Learning Temporally Causal Latent Processes from General Temporal Data

Weiran Yao[†*], Yuewen Sun[‡*], Alex Ho[◊], Changyin Sun[‡], Kun Zhang[†•]

[†] Carnegie Mellon University
[‡] Southeast University
[◊] Rice University
[•] Mohamed bin Zayed University of Artificial Intelligence
(* Equal contribution)

Motivation

- **Causal discovery** is essential in empirical science and artificial intelligence.
- **Identifiability** on the causal latent variables is at the core of causal discovery.

(1) Data generation process:

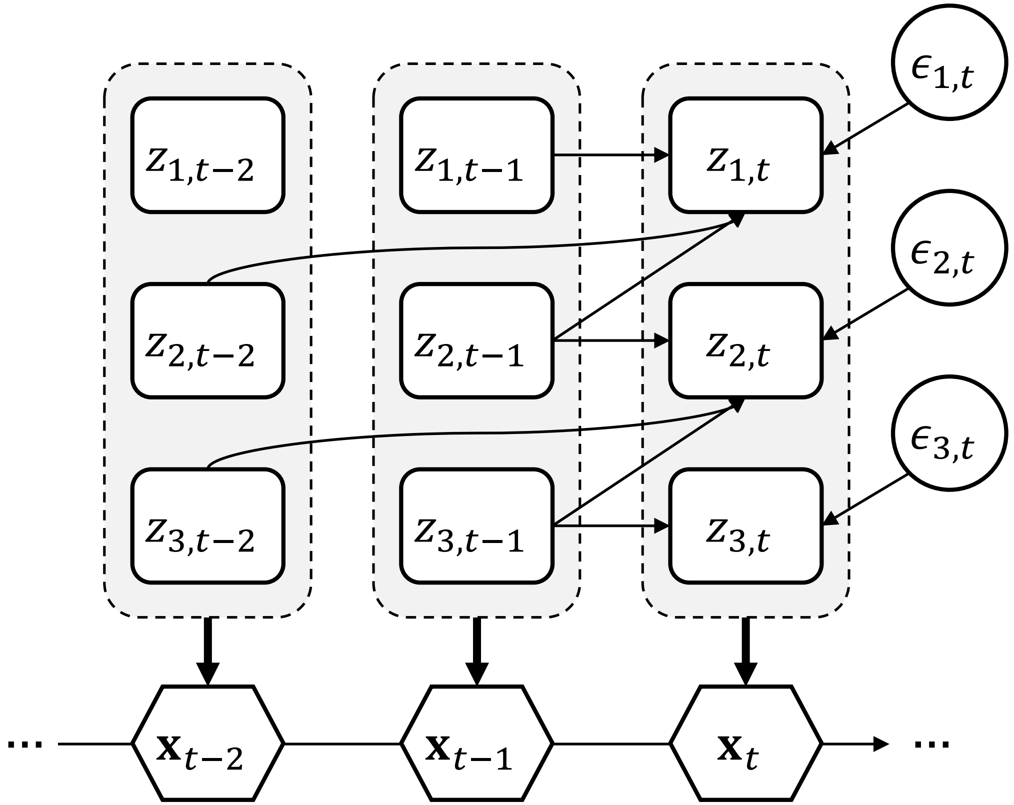
$$\mathbf{x}_t = g(\mathbf{z}_t)$$

(2) Causally-related latent factors:

$$z_{it} = f_i(\mathbf{Pa}(z_{it}), \epsilon_{it})$$

(3) Component-wise identifiability:

$$p_{\hat{g},\hat{f},\hat{p}_\epsilon}(\mathbf{x}_t) = p_{g,f,p_\epsilon}(\mathbf{x}_t) \Rightarrow \hat{g} = g \circ T \circ \pi$$



In this work, we aim to

(1) identify temporally causally-related latent factors under

- ☐ **Parametric**: linear transition + addictive noise

$$\mathbf{z}_t = \sum_{\tau=1}^{L} \mathbf{B}_\tau \mathbf{z}_{t-\tau} + \boldsymbol{\epsilon}_t \text{ with } \epsilon_{it} \sim p_{\epsilon_i}$$

- ☐ **Nonparametric & nonstationary**: nonlinear transition + nonstationary noise

$$z_{it} = f_i(\{z_{j,t-\tau} | z_{j,t-\tau} \in \mathbf{Pa}(z_{it})\}, \epsilon_{it}) \text{ with } \epsilon_{it} \sim p_{\epsilon_i|\mathbf{u}}.$$

(2) develop a theoretically-grounded estimation framework to

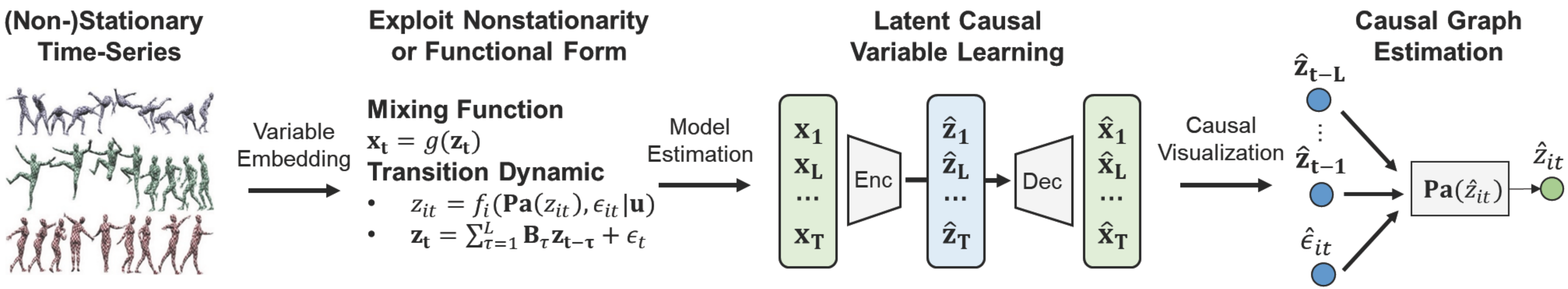
- ☐ enforce our conditions through proper constraints in causal process prior.

■ This is particularly challenging in that

(1) neither sparsity nor minimality assumptions are considered in our work,

(2) whether temporal structure contributes to latent causal discovery is unclear.

Latent tEmporally cAusal Processes (LEAP)



(1) Variable embedding

- ◻ observes stationary/nonstationary time series and
- ◻ assumes the functional/distributional forms of temporal statistics or the nonstationarity in noise.

(2) Model estimation

- ◻ builds upon the framework of sequential VAEs and
- ◻ enforces the proposed condition as constraints for identification.

(3) Causal visualization

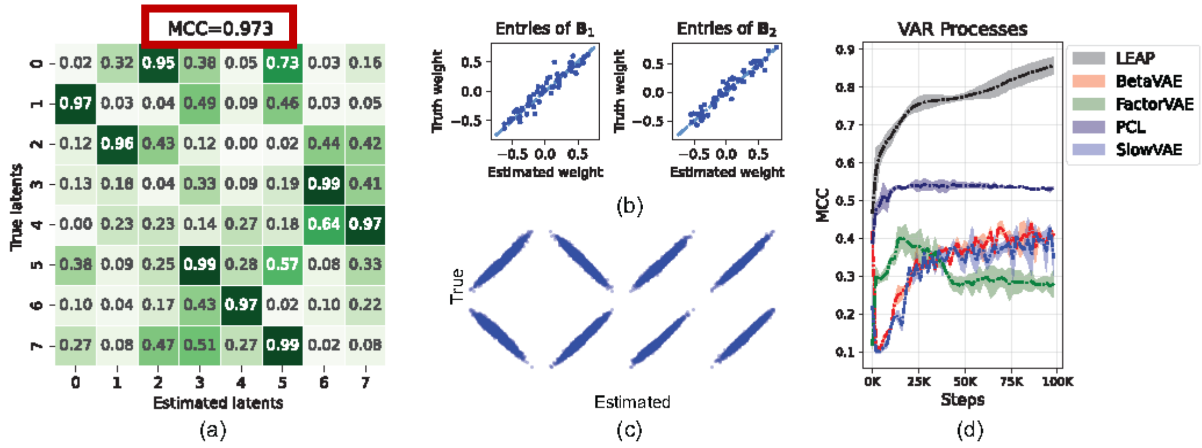
Experiment Result: Synthetic Dataset

ICLR | 2022
Tenth International Conference on Learning Representations

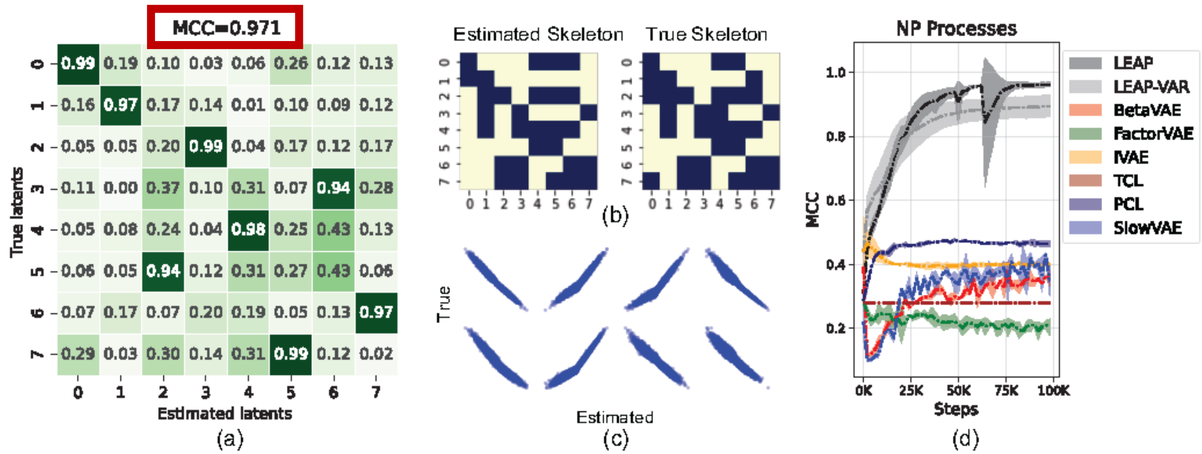
■ Parametric (VAR) and nonparametric (NP) conditions

The latent processes are successfully recovered.

Parametric



Nonparametric



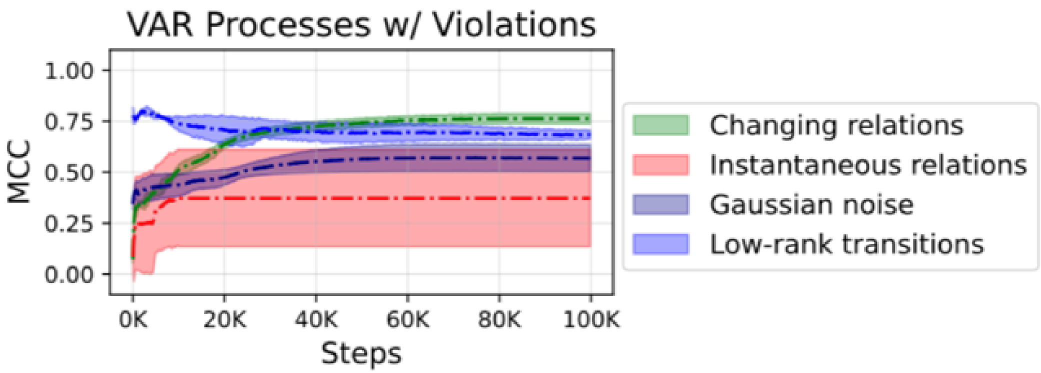
- Our approach achieves **high MCC** for the causally-related factors.
- Causal skeletons are recovered.
- Latent causal variables are estimated up to permutation and component-wise invertible transformation.
- Baselines without using history or assume independent sources **fail to** recover the latent processes.

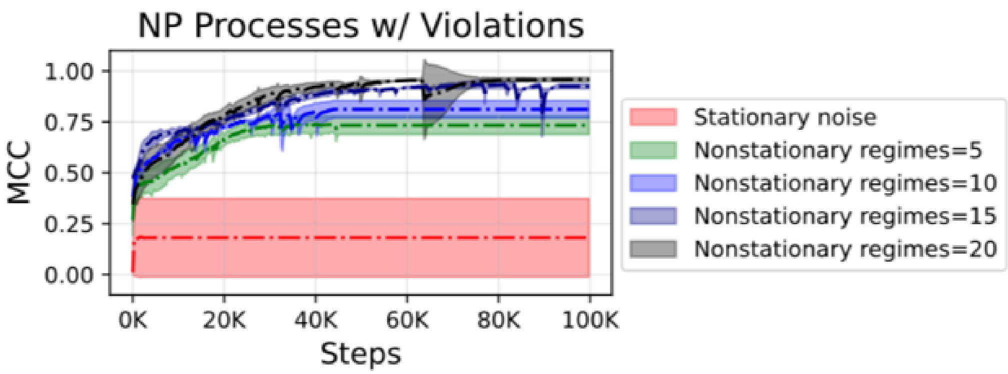
Experiment Result: Synthetic Dataset

ICLR | 2022
Tenth International Conference on Learning Representations

■ Robustness analysis

MCC for temporal data with assumption violations





[VAR] Our approach gains partial identifiability under changing causal relations or low-rank transitions.

[VAR] Instantaneous relations and Gaussian noise distort the results.

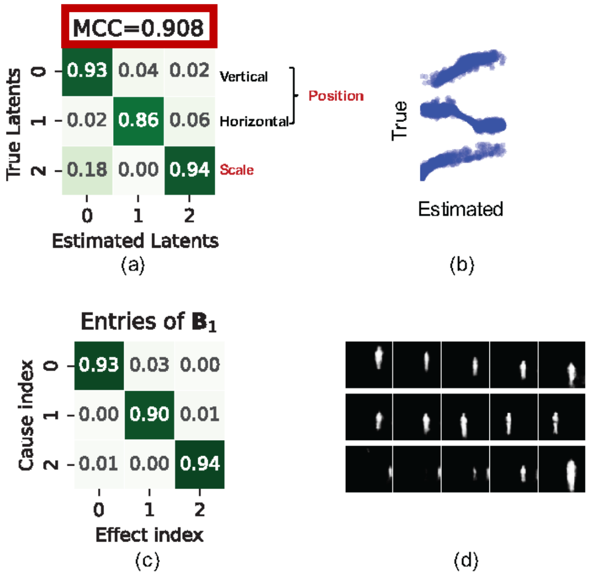
[NP] Nonstationarity is necessary for identifiability in nonparametric transition.

[NP] Our approach **does not strictly need** more than 2n+1 regimes.

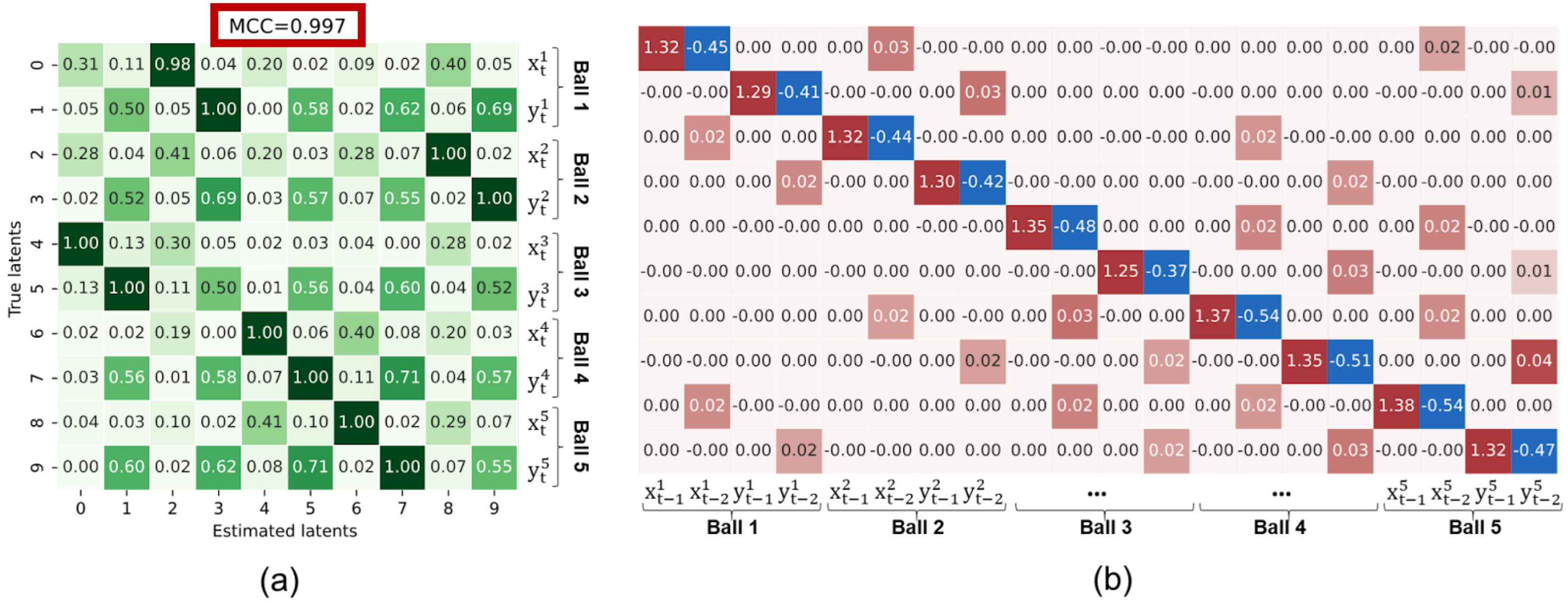
Experiment Result: KiTTiMask and Mass-Spring System

■ Parametric condition

The latent causal processes are successfully recovered.



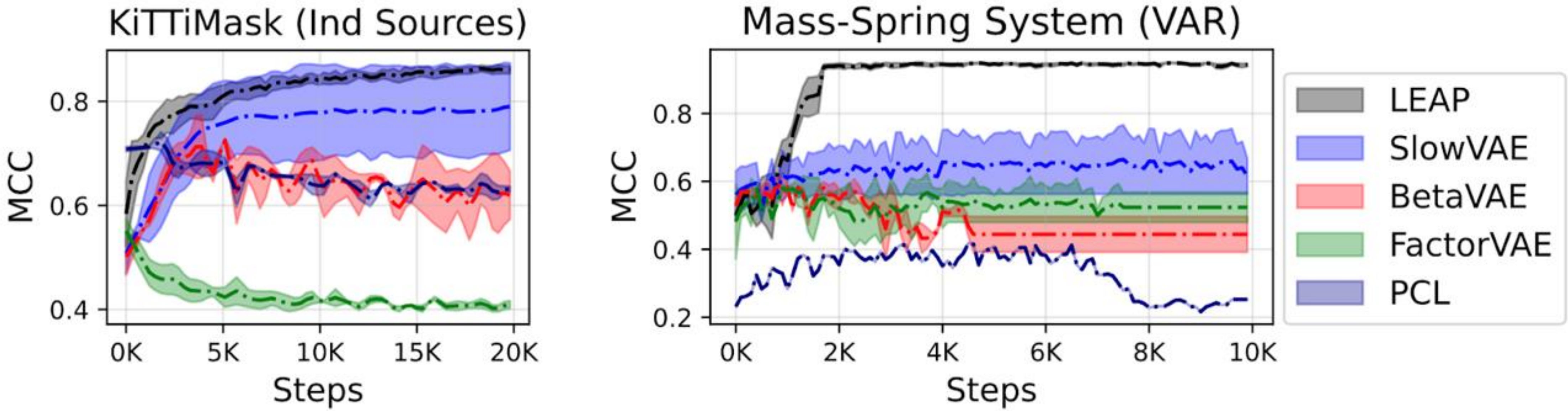
KiTTiMask



Mass-Spring System

Experiment Result: KiTTiMask and Mass-Spring System

- **Comparisons of MCC**



- **KiTTiMask**: SlowVAE performs slightly worse than our approach in independent source dependency.
- **Mass-Spring system**: our approach outperforms the baselines that do not use temporal constraints or assumes independent sources.

Experiment Result: CMU-MoCap

- **Comparisons of latent traversals**

BetaVAE

SlowVAE

LEAP

- LEAP represents the data with causally-related factors, thus can encode the data with fewer latent variables (3 vs 8) and smooth transitions dynamics.

Conclusion and Future Work

- ✓ Propose two provable conditions under which temporally causal latent processes can be identified from their observed nonlinear mixtures.
- ✓ Develop a theoretically-grounded training framework that enforces the assumed conditions through proper constraints.
- ✓ Experimental results on various datasets demonstrate that temporally causal latent processes are reliably identified from observed variables and LEAP considerably outperforms baselines.

- ☐ Extend our identifiability theories and framework to accommodate
  - ☐ instantaneous causal influence between latent causal processes, and
  - ☐ changeable causal influences across regimes.