# DMIR 组会 · 工作汇报 · 综述开题

## DMIR Group Meeting · Work Report · Review Proposal

陈炫志 (Xuanzhi CHEN) | May 26, 2023 | Guangzhou, China
xuanzhichen.42@gmail.com/xuanzhichen.github.io

---

基于不完全观察数据的时序因果关系发现综述

---

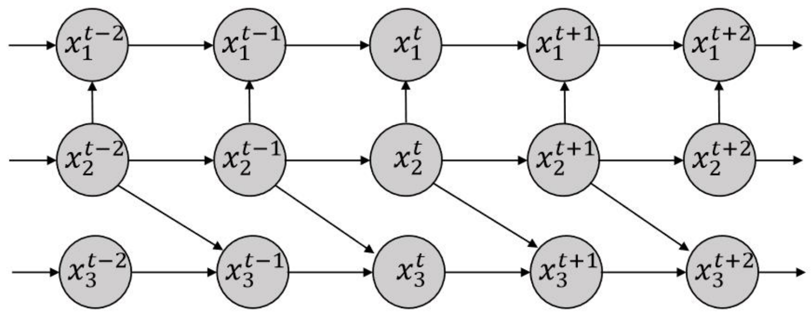
**A Survey on Causal Discovery with Incomplete Time-Series Data**

# CONTENT

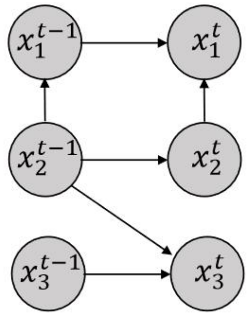
- **Background**
- **Frameworks as to TCD with Incomplete Data**
    - Constraint-Based Algorithms
    - Functional-Based Algorithms
    - Algorithms Based on Score Functions
    - Algorithms Based on Granger Causality
- **Summary**

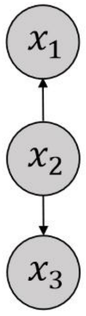
# CONTENT

- **Background**
- **Frameworks as to TCD with Incomplete Data**
  - Constraint-Based Algorithms
  - Functional-Based Algorithms
  - Algorithms Based on Score Functions
  - Algorithms Based on Granger Causality
- **Summary**

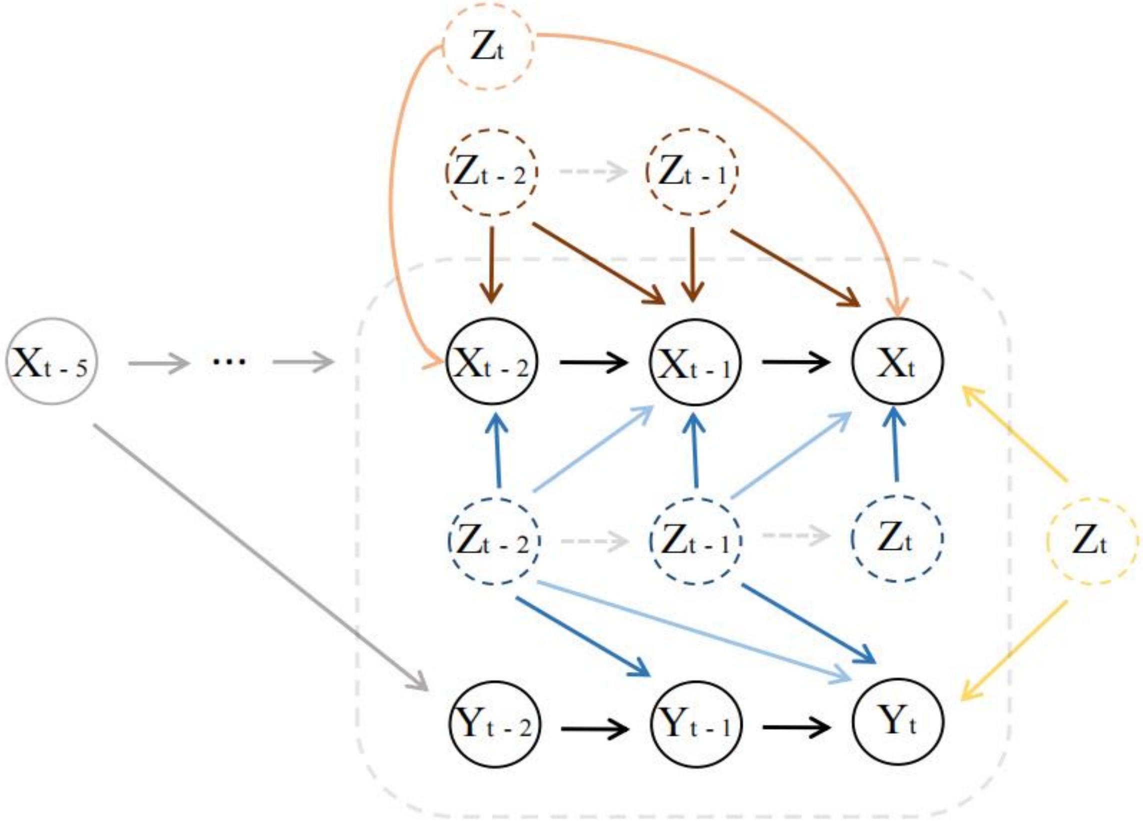


(a) Full time causal graph





(b) Window causal graph (c) Summary causal graph




- Three Types of Temporal Causal Graphs
- Four Types of Temporal Latent Confounders

# CONTENT

- Background
- **Frameworks as to TCD with Incomplete Data**
  - Constraint-Based Algorithms
  - Functional-Based Algorithms
  - Algorithms Based on Score Functions
  - Algorithms Based on Granger Causality
- Summary

# SVAR-FCI (Structure Vector Auto-Regression FCI)

## Motivation/Source

The TS-FCI algorithm cannot handle Contemporaneous Latent Confounding

## Primary Ideas/Intro

Partition the time-series dataset, run the FCI algorithm,
and prune the graph based on temporal context and homogeneity priors

$$\mathbf{x}_t = \left(x_t^1, \ldots, x_t^d\right)^T_{t_0 \leqslant t \leqslant T}, \quad \boxed{\mathbf{x}_t{}' = \left(x_t^1, \ldots, x_t^d, x_{t-1}^1, \ldots, x_{t-1}^d, \ldots, x_{t-p}^1, \ldots, x_{t-p}^d\right)^T_{t_0 \leqslant t \leqslant T}}.$$

| | | | |
|---|---|---|---|
| ts-FCI | SVAR-FCI | LPCMCI<br>Tiered Knowledge | DMAGs & DPAGs<br>SyPI |
| 2010 | 2018 | 2020 | 2021 |



Malinsky, Daniel, and Peter Spirtes. "Causal structure learning from multivariate time series in settings with unmeasured confounding." Proceedings of 2018 ACM SIGKDD workshop on causal discovery. PMLR, 2018.

# LPCMCI (Latent PCMCI)

## Motivation/Source

Momentary Conditional Independence Tests (MCI tests)

Confounding Effect Propagation: Time Lag and Autocorrelation

## Primary Ideas/Intro

Difine middle marks and constraint rules

(e.g., the ancestor-parent-rule to anticipate (time-lag) ancestor variables):

$$S_{\text{default}} := S, \;\; S := \arg\min I\left(x_t^i \,;\, x_{t-p}^j \mid S \cup S_{\text{default}}\right).$$

ts-FCI — 2010

SVAR-FCI — 2018

**LPCMCI**, Tiered Knowledge — **2020**

DMAGs & DPAGs, SyPI — 2021

Gerhardus, Andreas, and Jakob Runge. "High-recall causal discovery for autocorrelated time series with latent confounders." Advances in Neural Information Processing Systems 33 (2020): 12615-12625.

# Other Relevant Constraint-Based Methods

- **Tiered Background Knowledge**: contemporaneous latent confounding, ancestral graphs
- **Directed Maximal Ancestral Graphs, DMAGs**: subclasses within ancestral graphs
- **The SyPI Algorithm**: the concept of *sg-unconfounded* causal paths, variable selection

2010 — ts-FCI
2018 — SVAR-FCI
2020 — LPCMCI, **Tiered Knowledge**
2021 — **DMAGs & DPAGs**, **SyPI**

On the Completeness of Causal Discovery in the Presence of Latent Confounding with Tiered Background Knowledge

Andrews, Bryan, Peter Spirtes, and Gregory F. Cooper. "On the completeness of causal discovery in the presence of latent confounding with tiered background knowledge." PMLR, 2020.

Mastakouri, Atalanti A., Bernhard Schölkopf, and Dominik Janzing. "Necessary and sufficient conditions for causal feature selection in time series with latent common causes." International Conference on Machine Learning. PMLR, 2021.

# CONTENT

- Background
- **Frameworks as to TCD with Incomplete Data**
    - Constraint-Based Algorithms
    - Functional-Based Algorithms
    - Algorithms Based on Score Functions
    - Algorithms Based on Granger Causality
- Summary

# TS-LiNGAM (Time Series LiNGAM)

## Motivation/Source

Temporal extension concerning the Non-Gaussianity of LiNGAM approach

## Primary Ideas/Intro

A two-stage solution algorithm that combines the SVAR model estimation and the LiNGAM analysis

$$\mathbf{x}_t := \sum_{p=1}^{P} B_p \mathbf{x}_{t-p} + B_0 \mathbf{x}_t + \mathbf{n}_t \rightarrow \mathbf{x}_t := \sum_{p=0}^{P} \underbrace{(I - B_0)^{-1} B_p}_{B'_p} \mathbf{x}_{t-p} + \underbrace{(I - B_0)^{-1} \mathbf{n}_t}_{\mathbf{n}'_t} .$$

ts-LiNGAM — 2008
TiMINo — 2013
Algorithm GP-Model — 2015

Hyvärinen, Aapo, et al. "Estimation of a structural vector autoregression model using non-gaussianity." Journal of Machine Learning Research 11.5 (2010).

# GP-Model (Gaussian Process Model)

Notes: issue of miss data

## Motivation/Source

Treat timestamps as the special kind of latent confounders (causal effects vary over time)

$$\mathbf{x}(t) := \sum_{p=1}^{P} B_p(t-p)\mathbf{x}(t-p) + B_0(t)\mathbf{x}(t) + G(t) + \mathbf{n}(t).$$

## Primary Ideas/Intro

Gaussian Process Regression/Gaussian Prior (for latent confounders)/Two-Stage Algorithms

Algorithm

| ts-LiNGAM | TiMINo | GP-Model |
|---|---|---|
| 2008 | 2013 | 2015 |

Huang, Biwei, Kun Zhang, and Bernhard Schölkopf. "Identification of time-dependent causal model: A gaussian process treatment." Twenty-Fourth international joint conference on artificial intelligence. 2015.

# Other Relevant Functional-Based Methods

- **Algorithm**: (Non-Gaussian) Probability Transition Matrix/Generalized Residuals

$$\begin{pmatrix} \mathbf{x}_t \\ \mathbf{z}_t \end{pmatrix} = A \begin{pmatrix} \mathbf{x}_{t-1} \\ \mathbf{z}_{t-1} \end{pmatrix} + \mathbf{n}_t, \quad A := \begin{pmatrix} B & C \\ D & E \end{pmatrix}. \qquad R_t(U_1, U_2) = \begin{pmatrix} I \\ -U_1 \\ -U_2 \end{pmatrix}^T \begin{pmatrix} \mathbf{x}_t \\ \mathbf{x}_{t-1} \\ \mathbf{x}_{t-2} \end{pmatrix}.$$

- **TiMINo**: Nonlinearity/Independent Noise/Mark the latent confounder

$$x_t^i := f_i\left(\mathbf{pa}_{t-p}^i, \ldots, \mathbf{pa}_{t-1}^i, \mathbf{pa}_t^i\right) + n_t^i. \ \text{Where } \mathbf{pa}_t^i \subset \mathbf{x}_t^{N\setminus\{i\}}, \ \mathbf{pa}_{t-p}^i \subseteq \mathbf{x}_{t-p}^N, \ p > 0.$$

| 2008 | 2013 | 2015 |
|---|---|---|
| ts-LiNGAM | TiMINo | Algorithm<br>GP-Model |

Peters, Jonas, Dominik Janzing, and Bernhard Schölkopf. "Causal inference on time series using restricted structural equation models." Advances in neural information processing systems 26 (2013).
Geiger, Philipp, et al. "Causal inference by identification of vector autoregressive processes with hidden components." International Conference on Machine Learning. PMLR, 2015.

# CONTENT

- Background
- **Frameworks as to TCD with Incomplete Data**
  - Constraint-Based Algorithms
  - Functional-Based Algorithms
  - Algorithms Based on Score Functions
  - Algorithms Based on Granger Causality
- Summary

# DyNotears(Dynamic Notears)

## ➤ Motivation/Source

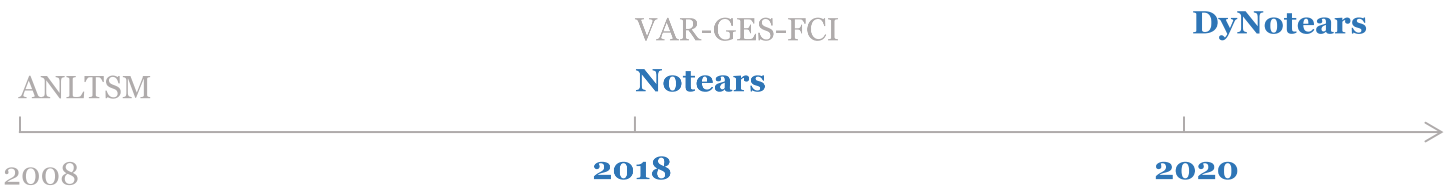
Extension of the Notears model in a temporal context based on non-Gaussianity or homoscedastic noise

## ➤ Primary Ideas/Intro

DBN from the perspective of SEMs / Continuous Optimization / Augmented Lagrangian.

$$\begin{cases} \mathbf{x}_t = W_t\,\mathbf{x}_t + \sum_{p=1}^{P} W_{t-p}\mathbf{x}_{t-p} + \mathbf{n}_t. \\ \mathbf{x}_t = W_t\,\mathbf{x}_t + A_p\,\mathbf{y}_p + \mathbf{n}_t. \end{cases}$$

$$\min_{\mathbf{W},\mathbf{A}} f(\mathbf{W},\mathbf{A}), \quad s.t. \quad h(\mathbf{W}) = 0.$$






Pamfil, Roxana, et al. "Dynotears: Structure learning from time-series data." International Conference on Artificial Intelligence and Statistics. PMLR, 2020.

# ANLTSM(Additive Non-Linear Time Series Model)

Notes: functional-based algorithms as well

## Motivation/Source

Stationarity Assumption / Conditional Independence Tests / Additive Models

$$x_t = B_t x_t + \sum_{p=1}^{P} f(\mathbf{x}_{t-p}) + C_t \mathbf{u}_t + \mathbf{n}_t \ .$$

## Primary Ideas/Intro

Transforming the acquisition of conditional independence into the acquisition of conditional (statistical) Expectations

$$CI(x_t^i \, , x_t^j \mid \mathbf{x}_t^{V\setminus\{i,j\}}) \Leftrightarrow \mathbb{E}\left[x_t^i \mid x_t^j, \ \mathbf{x}_t^{V\setminus\{i,j\}}, \ \mathbf{x}_{t-p}^{V}\right]$$

| 2008 | 2018 | 2020 |
|---|---|---|
| ANLTSM | VAR-GES-FCI<br>Notears | DyNotears |



Chu, Tianjiao, Clark Glymour, and Greg Ridgeway. "Search for Additive Nonlinear Time Series Causal Models." Journal of Machine Learning Research 9.5 (2008).

# CONTENT

- Background
- **Frameworks as to TCD with Incomplete Data**
  - Constraint-Based Algorithms
  - Functional-Based Algorithms
  - Algorithms Based on Score Functions
  - **Algorithms Based on Granger Causality**
- Summary

# Partial GC (Granger Causality)

## Motivation/Source

Limitations of conditional Granger causality analysis: Negative values of F-statistics

## Primary Ideas/Intro

A correction similar to partial correlation has been made,
compared to the F-statistics constructed in conditional Granger causality analysis.

$$S(i,j)=\begin{pmatrix} s_{jj} & s_{ji} \\ s_{ij} & s_{ii} \end{pmatrix}, \Sigma(i,j,k)=\begin{pmatrix} \Sigma_{jj} & \Sigma_{jk} & \Sigma_{ji} \\ \Sigma_{kj} & \Sigma_{kk} & \Sigma_{ki} \\ \Sigma_{ij} & \Sigma_{ik} & \Sigma_{ii} \end{pmatrix} \qquad F_{par}\left(x_t^j \rightarrow x_t^i\right)=\ln\left(\frac{s_{ii}-s_{ij}s_{jj}^{-1}s_{jj}}{\Sigma_{ii}-\Sigma_{ij}\Sigma_{jj}^{-1}\Sigma_{jj}}\right)$$

Granger — 1969
Partial-GC — 2008
likelihood-GC — 2019
DL-GC — 2019-2022



Guo, Shuixia, et al. "Partial Granger causality—eliminating exogenous inputs and latent variables." Journal of neuroscience methods 172.1 (2008): 79-93.

# Other Relevant Granger-Based Methods

- **TCDF ( Temporal Causal Discovery Framework ):**

  Neuro Networks/Attention Score/Times Step (p) Estimation

  $$\textit{For latent confounder } k\text{:}\quad (p_{ji} = p_{ij} = 0) \ \wedge \ (p_{kj} = p_{ki})$$

- **Temporal-VAE GC：** Proxy Variable/Variational Autoencoder

  $$GC\left(x^j \rightarrow x^i \mid \mathbf{x}^{V\setminus\{i,j\}},\ \hat{\mathbf{z}}\right),\ \ \hat{\mathbf{z}} \sim q\left(\hat{\mathbf{z}} \mid x^j, x^i, \mathbf{x}^{P\subset V}\right)$$

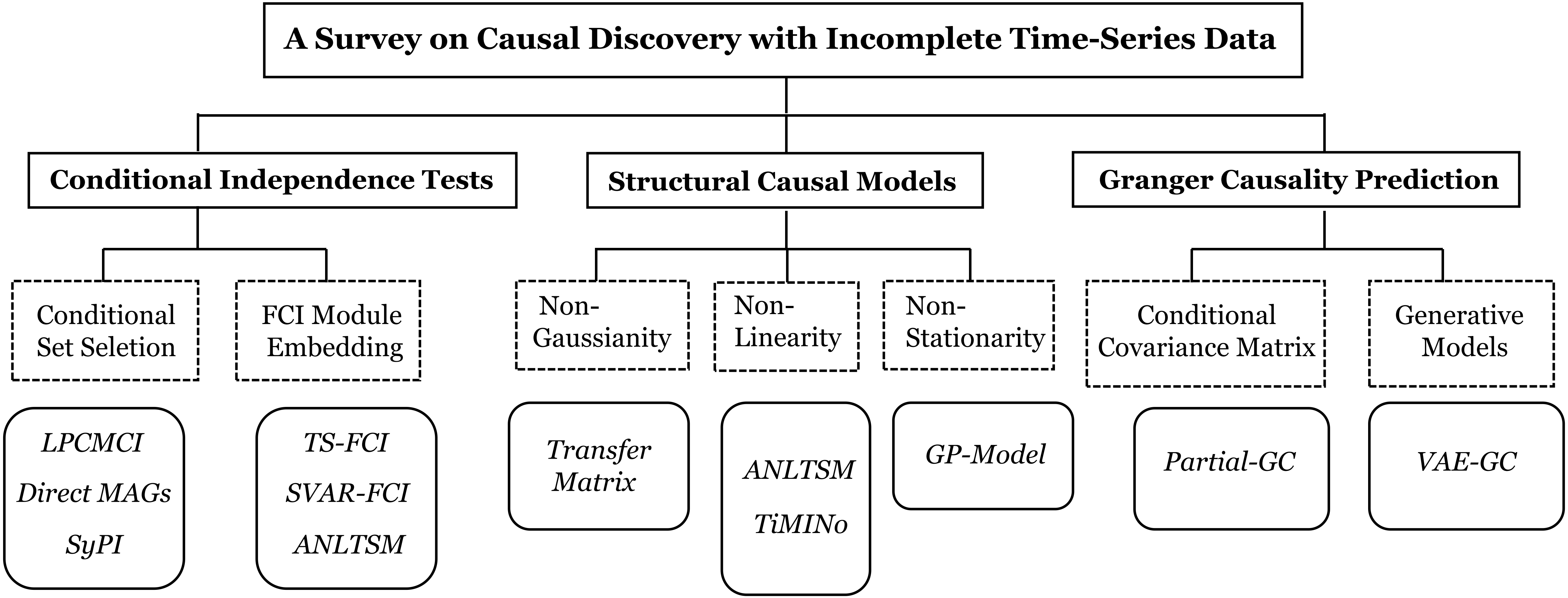
Granger | Partial-GC | likelihood-GC | **DL-GC**

1969 | 2008 | 2019 | **2019-2022**

Nauta, Meike, Doina Bucur, and Christin Seifert. "Causal discovery with attention-based convolutional neural networks." Machine Learning and Knowledge Extraction 1.1 (2019): 312-340.
Yin, Zexuan, and Paolo Barucca. "Deep recurrent modelling of Granger causality with latent confounding." Expert Systems with Applications 207 (2022): 118036.

# CONTENT

- Background
- Frameworks as to TCD with Incomplete Data
    - Constraint-Based Algorithms
    - Functional-Based Algorithms
    - Algorithms Based on Score Functions
    - Algorithms Based on Granger Causality
- **Summary**


A Survey on Causal Discovery with Incomplete Time-Series Data
Conditional Independence Tests
Structural Causal Models
Granger Causality Prediction
Conditional Set Seletion
FCI Module Embedding
Non-Gaussianity
Non-Linearity
Non-Stationarity
Conditional Covariance Matrix
Generative Models
LPCMCI
Direct MAGs
SyPI
TS-FCI
SVAR-FCI
ANLTSM
Transfer Matrix
ANLTSM
TiMINo
GP-Model
Partial-GC
VAE-GC

# Thanks · END

## DMIR Group Meeting · Work Report · Review Proposal

### A Survey on Causal Discovery with Incomplete Time-Series Data

**陈炫志 (Xuanzhi CHEN) | May 26, 2023 | Guangzhou, China**

xuanzhichen.42@gmail.com/xuanzhichen.github.io