

NeurIPS 2026

Coarsening Linear Non-Gaussian Causal Models with Cycles

Francisco Madaleno

Technical University of Denmark

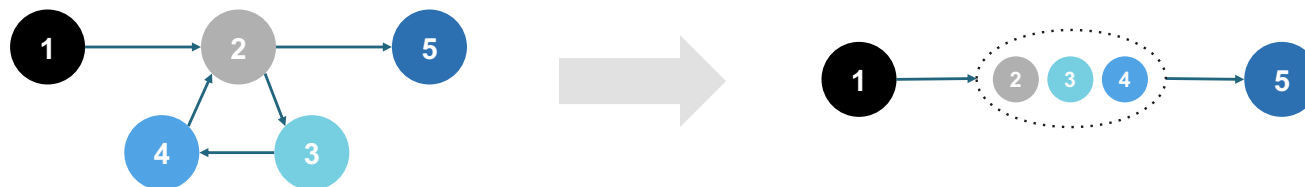
fmfsa@dtu.dk

Francisco C. Pereira

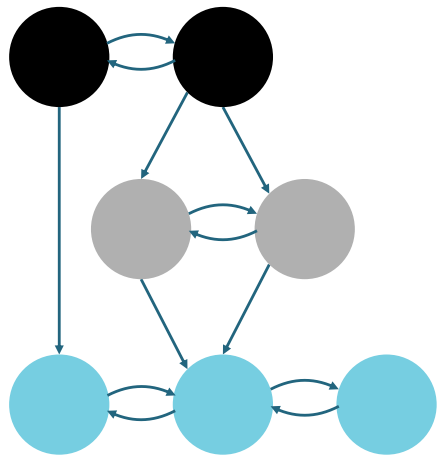
Technical University of Denmark

Alex Markham

University of Copenhagen

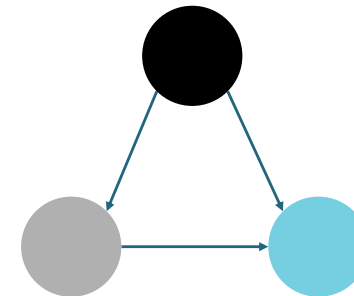


Motivation



High-dimensional cyclic graph G (with feedback loops)

Systems Biology signaling feedback loops $\rightarrow$ pathway states



Low-dimensional condensation DAG G^{sc}

Dynamical Systems coupled equilibrium vars $\rightarrow$ macro states

Which variables can be grouped, and what causal DAG structure remains?

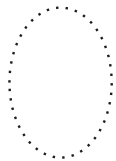
Notation



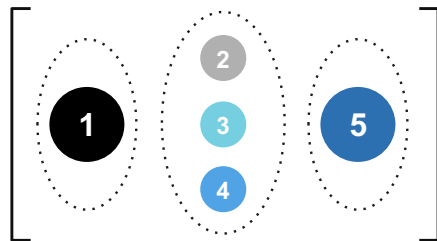
Variable, Node



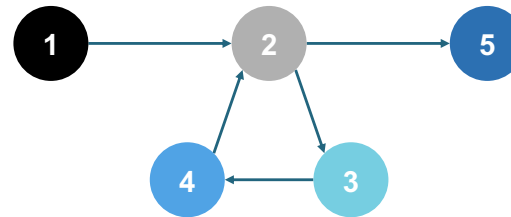
Edge (may form directed cycles)



Cluster, Part, Coarse Node (SCC)



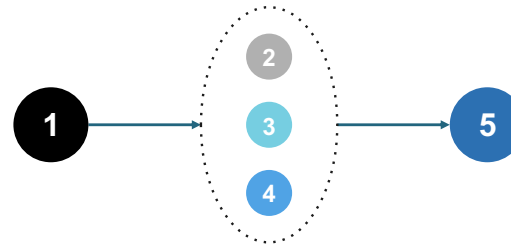
Partition, Coarsening Π



Cyclic Graph G

$$X = BX + \varepsilon$$

$$W = I - B$$



Cluster-DAG, C-DAG,
Condensation G^{sc}

What is a DAG-Coarsening?

For directed graph $G = (V, E)$, a DAG-coarsening is a DAG $G' = (\Pi, E')$ with $\chi : V \rightarrow \Pi$:

$$E' = \{ \chi(u) \rightarrow \chi(v) \mid u \rightarrow v \in E, \chi(u) \neq \chi(v) \}$$

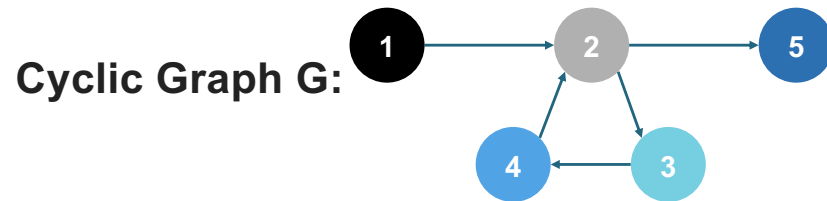
Rules out:

- Splitting any Strongly Connected Component (SCC)
- Merging across layers that creates cycles in G'

Preserves:

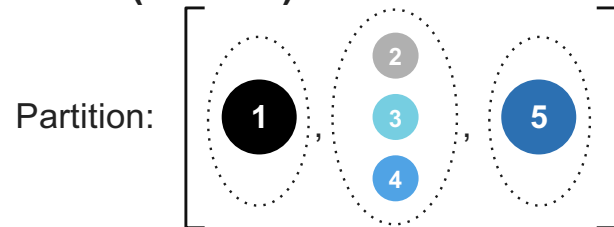
- Causal direction / partial order between groups
- *Markov property (Lemma 1):* $\sigma\text{-Markov}(G) \Rightarrow \text{Markov}(G^{\text{sc}})$

Valid vs Invalid Coarsenings



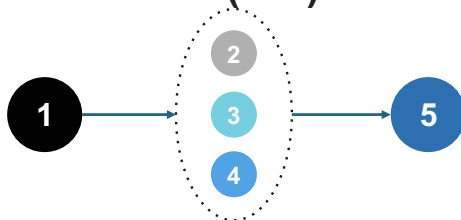
(3-cycle SCC {2,3,4}; Example 1 of Lacerda et al. [2008])

Valid (4 of 52):



$$\chi(2) = \chi(3) = \chi(4) = \{2, 3, 4\}$$

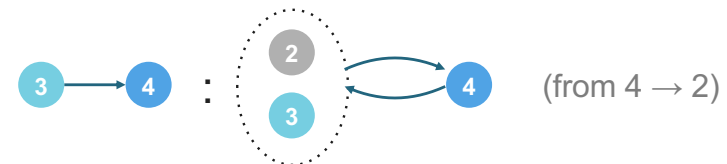
Coarsened (G^{sc}):



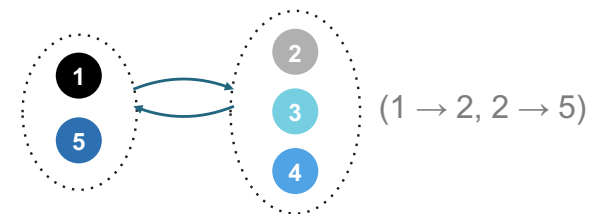
Absorbs cycle into SCC. Quotient is a DAG ✓

Invalid (48 of 52):

1. Split an SCC: $\chi(2) = \chi(3) = \{2, 3\}$, $\chi(4) = \{4\}$



2. Merge across layers: $\chi(1) = \chi(5) = \{1, 5\}$



Not a DAG ✗

The Coarsening Lattice & The SCC Floor

The DAG-coarsenings of any directed graph form a sublattice with a structural floor (Proposition 1).

- **The SCC Floor (Π^*):**

nodes in a common SCC must stay grouped:

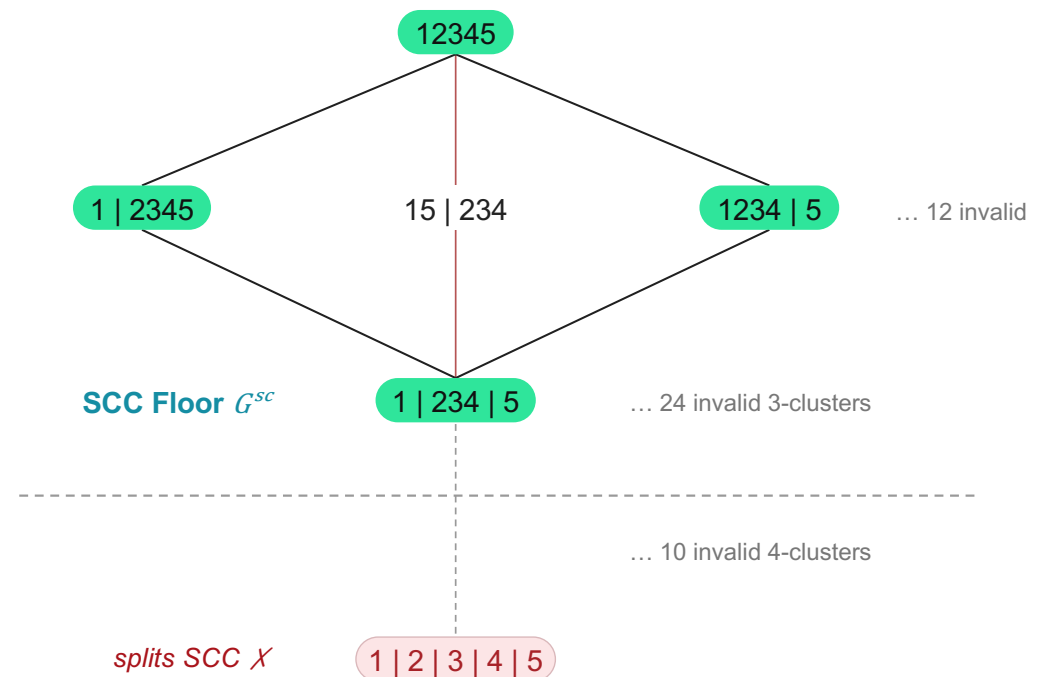
$$\chi(v) = \chi(w) \Leftrightarrow \text{sc}_G(v) = \text{sc}_G(w)$$

- **Condensation (G^{sc}):**

finest valid DAG-coarsening (least element of the sublattice); computable in $\mathcal{O}(d + |E|)$ via Tarjan

- **Resolves Cyclic Ambiguity:**

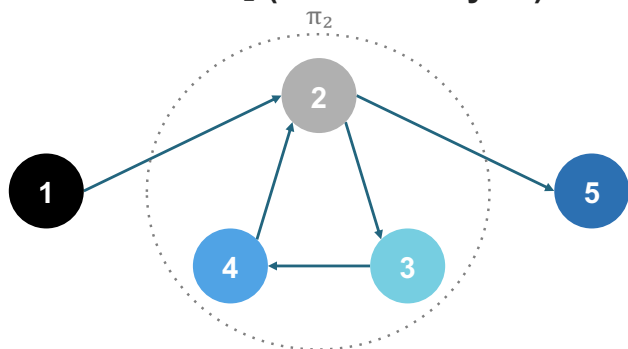
every refinement below the floor splits a cycle



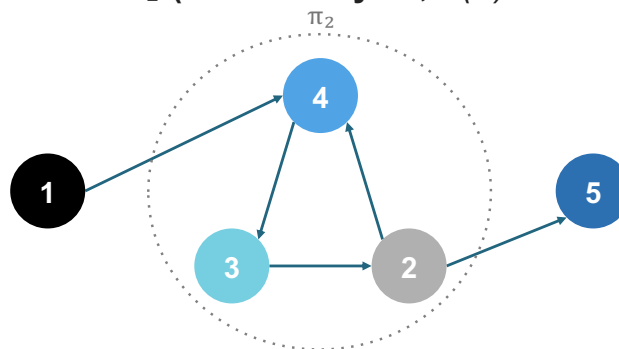
Condensation Identifiability (Theorem 1)

ICA identifies $W = I - B$ up to row permutation P : $\widehat{B}_p = I - \text{diag}(P\widehat{W})^{-1}P\widehat{W}$

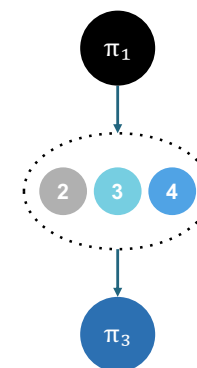
Member G_1 (clockwise cycle)



Member G_2 (reversed cycle, $P(X)$ identical)



Condensation G^{sc}



Theorem 1 (Condensation Identifiability):

If two linear non-Gaussian cyclic SCMs G_1, G_2 (no latent confounders, ≤ 1 Gaussian noise) induce the same observational distribution $P(X)$, then $G_1^{sc} = G_2^{sc}$ (same SCCs, same inter-SCC edges).

Lemma 2: Every admissible permutation maps each SCC onto itself and preserves inter-SCC edges

Identifiability & Causal Effects

Under our Assumptions, the condensation G^{sc} is identifiable from observational data alone.

Assumptions:

1. Linear Cyclic SCM

$$X = BX + \varepsilon \text{ with } \det(I - B) \neq 0$$

2. Independent Non-Gaussian Noise

mutually independent ε_i , at most one Gaussian component

3. Causal Sufficiency

no unobserved confounders between variables

Strictly weaker than prior cyclic LiNG works:

- No stability assumption ($\rho(B) < 1$) needed
- No disjoint-cycles assumption needed
- No faithfulness assumption needed

Downstream Payoff (Lemma 1 & §4.3):

- σ -Markov on $G \Rightarrow d$ -sep Markov on G^{sc}
- Whole-SCC hard interventions $\text{do}(X_\pi = c)$ identifiable via cluster-level g -formula

Algorithm 1: Condensation Recovery

Stage:

Practical instantiation:

Unmixing: Estimate demixing matrix $W = I - B$

➡ **FastICA**(X, d) to obtain $\widehat{W}$

Assignment: Pick one admissible permutation P

➡ **Hungarian**($\widehat{W}$) maximizing $\sum_i \log |(\widehat{PW})_{ii}|$

Thresholding: Normalize & prune noise below τ

➡ $\widehat{B} \leftarrow I - \text{diag}(\widehat{PW})^{-1} \widehat{PW}$, zero $|\widehat{B}_{ij}| < \tau$

Condensation: Extract SCCs & inter-SCC edges

➡ **Tarjan**($\text{supp}(\widehat{B})^T$) $\Rightarrow \widehat{G}^{\text{sc}} = (\widehat{\Pi}, \widehat{E}')$

Consistent (finite-sample error $\Pr[\widehat{G}_n^{\text{sc}} \neq G^{\text{sc}}] \leq \mathcal{O}(n^{-2})$, Prop. 3)

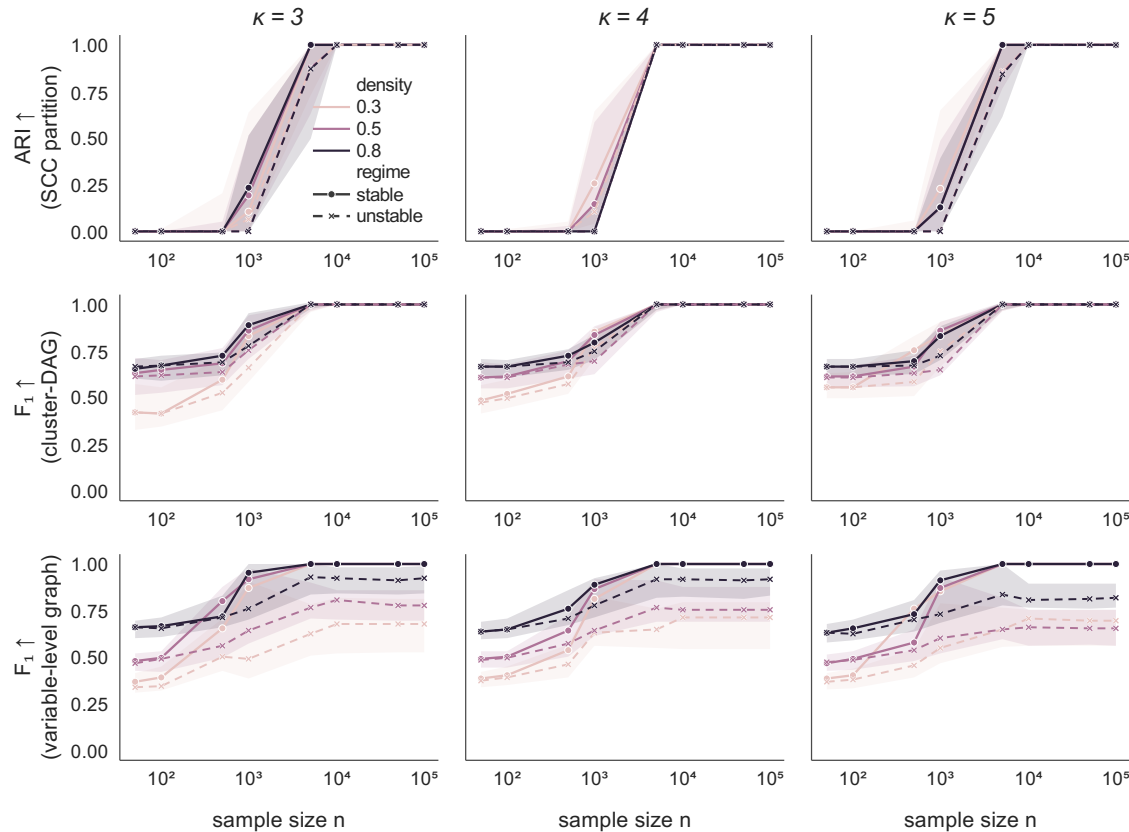
d : number of nodes n : number of samples

k : FastICA iterations τ : edge threshold

Complexity: $\mathcal{O}(k \cdot n \cdot d^2 + d^3)$ (bypasses $\mathcal{O}(d!)$ equivalence-class enumeration)

Synthetic Results

Linear Laplace cyclic SCMs, $d = 10$, SCCs $\kappa \in \{3, 4, 5\}$, density $\lambda \in \{0.3, 0.5, 0.8\}$, stable ($\rho \approx 0.9$) vs. unstable ($\rho = 1.5$)



Condensation G^{sc} (Rows 1–2)

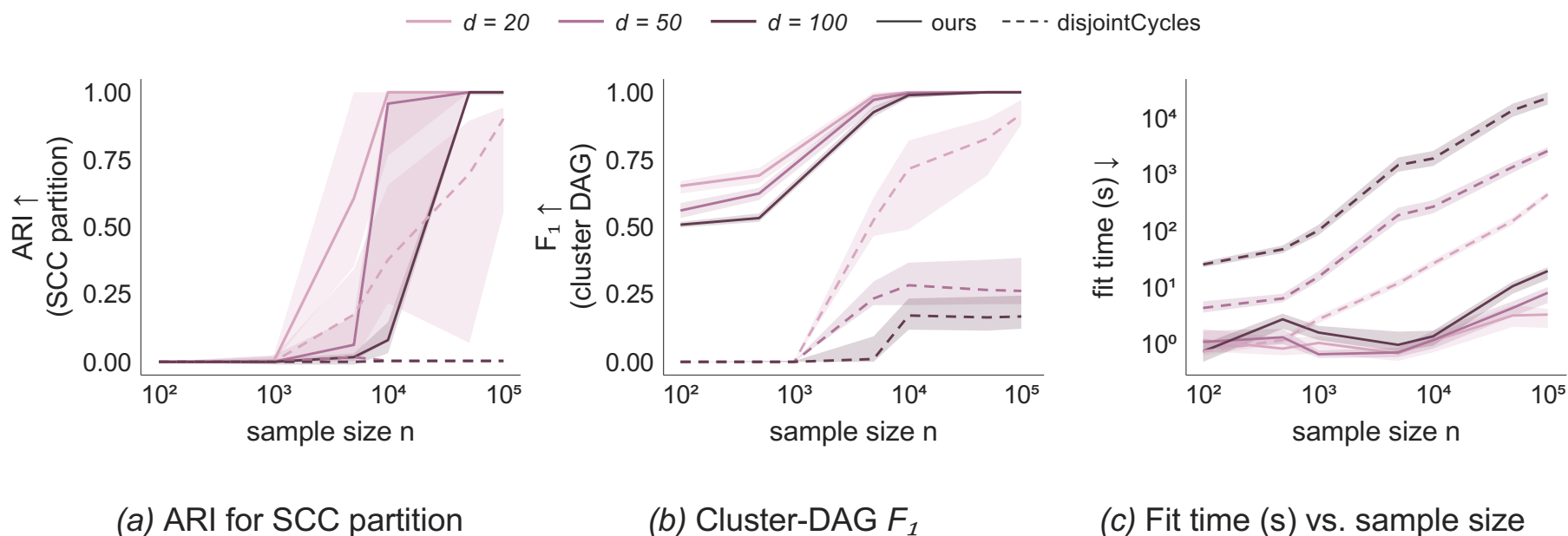
- **Partition ARI** $\rightarrow 1$ across all κ, λ
- **Cluster-DAG F_1** $\rightarrow 1$ in both stable (solid) and unstable (dashed) regimes ✓

Micro-Graph G (Row 3)

- Unstable regime ($\rho = 1.5$, dashed) **saturates below 1.0**
- Intra-SCC cycle orientation is unidentifiable without stability

Scalability

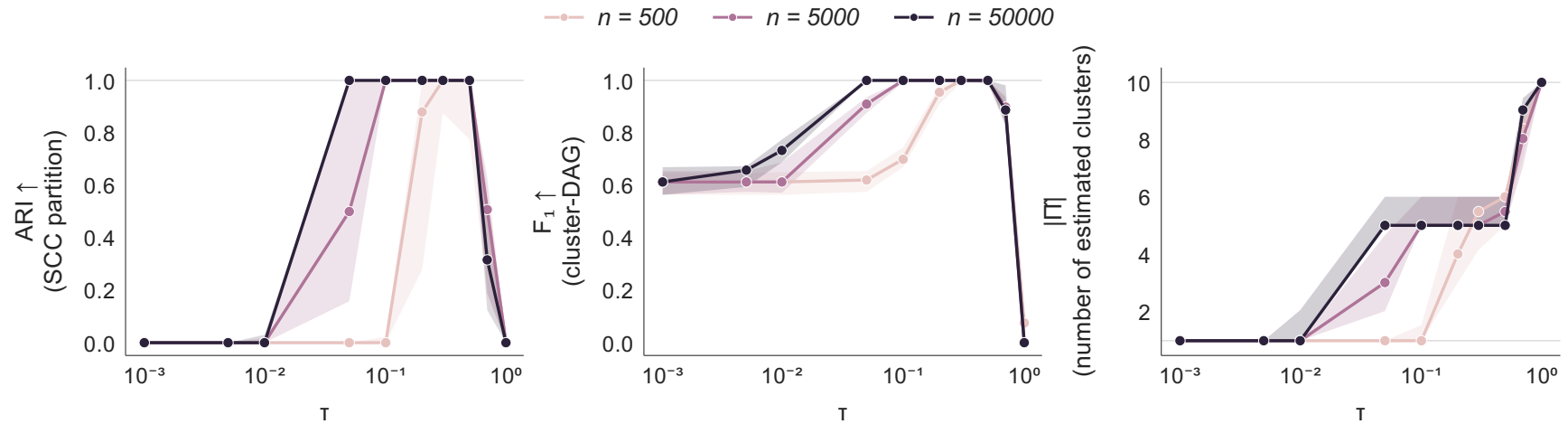
$d \in \{20, 50, 100\}$, density $\lambda = 0.5$, $\kappa = 10$ SCCs, 10 seeds — Ours (solid) vs. disjointCycles [Drton et al., 2025] (dashed)



Accurate with overlapping cycles ($d \geq 50$) where disjoint-cycle methods plateau — and faster (10^1 s vs. 10^4 s)

Threshold Sensitivity & Failure Modes

Proposition 2 & 3: $\Pr[\hat{G}_n^{\text{sc}} \neq G^{\text{sc}}] \leq \frac{K_1}{n^2(\beta_{\min} - \tau)^4}$ (splitting) + $\frac{K_2}{n^2\tau^4}$ (merging), $d = 10$, $\kappa = 4$, $\lambda = 0.5$



$\tau \rightarrow 0$ (Too Small)

Spurious edges close cross-SCC paths $\Rightarrow$ **merges SCCs** ($|\Pi| = 1$)

Safe Band ($\tau \in [0.1, 0.5]$)

Exact recovery ($\text{ARI} = 1$, $F_1 = 1$);
error-free band widens as $\mathcal{O}(n^{-2})$

$\tau \rightarrow 1$ (Too Large)

True intra-SCC edges dropped
 $\Rightarrow$ **splits SCCs** ($|\Pi| = d = 10$)

Ablations & Causal Effect Estimation

1. Permutation Selection Ablation ($d = 20, \kappa = 5, n = 5 \cdot 10^4$)

Selection Rule	Stable	Unstable	Selection Time
Enum. + First-Stable [Lacerda 08]	100%	96.6%	6.08 s
Enum. + Uniform Admissible	98.0%	98.6%	6.08 s
Ours (Hungarian Assignment)	100%	100%	0.04 ms
Random Permutation (Control)	0%	0%	0.03 ms

2. Comparison with GroupLiNGAM ($d = 10, \kappa = 4, n = 5000$)

- **Ours:** ARI = 1.00 (stable & unstable), fit time < 1 s
- **GroupLiNGAM:** ARI = 0.84 (stable) / 0.69 (unstable), fit time $\approx$ 11 hour

3. Whole-SCC Hard Interventions (§ 4.3)

Hard intervention $\text{do}(X_s = c)$ on a target SCC S is identified on G^{sc} via the cluster-level g-formula:

$$p(x_{V \setminus S} \mid \text{do}(X_S = c)) = \prod_{\pi \in \Pi_S} p(x_\pi \mid x_{\text{pa}(\pi)})$$

Regime (40 SCMs)	Err (n = 100)	Err ($5 \cdot 10^4$)	Oracle Match
Stable ($\rho \approx 0.9$)	0.289	0.013	100%
Unstable ($\rho = 1.5$)	0.219	0.008	100%

Requires **no intra-SCC edge orientation**:
reduces cyclic cluster interventions to acyclic C-DAGs.

Any admissible permutation recovers G^{sc} — Hungarian avoids $\mathcal{O}(d!)$ enumeration

Broader Connections

The cyclic coarsening framework connects and unifies key paradigms in causal discovery:

SCC Condensation

$$\chi(v) = \chi(w) \iff \text{sc}_G(v) = \text{sc}_G(w)$$

*structural floor of the DAG-coarsening lattice;
invariant across the cyclic LiNG equivalence class*

Block Ordering

Finest exogenous block partition [GroupLiNGAM, Kawahara et al. 2010]

coincides with G^{sc} under sufficiency, but requires faithfulness and $O(2^d)$ subset search

Interventional & C-DAG

$$\chi(v) = \chi(w) \iff \text{l-an}(v) = \text{l-an}(w) \text{ [Madaleno et al. 2026; Anand et al. 2023]}$$

Summary

Framework

DAG-coarsening for directed graphs with cycles

SCC partition is the structural floor of the coarsening lattice

Theory

condensation G^{sc} is invariant across the LiNG equivalence class

identifiable from observations without stability or disjoint cycles

Algorithm

FastICA + Hungarian + Tarjan: $O(knd^2 + d^3)$ time,

enables whole-SCC effect estimation via cluster g-formula



github.com/fmfsa/coarsening-cycles

Thank you!

fmfsa@dtu.dk