k \in C)} & \text{non-collider} \\ \xrightarrow{\text{otherwise}} & \text{unassigned} \end{cases}$$

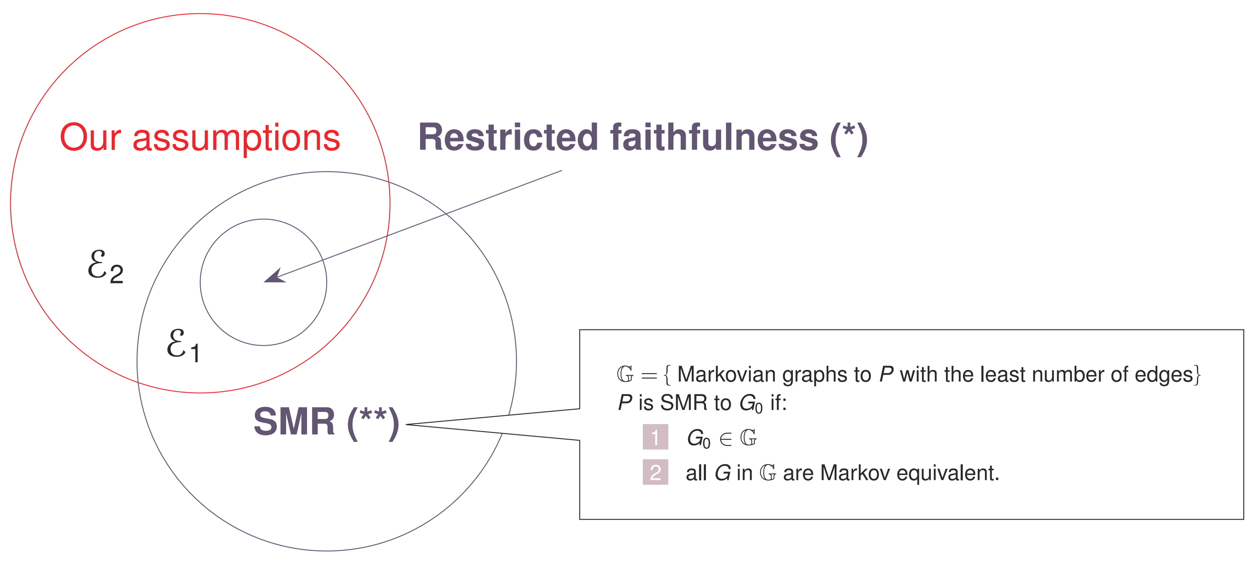
- Propagation step replaced with Mixed Integer Linear Programming to solve for DAG

Theorem

From observational distribution P and true causal graph G_0:

Me-LoNS return G_0 (up to MEC)
$\Updownarrow$
P satisfies our assumptions with G_0.

How do our assumptions compare with **existing** causal discovery assumptions?



An alternative to some existing causal discovery approaches! :)

How reasonable are our assumptions?

How strong is:

- Collider-stability?

  Pairwise Markov property $\Rightarrow$ Collider stability

- V-OUS?

  1. Composition (e.g. Gaussians)
  2. Conditional exchangability on non-colliders

  $\Rightarrow$ V-OUS

  Note:

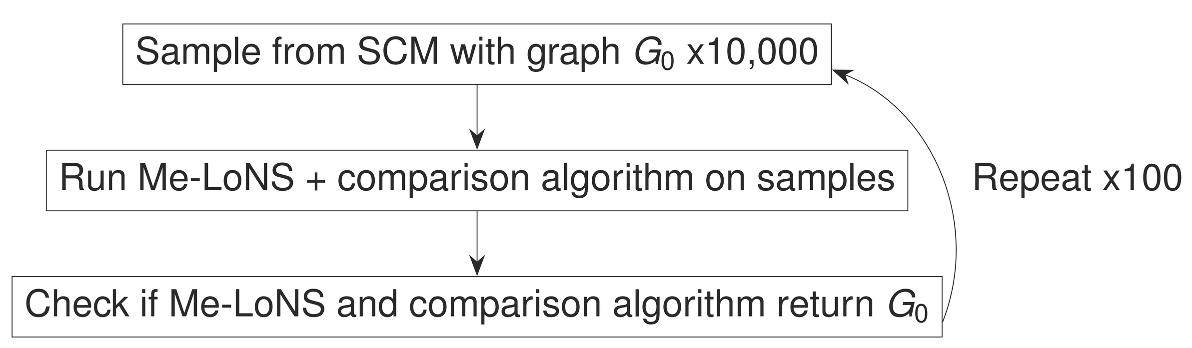
  Conditional exchangability here refers to the exchangability of the conditional distribution

- Modiifed V-stability?

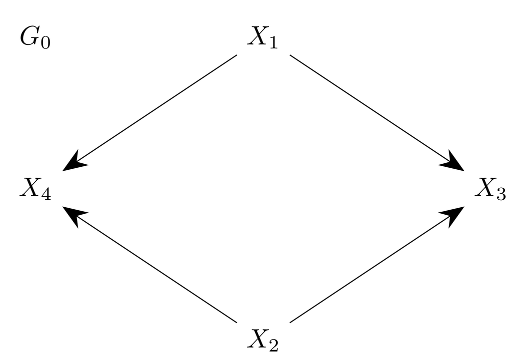
  Singleton-transitivity (e.g. Gaussians) $\Rightarrow$ Modified V-stability

Weaker than common assumptions! :)

Simulation Comparisons



(C)PC (*) comparison

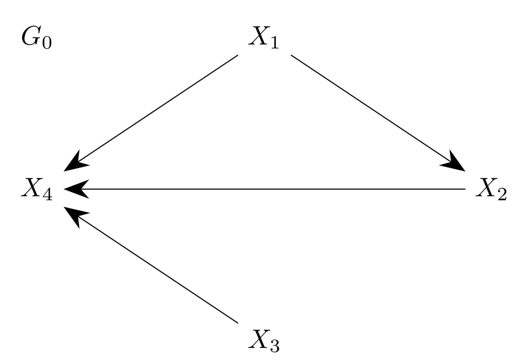


$$\epsilon_i \overset{\text{i.i.d.}}{\sim} N(0,1),\ i = 1,2,3,4$$
$$X_1 = \epsilon_1$$
$$X_2 = \epsilon_2$$
$$X_3 = -6X_1 + 2X_2 + \epsilon_3$$
$$X_4 = 3X_1 + 4X_2 + \epsilon_4$$

SCM $\mathcal{E}_1$

PC	Me-LoNS
8%	90%

SP (**) comparison



$$\epsilon_i, \phi_j \overset{\text{i.i.d.}}{\sim} Bern(\tfrac{1}{2}),\ i = 1,..,4,\ j = 1,...,5$$
$$X_1 = (\phi_1, \phi_2, \epsilon_1)$$
$$X_2 = (X_1^1, \phi_3, \epsilon_2)$$
$$X_3 = (\phi_4, \phi_5, \epsilon_3)$$
$$X_4 = (X_1^1 + X_3^1, X_2^1 + X_2^2, X_2^2, \epsilon_4)$$

SCM $\mathcal{E}_2$

GRaSP	Me-LoNS
56%	94%

Scan for paper!

kai.teh.21@ucl.ac.uk