Generalized Independent Noise Condition for Estimating Latent Variable Causal Graphs (NeurIPS 2020)

Feng Xie*[1,2], Ruichu Cai*[1,3], Biwei Huang[4], Clark Glymour[4], Zhifeng Hao[1,5], Kun Zhang*[4]

[1]Guangdong University of Technology, [2]Peking University, [3]Pazhou Lab

[4]Carnegie Mellon University, [5]Foshan University

Speaker: Feng Xie

Outline

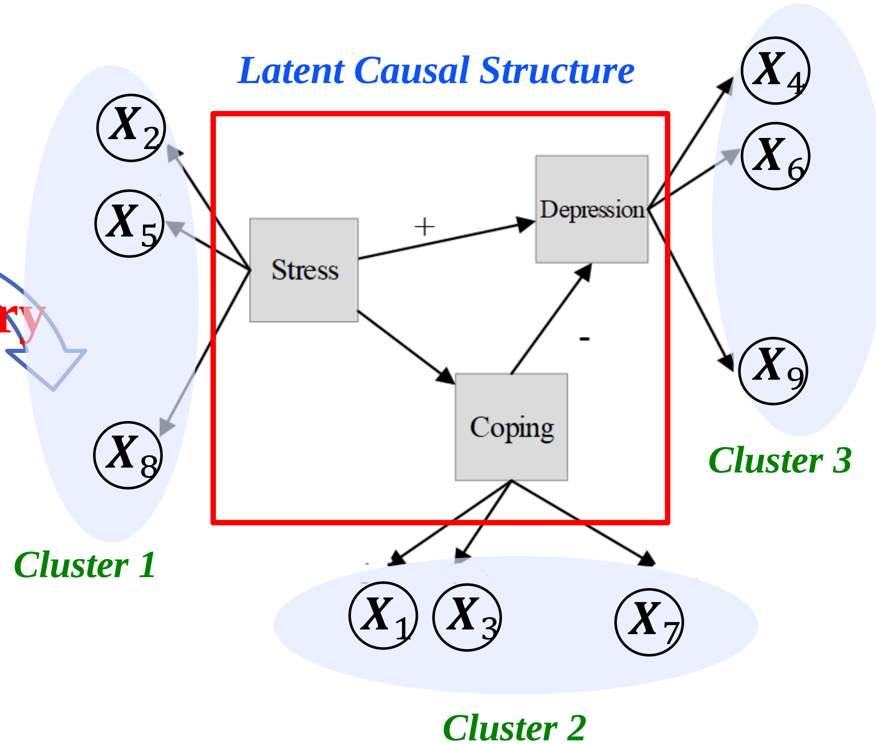
- Background and Related works
- IN Condition to GIN Condition
- Estimating LiNGLaM Based on GIN
- Experiments and Application
- Conclusion and Further work

Background

Observational dataset

$\mathbf{X}$					
	$\boldsymbol{X}_1$	X_{11}	X_{12}	$\ldots$	X_{1n}
	$\boldsymbol{X}_2$	X_{21}	X_{22}	$\ldots$	X_{2n}
	$\boldsymbol{X}_3$	X_{31}	X_{32}	$\ldots$	X_{3n}
	$\boldsymbol{X}_4$	X_{41}	X_{42}	$\ldots$	[illegible]
	$\boldsymbol{X}_5$	X_{51}	X_{52}	$\ldots$	X_{5n}
	$\boldsymbol{X}_6$	X_{61}	X_{62}	$\ldots$	X_{6n}
	$\boldsymbol{X}_7$	X_{71}	X_{72}	$\ldots$	X_{7n}
	$\boldsymbol{X}_8$	X_{81}	X_{82}	$\ldots$	X_{8n}
	$\boldsymbol{X}_9$	X_{91}	X_{92}	$\ldots$	X_{9n}
		1	2	$\ldots$	n

Causal Discovery



Stress, **Depression**, and **Coping** are latent variables.

Open problem:

It is hard to find latent variables and their underlying causal structure from observarional dataset X.

Related Works

Causal Discovery Methods On Latent Variables

- The classic constraint-based methods, such as, FCI (Spirtes, et al., UAI'1995), and RFCI (Colombo, et al., Ann Stat'2012).
- The Over-complete ICA methods, such as, LvLiNGAM (Hoyer et al., IJAR'2008).
- Tetrad constraints methods, such as, BPC (Silva et al., JMLR'2006), and FOFC (Kummerfeld and Ramsey, KDD'2016).
- Non-Gaussian-based methods, such as, noisy ICA(Shimizu et al., NC'2009), ECA (Anandkumar et al., ICML' 2013), and Triad condition (Cai et al, NeurIPS'2019).

☑ **Limitations**:

- The first two strategies do not focus on the structure of latent variables;
- The third strategy needs more pure measurement variables and output Markov equivalence classes;
- The last one, some extract only second-order statistics in identifying latent factors, and some do not apply to the case where there are multiple latent variables behind two observed variables.

Outline

- Background and Related works
- IN Condition to GIN Condition
- Estimating LiNGLaM Based on GIN
- Experiments and Application
- Conclusion and Further work

IN Condition (Shimizu et al., JMLR'11)

$$\boldsymbol{Z} \longrightarrow Y$$

DEF[Independent Noise(IN), condition] $(\boldsymbol{Z}, Y)$ follows the IN condition iff the residual of regressing Y on $\boldsymbol{Z}$, $Y - \tilde{\omega}^{\mathrm{T}}\boldsymbol{Z}$, is independent from $\boldsymbol{Z}$.

➢ **Graphical criterion**

Let $\boldsymbol{Z}$ and Y be variables in a Linear, Non-Gaussian Acyclic Model (LiNGAM). $\boldsymbol{Z}$ and Y satisfies the IN condition iff

- All variables in $\boldsymbol{Z}$ are causally earlier than Y, and
- the common cause for Y and each variable in $\boldsymbol{Z}$, if there is any, is in $\boldsymbol{Z}$.

IN Condition (Shimizu et al., JMLR'11)

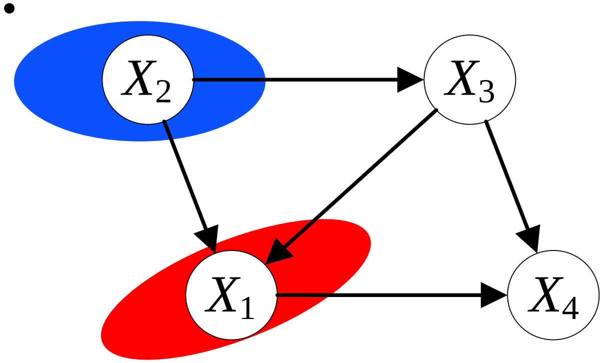
$$\boldsymbol{Z} \longrightarrow Y$$

DEF[Independent Noise(IN), condition] $(\boldsymbol{Z}, Y)$ follows the IN condition iff the residual of regressing Y on $\boldsymbol{Z}$, $Y - \tilde{\omega}^{\mathrm{T}}\boldsymbol{Z}$, is independent from $\boldsymbol{Z}$.

➢ **Graphical criterion**

Let $\boldsymbol{Z}$ and Y be variables in a Linear, Non-Gaussian Acyclic Model (LiNGAM). $\boldsymbol{Z}$ and Y satisfies the IN condition iff

- All variables in $\boldsymbol{Z}$ are causally earlier than Y, and
- the common cause for Y and each variable in $\boldsymbol{Z}$, if there is any, is in $\boldsymbol{Z}$.



Let $\boldsymbol{Z} = \{X_2\}$ and $Y = \{X_1\}$.

- $(\boldsymbol{Z}, Y)$ follows the IN condition.
- $\{X_2\}$ is causally earlier than $\{X_1\}$ and their common cause ($\{X_2\}$) are both in $\boldsymbol{Z}$.

IN Condition (Shimizu et al., JMLR'11)

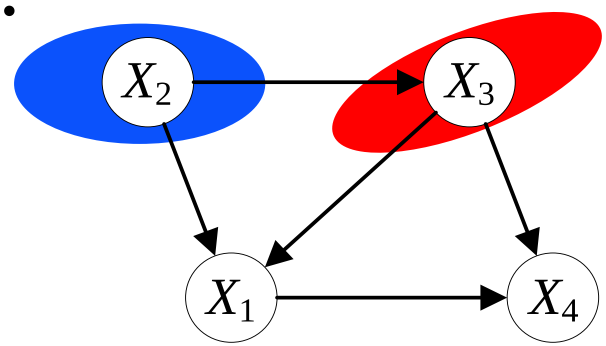
$$\boldsymbol{Z} \longrightarrow Y$$

DEF[Independent Noise(IN), condition] $(\boldsymbol{Z}, Y)$ follows the IN condition iff the residual of regressing Y on $\boldsymbol{Z}$, $Y - \tilde{\omega}^{\mathrm{T}}\boldsymbol{Z}$, is independent from $\boldsymbol{Z}$.

- **Graphical criterion**

Let $\boldsymbol{Z}$ and Y be variables in a Linear, Non-Gaussian Acyclic Model (LiNGAM). $\boldsymbol{Z}$ and Y satisfies the IN condition iff

- All variables in $\boldsymbol{Z}$ are causally earlier than Y, and
- the common cause for Y and each variable in $\boldsymbol{Z}$, if there is any, is in $\boldsymbol{Z}$.



Let $\boldsymbol{Z} = \{X_2\}$ and $Y = \{X_3\}$.

- $(\boldsymbol{Z}, Y)$ follows the IN condition.
- $\{X_2\}$ is causally earlier than $\{X_3\}$ and their common cause ($\{X_2\}$) are both in $\boldsymbol{Z}$.

IN Condition (Shimizu et al., JMLR'11)

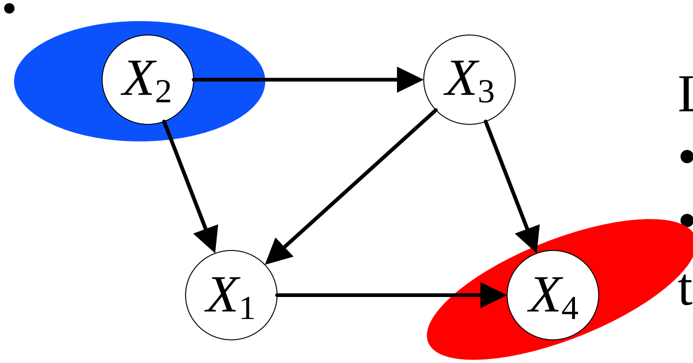
$$\boldsymbol{Z} \longrightarrow Y$$

DEF[Independent Noise(IN), condition] $(\boldsymbol{Z}, Y)$ follows the IN condition iff the residual of regressing Y on $\boldsymbol{Z}$, $Y - \tilde{\omega}^{\mathrm{T}}\boldsymbol{Z}$, is independent from $\boldsymbol{Z}$.

- **Graphical criterion**

Let $\boldsymbol{Z}$ and Y be variables in a Linear, Non-Gaussian Acyclic Model (LiNGAM). $\boldsymbol{Z}$ and Y satisfies the IN condition iff

- All variables in $\boldsymbol{Z}$ are causally earlier than Y, and
- the common cause for Y and each variable in $\boldsymbol{Z}$, if there is any, is in $\boldsymbol{Z}$.



Let $\boldsymbol{Z} = \{X_2\}$ and $Y = \{X_4\}$.

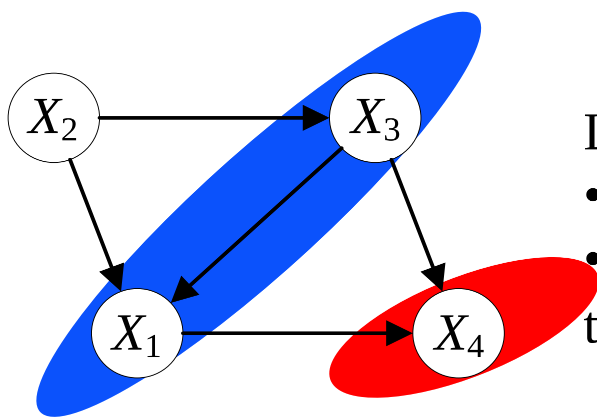
- $(\boldsymbol{Z}, Y)$ follows the IN condition.
- $\{X_2\}$ is causally earlier than $\{X_4\}$ and their common cause ($\{X_2\}$) are both in $\boldsymbol{Z}$.

IN Condition (Shimizu et al., JMLR'11)

$$\boldsymbol{Z} \longrightarrow Y$$

DEF[Independent Noise(IN), condition] $(\boldsymbol{Z}, Y)$ follows the IN condition iff the residual of regressing Y on $\boldsymbol{Z}$, $Y - \tilde{\omega}^{\mathrm{T}}\boldsymbol{Z}$, is independent from $\boldsymbol{Z}$.

- Let $\boldsymbol{Z}$ and Y be variables in a Linear, Non-Gaussian Acyclic Model (LiNGAM). $\boldsymbol{Z}$ and Y satisfies the IN condition iff
  - All variables in $\boldsymbol{Z}$ are causally earlier than Y, and
  - the common cause for Y and each variable in $\boldsymbol{Z}$, if there is any, is in $\boldsymbol{Z}$.



Let $\boldsymbol{Z} = \{X_1, X_3\}$ and $Y = \{X_4\}$.
- $(\boldsymbol{Z}, Y)$ follows the IN condition.
- $\{X_1, X_3\}$ is causally earlier than $\{X_4\}$ and their common cause ($\{X_1, X_3\}$) are both in $\boldsymbol{Z}$.

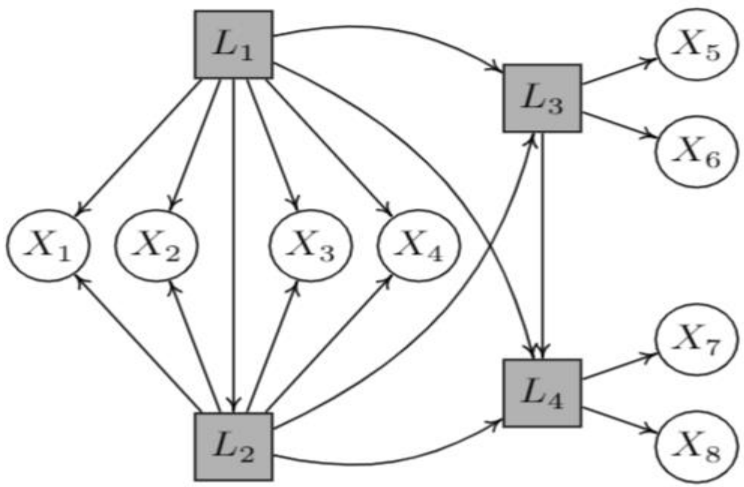
Is it possible to solve the latent-variable problem similar to IN condition?

Linear Non-Gaussian Latent Variable Model (LiNGLaM)

- **Measured variables** (e.g., answer scores in psychometric questionnaires) may not be directly causally related but were *generated by causally related latent Variables*

- Assume variables were generated by the Linear, Non-Gaussian Latent Variable Model (LiNGLaM)

- The whole model is **identifiable** with the Generalized Independent Noise (GIN) condition

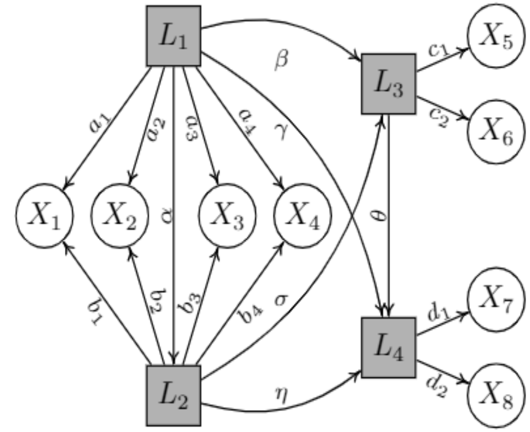


A simple structure that satisfies LiNGLaM

Task: find latent variables and their underlying causal structure from these observed data.

Basic Idea

Using Measured Variables as Surrogate

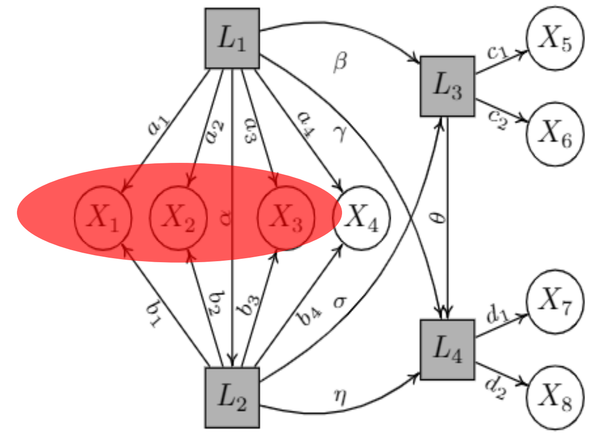


Basic Idea

Using Measured Variables as Surrogate

Consider $\boldsymbol{Y} = \{X_1, X_2, X_3\}$, we have

$$\underbrace{\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}}_{\boldsymbol{A}} \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} + \underbrace{\begin{bmatrix} \varepsilon_{X_1} \\ \varepsilon_{X_2} \\ \varepsilon_{X_3} \end{bmatrix}}_{\mathbf{E_Y}}$$



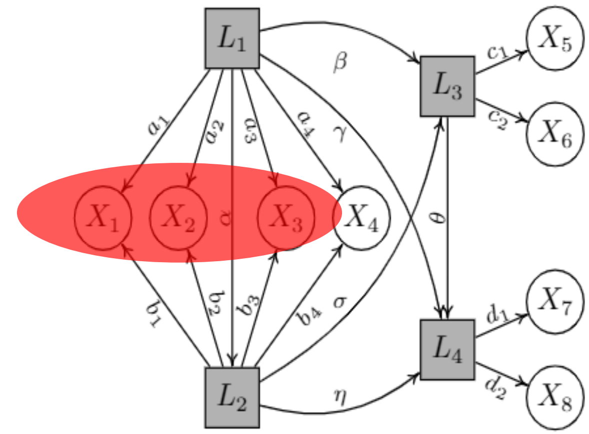
$\exists$ nonzero vector $\boldsymbol{\omega}$ s.t. $\boldsymbol{\omega} \cdot \mathbf{Cov}(\boldsymbol{Y}, \boldsymbol{L}_{\mathbf{1,2}}) = \mathbf{0}$, then $\boldsymbol{\omega A} = \mathbf{0}$, so $\omega^{\mathrm{T}} \boldsymbol{Y} = \omega^{\mathrm{T}} \mathbf{E}$ is independent from L.

Basic Idea

Using Measured Variables as Surrogate

Consider $\boldsymbol{Y} = \{X_1, X_2, X_3\}$, we have

$$\underbrace{\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}}_{\boldsymbol{A}} \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} + \underbrace{\begin{bmatrix} \varepsilon_{X_1} \\ \varepsilon_{X_2} \\ \varepsilon_{X_3} \end{bmatrix}}_{\mathbf{E_Y}}$$



$\exists$ nonzero vector $\boldsymbol{\omega}$ s.t. $\boldsymbol{\omega} \cdot \mathbf{Cov}(\boldsymbol{Y}, \boldsymbol{L_{1,2}}) = \mathbf{0}$, then $\boldsymbol{\omega A} = \mathbf{0}$, so $\omega^{\mathrm{T}}\boldsymbol{Y} = \omega^{\mathrm{T}}\mathbf{E}$ is independent from L.

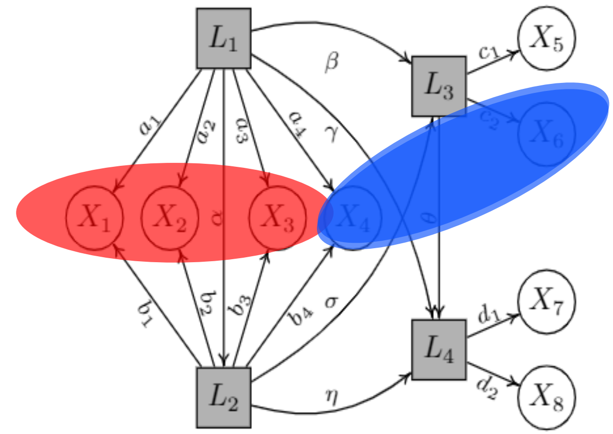
However, we don't have access to L_1 and L_2

Basic Idea

Using Measured Variables as Surrogate

Consider $\boldsymbol{Y} = \{X_1, X_2, X_3\}$, we have

$$\underbrace{\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}}_{\boldsymbol{A}} \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} + \underbrace{\begin{bmatrix} \varepsilon_{X_1} \\ \varepsilon_{X_2} \\ \varepsilon_{X_3} \end{bmatrix}}_{\mathbf{E_Y}}$$



$\exists$ nonzero vector $\boldsymbol{\omega}$ s.t. $\boldsymbol{\omega} \cdot \mathbf{Cov}(\boldsymbol{Y}, \boldsymbol{L_{1,2}}) = \mathbf{0}$, then $\boldsymbol{\omega A} = \mathbf{0}$, so $\omega^{\mathrm{T}}\boldsymbol{Y} = \omega^{\mathrm{T}}\mathbf{E}$ is independent from L.

However, we don't have access to L_1 and L_2

Fortunately, use $\boldsymbol{Z} = (X_4, X_6)^{\mathrm{T}}$ instead, we have

$\exists$ nonzero vector $\boldsymbol{\omega}$ s.t. $\boldsymbol{\omega} \cdot \mathbf{Cov}(\boldsymbol{Y}, \boldsymbol{Z}) = \mathbf{0}$

$$\boldsymbol{\omega} = [a_2 b_3 - b_2 a_3, b_1 a_3 - a_1 b_3, a_1 b_2 - b_1 a_2]^T$$

$\Rightarrow \omega^{\mathrm{T}}\boldsymbol{Y}$ is independent from L.

GIN Condition

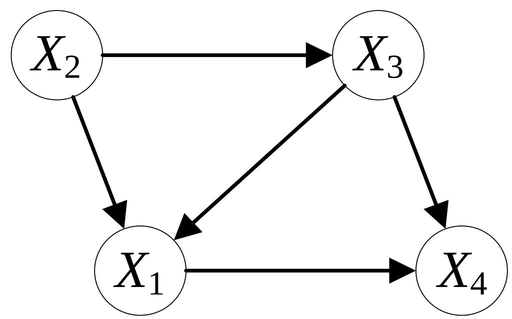
DEF[Generalized Independent Noise(GIN), condition] $(\boldsymbol{Z}, \boldsymbol{Y})$ follows the GIN condition iff there exists non-zeros ω such that $\omega^{\mathrm{T}}\mathbb{E}[\boldsymbol{Y}\boldsymbol{Z}^{\mathrm{T}}] = 0$ and $\omega^{\mathrm{T}}\boldsymbol{Y}$ is independent from $\boldsymbol{Z}$.

GIN Condition

DEF[Generalized Independent Noise(GIN), condition] $(\boldsymbol{Z}, \boldsymbol{Y})$ follows the GIN condition iff there exists non-zeros ω such that $\omega^{\mathrm{T}}\mathbb{E}[\boldsymbol{Y}\boldsymbol{Z}^{\mathrm{T}}] = 0$ and $\omega^{\mathrm{T}}\boldsymbol{Y}$ is independent from $\boldsymbol{Z}$.

- IN condition can be seen as a special case of the GIN condition (See more details in the Proposition 2).

For example, $(\{X_2, X_3\}, \{X_1\})$ satisfy IN contiditon iff $(\{X_2, X_3\}, \{X_1, X_2, X_3\})$ satisfy GIN condition



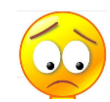
GIN Condition

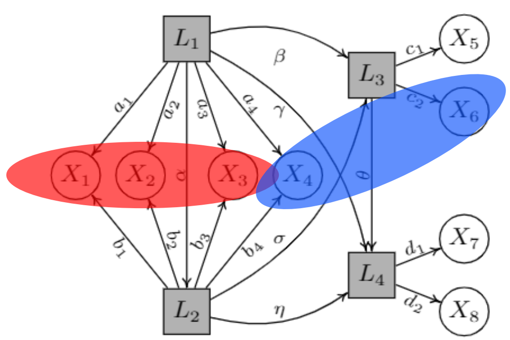
DEF[Generalized Independent Noise(GIN), condition] $(\boldsymbol{Z}, \boldsymbol{Y})$ follows the GIN condition iff there exists non-zeros ω such that $\omega^{\mathrm{T}}\mathbb{E}[\boldsymbol{Y}\boldsymbol{Z}^{\mathrm{T}}] = 0$ and $\omega^{\mathrm{T}}\boldsymbol{Y}$ is independent from $\boldsymbol{Z}$.

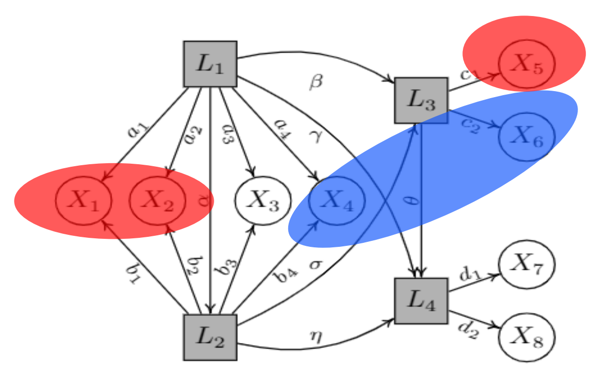
- IN condition can be seen as a special case of the GIN condition (See more details in the Proposition 2).

For example, $(\{X_2, X_3\}, \{X_1\})$ satisfy IN contiditon iff $(\{X_2, X_3\}, \{X_1, X_2, X_3\})$ satisfy GIN condition

Graphical criterion

- Let $\boldsymbol{Z}$ and $\boldsymbol{Y}$ be variables in a LiNGLaM. If $(\boldsymbol{Z}, \boldsymbol{Y})$ follows the GIN condition, there is an exogenous subset of the common cause of $\boldsymbol{Y}$ to d-separate from $\boldsymbol{Y}$ from $\boldsymbol{Z}$.
  - ✓ $(\{X_4, X_6\}, \{X_1, X_2, X_3\})$; $(\{X_3, X_4\}, \{X_1, X_2, X_5\})$ ... satisfy GIN 
  - ✓ $(\{X_4, X_6\}, \{X_1, X_2, X_5\})$; $(\{X_3, X_6\}, \{X_1, X_2, X_5\})$ ... don't satisfy GIN 





Determine where the latent variables are

GIN Condition

DEF[Generalized Independent Noise(GIN), condition] ($\boldsymbol{Z}, \boldsymbol{Y}$) follows the GIN condition iff there exists non-zeros ω such that $\omega^{\mathrm{T}}\mathbb{E}[\boldsymbol{Y}\boldsymbol{Z}^{\mathrm{T}}] = 0$ and $\omega^{\mathrm{T}}\boldsymbol{Y}$ is independent from $\boldsymbol{Z}$.

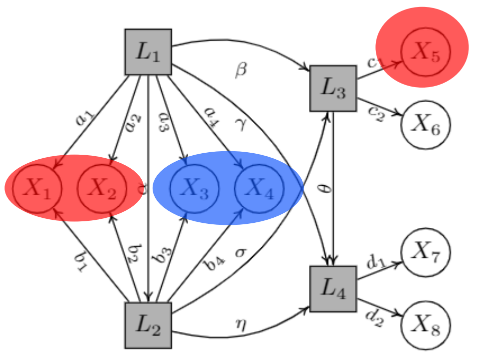
- IN condition can be seen as a special case of the GIN condition (See more details in the Proposition 2).

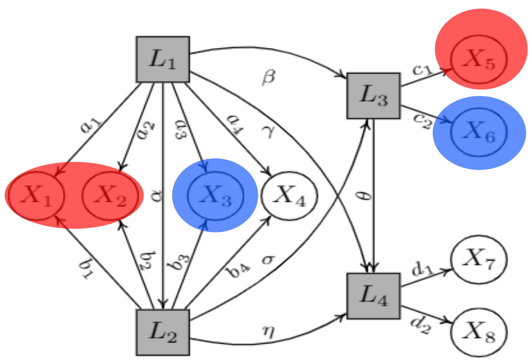
For example, ($\{X_2, X_3\}, \{X_1\}$) satisfy IN contiditon iff ($\{X_2, X_3\}, \{X_1, X_2, X_3\}$) satisfy GIN condition

Graphical criterion

➢ Let $\boldsymbol{Z}$ and $\boldsymbol{Y}$ be variables in a LiNGLaM. If ($\boldsymbol{Z}, \boldsymbol{Y}$) follows the GIN condition, there is an exogenous subset of the common cause of $\boldsymbol{Y}$ to d-separate from $\boldsymbol{Y}$ from $\boldsymbol{Z}$.

- ✓ ($\{X_4, X_6\}, \{X_1, X_2, X_3\}$); ($\{X_3, X_4\}, \{X_1, X_2, X_5\}$) ... satisfy GIN 
- ✓ ($\{X_4, X_6\}, \{X_1, X_2, X_5\}$); ($\{X_3, X_6\}, \{X_1, X_2, X_5\}$) ... don't satisfy GIN 





Determine the causal order of the latent variables

Outline

- Background and Related works
- IN Condition to GIN Condition
- Estimating LiNGLaM Based on GIN
- Experiments and Application
- Conclusion and Further work

Estimating LiNGLaM Based on GIN

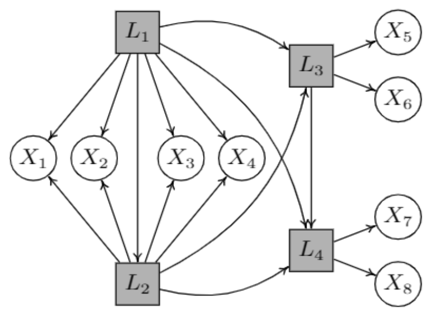
We proposed a *two-steps* algorithm.

- Step 1: find *causal clusters* (variables sharing the same latent variables as parents);
- Step 2: determine *causal order* of the latent variables;
- Estimate the coefficients if needed

Estimating LiNGLaM Based on GIN

We proposed a *two-steps* algorithm.

- Step 1: find *causal clusters* (variables sharing the same latent variables as parents);
- Step 2: determine *causal order* of the latent variables;
- Estimate the coefficients if needed

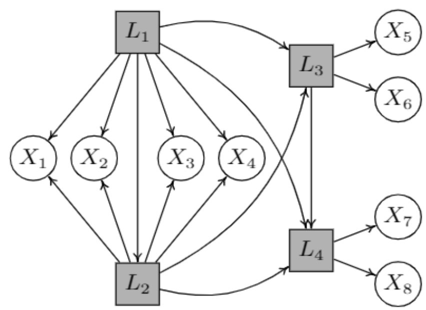


Ground-truth graph

Estimating LiNGLaM Based on GIN

We proposed a *two-steps* algorithm.

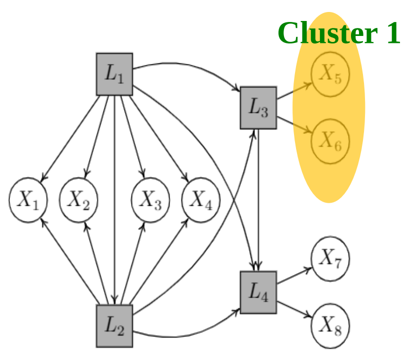
- Step 1: find *causal clusters* (variables sharing the same latent variables as parents);
- Step 2: determine *causal order* of the latent variables;
- Estimate the coefficients if needed



Ground-truth graph

E.g., Test |latent|=1, we have
$(\{X_1, \ldots, X_4, X_7, X_8\}, \{X_5, X_6\})$ satisfies GIN.
Thus, $\{X_5, X_6\}$ is a cluster.
Similarly,
$(\{X_1, \ldots, X_4, X_5, X_6\}, \{X_7, X_8\})$ satisfies GIN.
Thus, $\{X_7, X_8\}$ is a cluster.

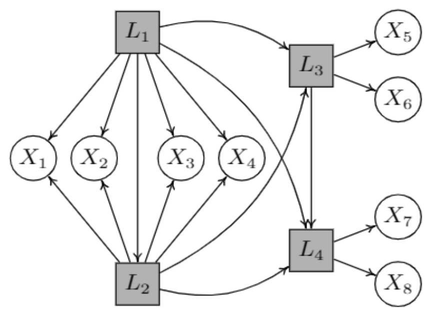
Run Step 1



Estimating LiNGLaM Based on GIN

We proposed a *two-steps* algorithm.

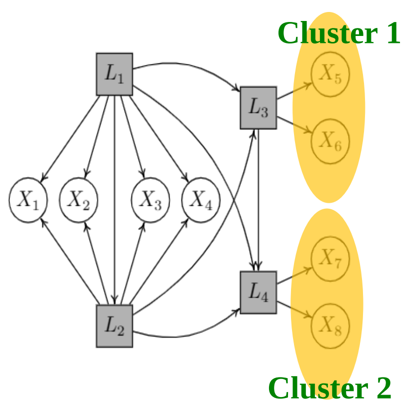
- Step 1: find *causal clusters* (variables sharing the same latent variables as parents);
- Step 2: determine *causal order* of the latent variables;
- Estimate the coefficients if needed



Ground-truth graph

E.g., Test |latent|=1, we have
$(\{X_1, \dots, X_4, X_7, X_8\}, \{X_5, X_6\})$ satisfies GIN.
Thus, $\{X_5, X_6\}$ is a cluster.
Similarly,
$(\{X_1, \dots, X_4, X_5, X_6\}, \{X_7, X_8\})$ satisfies GIN.
Thus, $\{X_7, X_8\}$ is a cluster.

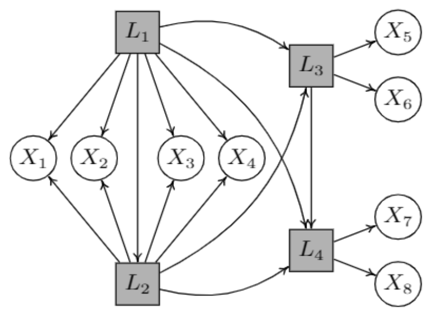
Run Step 1



Estimating LiNGLaM Based on GIN

We proposed a *two-steps* algorithm.

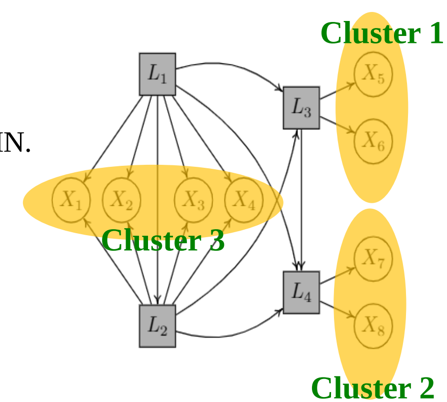
- Step 1: find *causal clusters* (variables sharing the same latent variables as parents);
- Step 2: determine *causal order* of the latent variables;
- Estimate the coefficients if needed



Ground-truth graph

E.g., Test |latent|=2, we have ($\{X_5, X_6, X_7, X_8\}, \{X_1, X_2, X_3\}$) satisfies GIN. Thus, $\{X_1, X_2, X_3\}$ is a cluster.

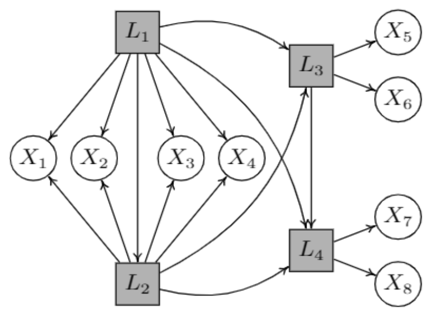
Run Step 1



Estimating LiNGLaM Based on GIN

We proposed a *two-steps* algorithm.

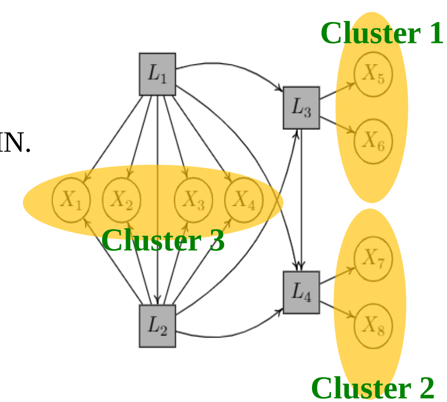
- Step 1: find *causal clusters* (variables sharing the same latent variables as parents);
- Step 2: determine *causal order* of the latent variables;
- Estimate the coefficients if needed



Ground-truth graph

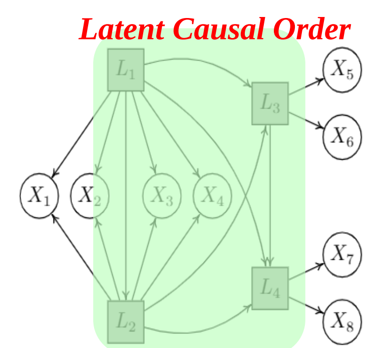
E.g., Test |latent|=2, we have ($\{X_5, X_6, X_7, X_8\}$, $\{X_1, X_2, X_3\}$) satisfies GIN. Thus, $\{X_1, X_2, X_3\}$ is a cluster.

Run Step 1



E.g., ($\{X_3, X_4\}$, $\{X_1, X_2, X_5\}$) satisfies GIN, and ($\{X_3, X_4\}$, $\{X_1, X_2, X_7\}$) satisfies GIN.
We have $\{L_1, L_2\}$ is ***causally earlier*** than L_3 and L_4.

Run Step 2



Outline

- Background and Related works
- IN Condition to GIN Condition
- Algorithm and a Toy Example
- Experiments and Application
- Conclusion and Further work

Experiments

We simulate data following the LiNGLaM, including 4 cases, with different DAG structures for and measurement variables and latent variables.

Goal: ***find clusters*** (determine the location of latent variables)?

- ***Latent oimission***: measure omitted latent variables
- ***Latent commission***: measure falsely detected latent variables
- ***Mismeasurements***: measure the misclassification of observed variables

Table 1: Results with GIN, LSTC, FOFC, and BPC for learning causal clusters.

		Latent omission				Latent commission				Mismeasurements			
Algorithm		GIN	LSTC	FOFC	BPC	GIN	LSTC	FOFC	BPC	GIN	LSTC	FOFC	BPC
Case 1	500	0.00(0)	0.00(0)	1.00(10)	0.50(10)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)
	1000	0.00(0)	0.00(0)	1.00(10)	0.50(10)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)
	2000	0.00(0)	0.00(0)	1.00(10)	0.50(10)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)
Case 2	500	0.10(2)	0.20(4)	0.9(10)	0.50(10)	0.00(0)	0.05(1)	0.00(0)	0.00(0)	0.12(2)	0.12(4)	0.00(0)	0.20(10)
	1000	0.05(1)	0.15(3)	1.00(10)	0.50(10)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.04(1)	0.12(3)	0.00(0)	0.20(10)
	2000	0.00(0)	0.00(0)	1.00(10)	0.50(10)	0.00(0)	0.02(2)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.20(10)
Case 3	500	0.20(3)	0.20(3)	0.13(9)	0.10(1)	0.00(0)	0.03(3)	0.00(0)	0.00(0)	0.19(3)	0.17(3)	0.00(0)	0.00(0)
	1000	0.06(2)	0.13(2)	0.16(10)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.06(2)	0.00(0)	0.00(0)	0.00(0)
	2000	0.00(0)	0.00(0)	0.50(10)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.00(0)
Case 4	500	0.13(4)	0.40(6)	0.90(10)	0.63(10)	0.00(0)	0.23(5)	0.00(0)	0.00(0)	0.04(2)	0.15(6)	0.02(2)	0.06(4)
	1000	0.10(3)	0.26(6)	0.93(10)	0.66(10)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.05(3)	0.11(2)	0.01(1)	0.02(2)
	2000	0.03(1)	0.32(6)	1.00(10)	0.70(10)	0.00(0)	0.00(0)	0.00(0)	0.00(0)	0.04(1)	0.11(3)	0.00(10)	0.00(0)

Note: The number in parentheses indicates the number of occurrences that the current algorithm *cannot* correctly solve the problem.

The lower the better

Our proposed algorithm is more efficient and can find all latent variables!

Application to Teacher's Burnout Data

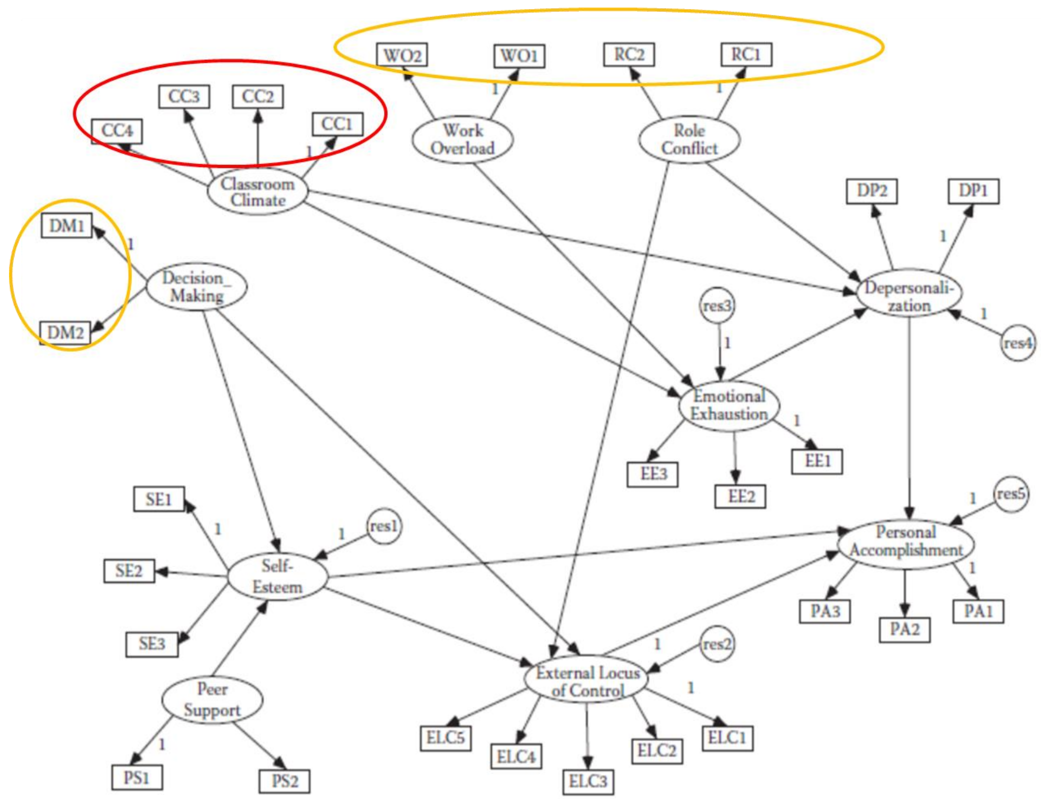
We apply our algorithm to discover the underlying causal structure behind the Teacher's Burnout. The data set contains 28 measured variables (See [Byrne, 2010]).

- Discovered clusters and causal order of the latent variables:

Causal Clusters	Observed variables
S_1 (1)	$RC_1, RC_2, WO_1, WO_2, DM_1, DM_2$
S_2 (1)	CC_1, CC_2, CC_3, CC_4
S_3 (1)	PS_1, PS_2
S_4 (1)	$ELC_1, ELC_2, ELC_3, ELC_4, ELC_5$
S_5 (2)	$SE_1, SE_2, SE_3, EE_1, EE_2, EE_3, DP_1, PA_3$
S_6 (3)	DP_2, PA_1, PA_2

$$\tilde{L}(S_1) \succ L(S_2) \succ L(S_3) \succ L(S_5) \succ L(S_4) \succ L(S_6).$$

(from root to leaf)



Hypothesized model by experts [Byrne, 2010]

Most of results are consistent with the hypothesized model

Outline

- Background and Related works
- IN Condition to GIN Condition
- Algorithm and a Toy Example
- Experiments and Application
- Conclusion and Further work

Conclusion and Further work

Conclusion

- Essential to learn **hidden causal representation**
- The whole latent variable causal model is **identifiable** under suitable assumptions
- GIN condition is a **powerful** extension of IN condition

Further work

- ✓ Develope an efficient algorithm that is able to recover the the LiNGLaM with **directed edges between observed variables** in a principled way;
- ✓ Show the (partial) **identifiability of the causal coefficients** in the model and develop an estimation method for them.

Thank you for your attention!