Stationarity-Aware Causal Discovery in Time Series via Minimal Separating Sets

Shanyun Gao[1], Raghavendra Addanki[2], Tong Yu[2], Ryan A. Rossi[2], Qifan Song[1], Murat Kocaoglu[3]
[1] Purdue University, [2] Adobe Research, [3] Johns Hopkins University.

Motivation

Existing constraint-based methods for time series leverage stationarity for conditional independence testing and reduce the problem to learning the parents of variables **at the final time point.**

However, their separating set search strategy mimics the PC algorithm:

1. **Stationarity** is used only to enforce time-invariant relations and to orient edges **after** dependencies have been estimated;

2. **Separating sets** are used only to establish non-adjacency (i.e., whether a separating set **exists**), not to exploit the information contained in the set itself.

Key Idea

We leverage the **stationarity** of time series **during** the discovery process.

Under stationarity and autoregressive structure, some **separating sets individually** enable deriving a **superset** of the target variable's **parents**.

Lag-Anchored Separating Sets (LASS)

Definition (Stationary SCM). A Stationary Structural Causal Model (SCM) is a tuple $\mathcal{M} = \langle V, \mathcal{F}, \mathcal{E}, \mathbb{P} \rangle$ such that each variable $X_t^j \in V$, $t > \tau_{\max}$ follows:

$$X_t^j = f_j\big(\mathrm{Pa}(X_t^j), \epsilon_t^j\big),\ j \in [n],\ t \in [\tau_{\max}+1, T],$$

where $f_j \in \mathcal{F}$ is time-invariant, $\epsilon_t^j \in \mathcal{E}$ is an unobserved exogenous variable, and $\{\epsilon_t^j\}_{j\in[n],t\in[T]}$ are jointly independent with probability measure $\mathbb{P}$.

Definition (Minimal Separating Set). A minimal separating set $\mathbf{M}$ between variables X and Y is a set that satisfies:

$$X \perp\!\!\!\perp Y \mid \mathbf{M} \quad \text{and} \quad X \not\perp\!\!\!\perp Y \mid \mathbf{M}_{\text{sub}} \quad \forall \mathbf{M}_{\text{sub}} \subsetneq \mathbf{M}$$

A toy example in Fig. 1:

$\{Z_7, Z_8, H_7, H_8\}$ is a Minimal Separating Set between X_4 and H_9,
$\{Y_5, X_6, Z_7, H_7, H_8\}$ is a Minimal Separating Set between X_4 and H_9,
$\{Z_6, Z_7, Z_8, H_6, H_7\}$ is a Minimal Separating Set between X_4 and H_9,
$\{X_7, Y_7\}$ is the only Minimal Separating Set between X_8 and Z_9,
$\{X_7, Y_7, Z_8\}$ is a Minimal Separating Set between X_6 and Z_9.

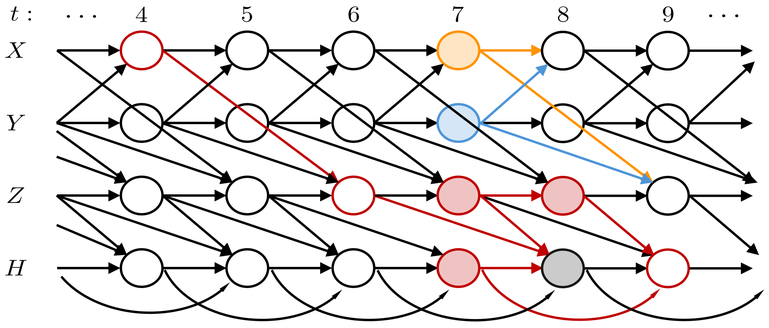


Figure 1: Illustration of a Stationary SCM.

Definition (Lag-Anchored Separating Set). Let X_s and Y_t be two variables with $s < t$. A set $\mathbf{M}$ is a Lag-Anchored Separating Set (**LASS**) between X_s and Y_t if it satisfies:

$$X_s \perp\!\!\!\perp Y_t \mid \mathbf{M}, \quad X_s \not\perp\!\!\!\perp Y_t \mid \mathbf{M}_{\text{sub}} \quad \forall \mathbf{M}_{\text{sub}} \subsetneq \mathbf{M} \quad \text{and} \quad Y_{t-1} \in \mathbf{M}$$

A toy example in Fig. 1:

$\{Z_7, Z_8, H_7, H_8\}$ is a LASS between X_4 and H_9,
$\{Y_5, X_6, Z_7, H_7, H_8\}$ is a LASS between X_4 and H_9,
$\{Z_6, Z_7, Z_8, H_6, H_7\}$ is a Minimal Separating Set between X_4 and H_9 but it is not a LASS,
$\{X_7, Y_7\}$ is the only Minimal Separating Set between X_8 and Z_9 but it is not a LASS,
$\{X_7, Y_7, Z_8\}$ is a LASS between X_6 and Z_9.

Existence of LASS

Lemma 1. For any target variable Y_t and any source variable X_s ($s < t - \tau_{\max}$), there must exist a LASS between X_s and Y_t.

Property of LASS

Lemma 2. (Characterization of LASS). Let $\mathbf{M}$ be a LASS of Y_t.
For any $Z_u \in \mathbf{PA}(Y_t) \setminus \mathbf{M}$, the following holds:

$$Z_{u-1} \in \mathbf{M}$$

A toy example in Fig. 1:

$\mathbf{PA}(H_9) = \{Z_7, Z_8, H_7, H_8\}$. $\{Y_5, X_6, Z_7, H_7, H_8\}$ is a LASS between X_4 and H_9.
Since $Z_8 \in \mathbf{PA}(Y_t) \setminus \mathbf{M}$, $Z_7 \in \mathbf{M}$.

Theorem 1. (Construction of a Superset of the Parent Set). Let $\mathbf{M}$ be a LASS of Y_t. We have:

$$\mathbf{PA}(Y_t) \subset \mathbf{M} \cup \mathbf{M}_{+1} \quad \text{where } \mathbf{M}_{+1} := \{X_{u+1} : X_u \in \mathbf{M}, u+1 < t\}$$

A toy example in Fig. 1:

$\mathbf{PA}(H_9) = \{Z_7, Z_8, H_7, H_8\}$. $\{Y_5, X_6, Z_7, H_7, H_8\}$ is a LASS between X_4 and H_9, hence:

$$\mathbf{PA}(H_9) \subset \mathbf{M} \cup \mathbf{M}_{+1} = \{Y_5, Y_6, X_6, X_7, Z_7, Z_8, H_7, H_8\}$$

Minimal-Separating-Set-based Causal Discovery (MSSD)

For any univariate component time series $\mathbf{X}^j \in \mathbf{V}, j \in [n]$, find a LASS $\mathbf{M}$ of X_t^j and then for any variable $X_s^k \in \mathbf{M} \cup \mathbf{M}_{+1}$,

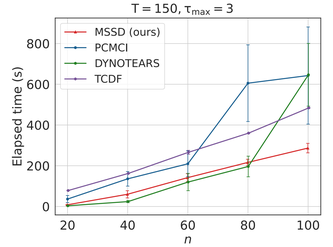
$$\text{remove } X_s^k \text{ from } \mathbf{M} \cup \mathbf{M}_{+1} \text{ if } X_s^k \perp\!\!\!\perp X_t^j \mid \mathbf{M} \cup \mathbf{M}_{+1} \setminus X_s^k$$

Computation Efficiency: With a consistent CI test, for an n-variate time series:

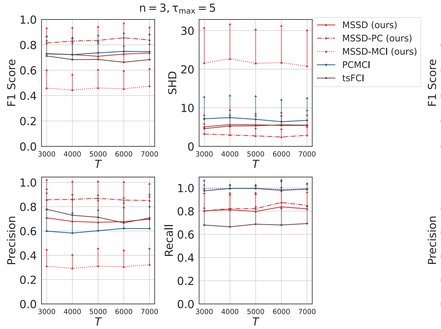
Algorithms	# variable pairs tested per variable	#Total CI tests needed[1]
MSSD (ours) [1]	1	$O(n^2\tau_{\max})$
PCMCI [2]	$n\tau_{\max}$	$O(n^3\tau_{\max}^2)$

[1]Assume that the dimensionality of the conditioning tests is constant.

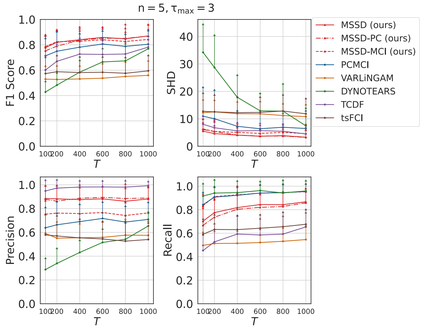
A single LASS can recover a superset of all parents.



Simulations



5 algorithms are tested on 3−multivariate **binary-valued time series** with different time length. Every marker corresponds to the average performance over 50 trials.



8 algorithms are tested on 5− multivariate **continuous-valued time series** with Gaussian noise. Every marker corresponds to the average performance over 100 trials.

[1] Gao, S., Addanki, R., Yu, T., Rossi, R.A., Song, Q. and Kocaoglu, M., Stationarity-Aware Causal Discovery in Time Series via Minimal Separating Sets. In *The 29th International Conference on Artificial Intelligence and Statistics*.
[2] Runge, J., Nowack, P., Kretschmer, M., Flaxman, S. and Sejdinovic, D., 2019. Detecting and quantifying causal associations in large nonlinear time series datasets. *Science advances*, *5*(11), p.eaau4996.