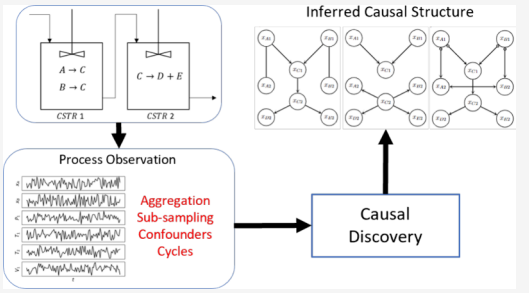
I&EC research
Industrial & Engineering Chemistry Research
pubs.acs.org/IECR
Article

Causal Discovery for Topology Reconstruction in Industrial Chemical Processes

Harman Dewantoro,* Alexander Smith,* and Prodromos Daoutidis*

Cite This: *Ind. Eng. Chem. Res.* 2024, 63, 11530−11543
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: This paper explores the application of causal discovery frameworks to infer the topology of industrial chemical processes, which is crucial for operational decision-making and system understanding. While traditional data-driven methods entail process interventions, causal discovery offers a noninvasive approach. Challenges such as temporal aggregation, subsampling, and unobserved confounders, which can lead to false predictions, are emphasized in the paper. Through simulation case studies, the performance of various causal discovery methods under different observation scenarios is evaluated. Our findings underscore the importance of simultaneously considering instantaneous and lagged causal relations, highlight the suitability of structural equation modeling for temporally aggregated processes, and caution against misinterpretation of subsampled data. Additionally, we demonstrate the utility of the Wiener separation in identifying unobserved confounders, which is essential for navigating the complexity of industrial processes.



1. INTRODUCTION

Digitalization has significantly increased the amount of available data in industrial processes, leading to a large data repository.[1] Industrial chemical processes are also increasingly reliant on interconnected and automated systems with integrated machinery and sensors. Inferring the *topology* of such systems, i.e. the influences and interconnections between process variables, can provide a strong basis for critical operational decisions.[2] For example, knowledge of the system's topology is necessary to devise a feasible control strategy.[3−5] In addition, it is critical for maintaining process safety, fault diagnosis, and mitigation.[6−8] Topological information derived from big data can also help the design and operation of industrial processes by improving sustainability and market competitiveness.[9−11]

Data-driven methods, such as system identification and estimation, are widely applied to historical process data to identify a suitable process model that encodes the topology of the process.[12−14] However, this approach requires process intervention that deliberately perturbs the system from normal operating conditions, leading to downtime and process inefficiency that can impact operational costs. Furthermore, it requires inherent knowledge of the system to properly select a working model from a large class of models. The increasing complexity and size of chemical processes can make it more challenging for the user to select a proper model while avoiding overfitting in the presence of process and measurement noise. Machine learning-based identification is often employed to identify system behaviors.[15,16] However, frameworks that utilize black-box models lack interpretability and guarantee of stability, which makes them unsuitable for risk-sensitive tasks and decision-making in industrial processes.

The field of causal discovery offers an alternative data-driven approach to identify the topology of a large and complex system and has found success in the fields of neuroscience,[17−21] medicine,[22,23] genetics,[24,25] finance,[26] and ecology.[27−29] Causal discovery focuses on identifying causal relationships between process variables beyond correlation, which helps in understanding the underlying mechanisms driving a system and how information propagates throughout the system. The inferred topology is then represented as a graph consisting of nodes and directed edges that are easily interpretable. Given the complex and nonlinear nature of chemical processes, causal discovery algorithms are ideal for uncovering system connectivity without process interventions or a-priori model selection.

Previous works have applied causal discovery algorithms to chemical and industrial processes in the context of root-cause diagnosis and fault detection. Different classes of causal discovery frameworks based on Bayesian networks and temporal causal analysis have been applied to alarms associated with different parts of processes to identify how disturbances propagate through the system and apply proper mitigation.[30−33] Modified versions of the classical Granger Causality have been similarly applied to identify system-wide fault propagation.[34] In our study, our main motivation is to utilize causal discovery to

Received: March 25, 2024
Revised: June 5, 2024
Accepted: June 5, 2024
Published: June 22, 2024

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543

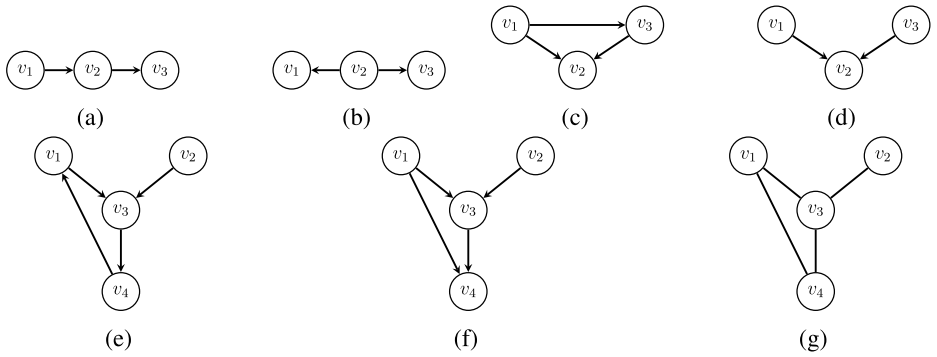


Figure 1. a) A chained triplet. b) A forked triplet. c) A collider structure. d) A v-structure. e) A directed cyclic graph with the cycle $\{v_1, v_3, v_4\}$. f) A DAG. g) The skeleton graph of Figures 1e and 1f.

reconstruct the topology of chemical processes to aid in process identification and control, emphasizing the importance of accurate topological prediction. Previous works have applied causal discovery to temporal observational process data in the context of topological reconstruction.[35−38] However, these approaches often overlook very fast (essentially instantaneous) causal effects induced by different process rates and assume perfect observation of the process.

In dealing with data derived from a real process, one must account for several factors that can lead to false causal predictions.[36,37,39] Low rates of process observation relative to the time scale of the process are often expressed as temporal aggregation or subsampling of the true process, leading to false causal predictions if not accounted for. Observation of real systems at low rates may be the result of data storage restrictions or sensors that observe at low rates. The presence of unobserved variables can also confound the results of causal discovery. Some variables in real processes that have causal influences on other variables might not be directly measurable by the available sensor array, which may lead to causal misprediction. Furthermore, if a subset of process variables in the system evolves at a significantly faster rate than other variables and the observation rate, the causal relation can appear as instantaneous, creating a combination of both instantaneous and lagged causal relations. Simultaneous consideration of both such types of causal relations is critical to recovering the topology of such a system accurately. Non-Gaussian noise and certain structures of the underlying system generating the data, such as control loops and recycle streams, can also make it more difficult to identify the true causal structure of the system. In this study, we aim to address the impact of these factors on the accuracy of the most common causal discovery frameworks and characterize the extent to which we can obtain reliable results using each framework. Through this analysis, we aim to develop guidelines to guide the implementation of different causal discovery methods so as to minimize false predictions with a minimum amount of process knowledge introduced.

The remaining part of this paper is structured as follows: Section 2 will discuss the basic definition of causality in dynamic systems and basic graphical terminology. Section 3 presents an overview of the causal discovery method used in this study. Section 4 describes the process models of two case studies. Section 5 presents and analyzes the results of applying the different causal discovery methods to the two case studies. Finally, Section 6 summarizes the main conclusions of the paper.

2. GRAPHICAL REPRESENTATION OF CAUSALITY

We will begin by formalizing the basic definition of causality and its relation with dynamic systems. Given two variables x and y, we say that a x causes y when interventions on the state of x induce changes in the state of y.[40] In dynamical systems, we are interested in uncovering system-wide causal relationships using causal discovery to predict the effects of intervention (whether deliberate or not). A *causal graph* summarizes the causal relationships within a system as a set of *nodes* representing the variables connected by *directed edges*, representing the direction of causal influence. Causal discovery algorithms take observational data corresponding to different process elements in the system as input and output a causal graph representing the predicted causal influences between every process element.

In a causal graph $\vec{G}$ containing a set of N nodes, $\mathcal{V} = \{v_1, \ldots, v_N\}$, with a set of directed edges $\vec{E}$, the causal influence of v_j to v_i is represented as an arrow pointing from v_j to v_i, $v_j \rightarrow v_i$, where we define node v_j as the *source* and node v_i as the *target*. A causal matrix $A \in \mathbb{R}^{N\times N}$ encodes the adjacency of the causal influences between the variables, where the causal influence $v_j \rightarrow v_i$ is encoded as $A_{ij} \neq 0$. If there is a directed edge from v_i to v_j, we say v_i is a *parent* of v_j and v_j is a *child* of v_i. A sequence of vertices v_1, ..., v_l is called a *path* starting from v_1 and ending in v_l if an edge exists between every ordered pair (v_{k-1}, v_k) in $\vec{G}$. The path is a *chain* if, for every consecutive pair (v_{k-1}, v_k) in the path, there is a directed edge from v_{k-1} to v_k. If there is a chain from v_i to v_j, we say v_i is a *ancestor* of v_j and v_j is a *descendant* of v_i. A vertex v_k in a path is known as *fork* if both v_{k-1} and v_{k+1} are children of v_k $(v_{k-1} \leftarrow v_k \rightarrow v_{k+1})$ and a *collider* if both v_{k-1} and v_{k+1} are parents of v_k $(v_{k-1} \rightarrow v_k \leftarrow v_{k+1})$. We say that a triplet of nodes v_{k-1}, v_k, and v_{k+1} is a *v-structure* if v_k is a collider with v_{k-1} and v_{k+1} as its parents but v_{k-1} and v_{k+1} are not connected. Figures 1a-1d show examples of a chain, fork, collider, and v-structure triplet. A path is a *cycle* if the first and last vertex of the path are equal. We say that the graph G is a *Directed Acyclic Graph* (DAG) if $\vec{G}$ does not contain any cycles. We obtain the *skeleton* graph G of the directed graph by removing all the directions of every edge in $\vec{G}$ to obtain a graph with only undirected edges. Figures 1e-1g illustrate the differences between a cyclic graph, a DAG, and their corresponding skeleton graph.

We associate each node $v_i \in \mathcal{V}$ to an observational time series variable with a length of $T \in \mathbb{Z}$ time points $x_i(t) = [x_i(0), x_i(1), \ldots, x_i(T)] \in \mathbb{R}^{1\times T}$. We define a vector of observational variables at every time point t as $x(t) = [x_1(t), x_2(t), \ldots, x_N(t)]^T \in \mathbb{R}^{N\times 1}$ and an observatio-

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543

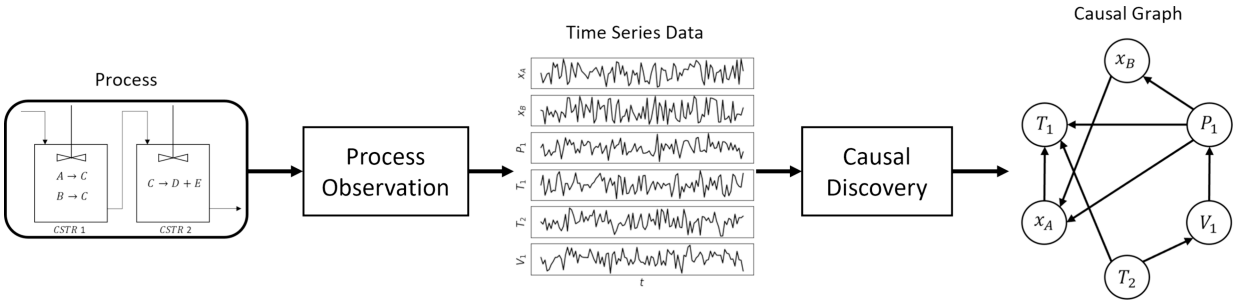


Figure 2. Illustration of the causal discovery workflow.

nal matrix as the concatenation of said vector across all time points $X = [x(0), x(1), \ldots, x(T)] \in \mathbb{R}^{N\times T}$. We can see that each node $v_i \in \mathcal{V}$ corresponds to the i^{th} row of $x(t)$ and X. Furthermore, we define the notation $X_{rem(i)}$ as the removal of the i^{th} row of the matrix X. Multiple indices, such as $X_{rem(ijk)}$, denote the removal of multiple rows of X.

3. CAUSAL DISCOVERY METHODS

Causal discovery methods attempt to identify causal relationships between the variables of a process given the observation matrix X generated by the process. The predicted causal relationships are then expressed as causal graphs. This workflow is illustrated in Figure 2. The remaining part of this section will provide an overview of the causal discovery frameworks considered in this study.

3.1. Temporal Frameworks. Temporal causal discovery frameworks determine the presence of causal connections between variables by assessing whether the past values of one variable can predict the future values of another. Consider a pair of variables v_i and v_j that we want to test for possible causal connection, and the set of variables $Z \subseteq \mathcal{V}\backslash\{v_i, v_j\}$ (where the operation $\mathcal{V}\backslash\{v_i, v_j\}$ indicates the removal of $\{v_i, v_j\}$ from $\mathcal{V}$) that can potentially contain information regarding v_i and v_j. Suppose that x_i, x_j, and $X_Z = X_{rem(ij)} \in \mathbb{R}^{(N-2)\times T}$ are their corresponding observed time series variables.

We are interested in identifying whether v_j temporally causes v_i while accounting for the information provided by Z by analyzing their corresponding time series variables x_i, x_j, and X_Z. The F-test is often used to measure the significance of a temporal causal relation.[41] Consider

$$F_{v_j\to v_i|Z} = \delta_t(x_{i,t}, x_{j,t-1:t-p}|x_{i,t-1:t-p}, X_{Z,t-1:t-p}) \tag{1}$$

where $F_{v_j\to v_i|Z}$ represents the confidence value of the causal relation $v_j \to v_i$ conditioned on Z. $x_{i,t}$ represents the value of x_i at time t while $x_{i,t-1:t-p}$, $x_{j,t-1:t-p}$, and $X_{Z,t-1:t-p}$ represent the values of x_i, x_j and X_Z from time $t-1$ to $t-p$. The value p is often known as the "order" and is typically far smaller than the overall length of the time series. Here, $\delta_t(\cdot|\cdot)$ is a scalar measure of the predictive significance of $x_{j,t-1:t-p}$ over $x_{i,t}$ conditioned on $x_{i,t-1:t-p}$ and $X_{Z,t-1:t-p}$. Generally, an $F_{v_j\to v_i|Z}$ value approaching zero indicates that the past values of v_j are irrelevant to future values of v_i, and an $F_{v_j\to v_i|Z}$ value approaching one indicates that the past values of v_j are significant to future values of v_i. A threshold value α is typically used to generate the final causal graph, where if $F_{v_j\to v_i|Z} > \alpha$, then $v_j \to v_i \in \vec{\mathcal{E}}$.

The specific measure of the scalar $\delta_t(\cdot|\cdot)$ varies across different temporal causal discovery frameworks. For one of the most widely used temporal causal discovery methods, the *Granger Causality* (GC), the scalar $\delta_t(x_{i,t}, x_{j,t-1:t-p}|x_{i,t-1:t-p}, X_{Z,t-1:t-p})$ reduces to the ratio of variances of the residuals of two *Vector Auto Regression* (VAR) models for $x_{i,t}$:

$$\delta_t(x_{i,t}, x_{j,t-1:t-p}|x_{i,t-1:t-p}, X_{Z,t-1:t-p}) = 1 - \frac{var_t(\epsilon_{UM}(t))}{var_t(\epsilon_{RM}(t))} \tag{2}$$

where $\epsilon_{UM}(t)$ is the residual of the VAR model for $x_{i,t}$ that accounts for $x_{j,t-1:t-p}$, $x_{i,t-1:t-p}$, and $X_{Z,t-1:t-p}$, also known as the *unrestricted model*, and $\epsilon_{RM}(t)$ is the residual of the VAR model for $x_{i,t}$ that accounts for only $x_{i,t-1:t-p}$ and $X_{Z,t-1:t-p}$, also known as the *restricted model*. $\text{var}_t(\epsilon_{UM})$ and $\text{var}_t(\epsilon_{RM})$ are the variances of both residuals across time. We see that if v_j is temporally relevant to v_i, then $\text{var}_t(\epsilon_{UM}(t)) \ll \text{var}_t(\epsilon_{RM}(t))$, leading to $F_{v_j\to v_i|Z} \to 1$ which implies $x_j \to x_i$.[42] Given that GC utilizes VAR models in its statistical measure, stationarity is a necessary feature of the observed time series for GC to identify temporal causal relations accurately. We can see that the set of conditioning variables Z plays an important role in ruling out spurious causal predictions by accounting for potential intermediate variables in Z between v_i and v_j. For example, consider three variables v_1, v_2, and v_3, and let x_1, x_2, and x_3 be the corresponding observed time series variables. Suppose that the underlying process is given by $x_1(t) = e_1(t)$, $x_2(t) = x_1(t-1) + e_2(t)$, and $x_3(t) = x_2(t-1) + e_3(t)$, corresponding to a true causal graph $v_1 \to v_2 \to v_3$. GC will predict a spurious causal link $v_1 \to v_3$ if v_2 is not accounted as a conditioning variable. The computational complexity of checking for pairwise Granger causality in a process with N variables, T samples, and a model order p is $O(NT^3p^3)$.[43] However, the problem becomes NP-hard when applied to processes with multiple time lags since it requires the user to check for multiple model orders p.

3.2. SEM-Based Methods. Unlike temporal frameworks, causal discovery frameworks based on the *Structural Equation Model* (SEM) lack a temporal element in their basic definition of causality. Interactions between variables in SEMs appear as instantaneous. Given the observational matrix $X \in \mathbb{R}^{N\times T}$, consider the linear SEM of the form

$$x(t) = Ax(t) + e(t) \tag{3}$$

where $x(t) \in \mathbb{R}^{N\times 1}$ is a vector of N observed variables corresponding to the t^{th} column of X, $e(t) \in \mathbb{R}^{N\times 1}$ is a vector of stationary independent random noise, and $A \in \mathbb{R}^{N\times N}$ is the coefficient matrix that maps the causal relation between elements of $\mathcal{V}$ with $A_{ii} = 0$ for all $i = 1, \ldots, N$ (no self-loops).

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543

In other words, A directly encodes a true directed causal graph $\vec{G}(\mathcal{V}, \vec{\mathcal{E}})$, where $v_j \rightarrow v_i \in \vec{\mathcal{E}}$ if $A_{ij} \neq 0$. In general, a SEM does not need to be linear. However, in this section, we will restrict our discussion only to data generated from eq 3.

Constraint-based algorithms can recover the causal structure of A by identifying conditional independence relationships between the observed variables in x. We can summarize constraint-based algorithms into two general steps: 1) generation of the skeleton graph by pruning the edges between conditionally independent variables using conditional independence tests, and 2) orientation of the skeleton graph. The orienting strategy of most constraint-based algorithms generally builds around identifying v-structures entailed by the inferred conditional independence relations in the skeleton. The remaining edges are then generally oriented by avoiding adding new conditional independence relations not entailed previously. For data with continuous variables and linear (or close to linear) relationships (such as the SEM), the *partial correlation* can be used to identify conditional independence relationships. For data with continuous variables and nonlinear relationships, kernel-based independence tests and information-theoretic measures such as conditional mutual information can be used to identify conditional independence relationships.[44]

One example of a constraint-based algorithm is the *Peter Clark* (PC) algorithm.[40] Given the skeleton graph inferred from conditional independence testing, the PC algorithm orients the undirected edges in the skeleton based on the graphical notion of independence known as *d-separation*, where an adjacent triplet $v_i - v_k - v_j$ in the skeleton is oriented as a v-structure $v_i \rightarrow v_k \leftarrow v_j$ if v_i and v_j are disconnected and v_k is not a member of the conditioning set that renders v_i and v_j independent. The remaining directed edges are oriented by avoiding new v-structures and cycles. The PC algorithm returns a *Complete Partially Direced Acyclic Graph* (CPDAG), a causal graph containing both undirected−and directed → edges. The undirected edges represent edges whose direction cannot be determined by the PC algorithm, entailing a class of graphs known as the *Markov Equivalence Class* (MEC) of DAGs. Given the sufficient size of X generated by eq 3 with an A that entails a DAG and assuming that there are no hidden confounding variables affecting x, the PC algorithm guarantees to return the correct MEC given that we infer the correct skeleton graph. The details of the PC algorithm are given in Algorithm 1 in Section 1 of the Supporting Information document.

There are several powerful extensions of the PC algorithm that account for potential latent or unobserved confounding variables, such as *Fast Causal Inference* (FCI).[40] FCI utilizes a more general notion of graphical independence, known as *m-separation*, and a graphical representation containing edge types beyond undirected and directed edges known as a *Partial Ancestral Graph* (PAG) that can capture latent confounders in an acyclic causal structure. We say that a variable C is a latent confounder between v_1 and v_2, if C is a fork element between v_1 and v_2, and C is not directly observable (e.g., latent). Unlike CPDAG (the output graph of the PC algorithm), a PAG encodes ancestral relations between the variables instead of direct parental relationships (note that all parental relations are also ancestral, but not all ancestral relations are parental). In a PAG, there are three kinds of edge points: > , -, and ∘. An edge $v_1 \rightarrow v_2$ means that v_1 is an ancestor to v_2, and v_2 is not an ancestor of v_1, implying that v_2 is a noncause of v_1. An edge $v_1 - v_2$ means that both v_1 and v_2 are ancestors of each other, implying that v_1 and v_2 are either direct or indirect causes of each other. Finally, the edge $\leftrightarrow$ implies that neither v_1 nor v_2 are ancestors of each other, implying that there exist one or more latent confounders as a common cause for both v_1 and v_2. A ∘ on either side of an edge implies that the algorithm cannot tell the proper edge point assignment. Similar to CPDAG, a PAG also represents a class of equivalent causal graphs and is theoretically guaranteed to return the correct class graphs given the correct skeleton graph and that the faithfulness assumption holds true. Details of the FCI algorithm are given in Algorithm 2 in Section 1 of the Supporting Information document. The NP-hard complexity skeleton discovery step of the constraint-based algorithm generally dominates the computational cost of the orientation step, as we must check for all possible combinations of conditioning variables for every conditional independence test of all possible pairs of variables. More efficient (but sometimes less informative) extensions of the aforementioned algorithms exist for larger but sparser graphs.[45−47] Specific properties of the dynamical interconnection can also be leveraged to improve the efficiency of uncovering the skeletal structure of a process.[48]

If the underlying process noise e in eq 3 is independent and non-Gaussian, the *Linear Non-Gaussian Acyclic Model* (LiNGAM) framework can be leveraged to identify causal relations between elements of $\mathcal{V}$ (rows of X).[49] Given X generated by eq 3 with non-Gaussian process noise, X can be decomposed into its independent components $X = BS$ using the *Independent Component Analysis* (ICA), where $S \in \mathbb{R}^{N\times T}$ is a matrix whose rows correspond to independent components of X that is mapped by the matrix $B \in \mathbb{R}^{N\times N}$. By drawing an analogy to eq 3, we can obtain an estimate for A from $X = BS$ by rescaling and permuting the rows of $I - B^{-1}$ given that A corresponds to a DAG. The computational complexity of LiNGAM is generally dominated by the ICA decomposition step. Typically, the more efficient fast-ICA is preferred over the ICA to handle larger data sets or high-dimensional data, reducing the computation complexity from around $O(N^cT)$, where $c \in [2, 3]$, to $O(NTr)$, where r is the number of iterations required for convergence and is usually significantly smaller than N.[50]

3.3. Hybrid Methods. Hybrid frameworks account for both lagged and instantaneous causal effects. An example of a hybrid framework can be found in,[51] where the authors incorporate a framework to identify conditional independence relations of dynamically interconnected variables known as the *Wiener-separation*. The measure is derived from the mathematics of Wiener filtration and is coupled with the PC algorithm, allowing the application of the PC algorithm and other constraint-based approaches to temporal systems. The *PC-Momentary Conditional Conditional Independence* (PCMCI) is a similar framework that accounts for lagged causal effects within the scope of the PC algorithm.[52] These methods open the way for applying more powerful and robust constraint-based methods, such as the FCI, to temporal systems.

The LiNGAM framework can also be extended to account for lagged causal relations. This extension is the *Vector Autoregressive Linear Non-Gaussian Acyclic Model* (VAR-LiNGAM). VAR-LiNGAM assumes a causal model described by the *Structural Vector Autoregression* (SVAR) model

$$x(t) = A_0 x(t) + \sum_{\tau=1}^{P} A_\tau x(t-\tau) + e(t) \tag{4}$$

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543

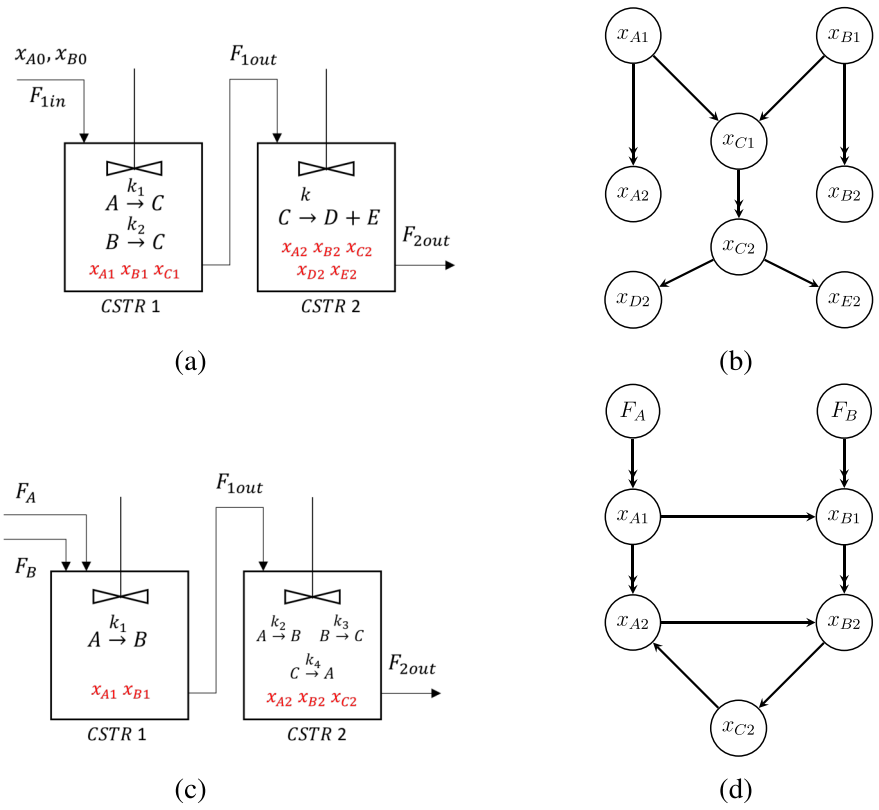


Figure 3. a) Process schematic of the first case study. b) True causal graph of the first case study. c) Process schematic of the second case study. d) True causal graph of the second case study.

where $A_0 \in \mathbb{R}^{N\times N}$ and $A_\tau \in \mathbb{R}^{N\times N}$ represent the instantaneous and lagged interactions between rows of x. Similarly, the rows of x in eq 4 correspond to a set of variables $\mathcal{V}$. The first term of eq 4 can be interpreted as a SEM defined by eq 3. Given the observational matrix X generated by eq 4, VAR-LiNGAM solves for an estimate of A_0 and A_τ by initially estimating a VAR model from x using the least-squares method and then applying LiNGAM to the residuals of the VAR model. Like LiNGAM, VAR-LiNGAM assumes that the underlying process noise is non-Gaussian and that A_0 corresponds to a DAG. A more efficient approach that yields the same result utilizes the *Multichannel blind deconvolution* to decompose x into dynamically independent components.[49] Similar to VAR-LiNGAM, the *extended Granger causality* is a variant of the Granger causality that determines the direction of instantaneous causal effects by leveraging the asymmetry of the residuals induced by non-Gaussian process noise using ICA.[53] Hybrid algorithms show NP-hard complexity with increasing numbers of variables, observations, and time-lags present in the process. Notably, the Wiener separation not only requires checking for all possible combinations of conditioning variables for every conditional independence test of all possible pairs of variables but also across all time lags up to the model order of the Wiener filter. The computational cost of hybrid methods can be significantly larger than that of either temporal frameworks or SEM-based frameworks.

4. CASE STUDIES AND PROCESS SIMULATION

We will consider two case studies (one with an acyclic structure and another containing a cyclic structure) to illustrate and analyze the effect of several nonideal observation factors on the reliability of the different causal discovery frameworks reviewed in the previous section.

4.1. Case Study 1. This case study considers two *Continuous Stirred Tank Reactor* (CSTR) in series with first-order reactions, operating at steady state. The process schematic is in Figure 3a. A feed stream with equal concentrations of A and B is fed into CSTR 1, where A and B react to form the intermediate product C. The exit stream is fed into CSTR 2, where C further reacts to form D and E.

The process can be modeled by the following delayed differential equation system

$$
\begin{aligned}
\frac{dx_{A1}}{dt} &= \frac{F_{1in}}{V_1}x_{A0} - k_1 x_{A1}(t) - \frac{F_{1out}}{V_1}x_{A1}(t)\\
\frac{dx_{B1}}{dt} &= \frac{F_{1in}}{V_1}x_{B0} - k_2 x_{B1}(t) - \frac{F_{1out}}{V_1}x_{B1}(t)\\
\frac{dx_{C1}}{dt} &= k_1 x_{A1}(t) + k_2 x_{B1}(t) - \frac{F_{1out}}{V_1}x_{C1}(t)\\
\frac{dx_{A2}}{dt} &= \frac{F_{1out}}{V_1}x_{A1}(t - \tau_{F_{1out}}) - \frac{F_{2out}}{V_2}x_{A2}(t)\\
\frac{dx_{B2}}{dt} &= \frac{F_{1out}}{V_1}x_{B1}(t - \tau_{F_{1out}}) - \frac{F_{2out}}{V_2}x_{B2}(t)\\
\frac{dx_{C2}}{dt} &= \frac{F_{1out}}{V_1}x_{C1}(t - \tau_{F_{1out}}) - k_3 x_{C2}(t) - \frac{F_{2out}}{V_2}x_{C2}(t)\\
\frac{dx_{D2}}{dt} &= k_3 x_{C2}(t) - \frac{F_{2out}}{V_2}x_{D2}(t)\\
\frac{dx_{E2}}{dt} &= k_3 x_{C2}(t) - \frac{F_{2out}}{V_2}x_{E2}(t)
\end{aligned}
\tag{5}
$$

where x_{A1}, x_{B1}, and x_{C1} are the concentrations of A, B, and C in CSTR 1, and x_{A2}, x_{B2}, x_{C2}, x_{D2}, and x_{E2} are the concentrations of

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543

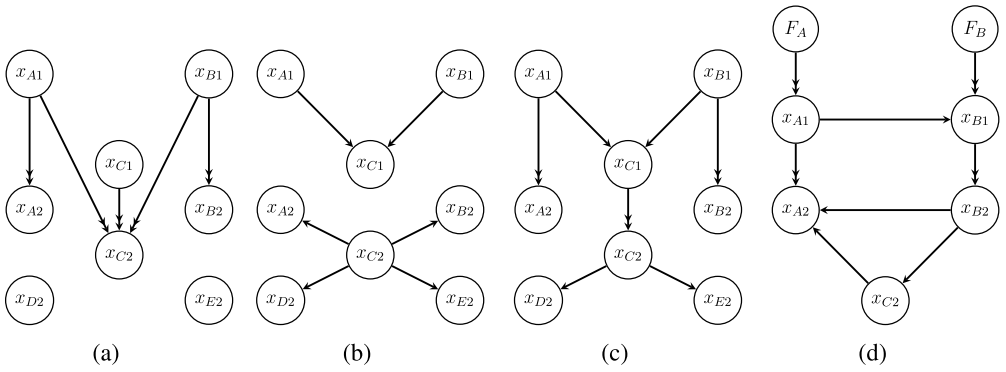


Figure 4. Causal graphs predicted by a) GC, b) LiNGAM, and c) VAR-LiNGAM for the first case study. d) A causal graph predicted by VAR-LiNGAM for the second case study.

A, B, C, D, and E in CSTR 2. V_1 and V_2 are the volumes of CSTR 1 and 2, respectively. F_{1in}, F_{1out}, and F_{2out} are the flow rates into CSTR 1, out of CSTR 1, and out of CSTR 2, respectively. The inlet flow rate of CSTR 1 F_{1out} contains equal concentrations of A and B, $x_{A0} = 0.5$ and $x_{B0} = 0.5$. The coefficients k_1 and k_2 are the first-order reaction rate constants for the reaction of A and B into C in CSTR 1, and k_3 is the first-order rate constant for the reaction of C to produce D and E in CSTR 2. The time delay $\tau_{F_{1out}}$ represents the transport delay through the pipe between the two CSTRs. Table S1 in Section 2 of the Supporting Information document summarizes the constants used in eq 6.

4.2. Case Study 2. The process shown in Figure 3c consists of two CSTRs in series with a first-order reaction network. Two separate feeds containing pure A and B are fed into CSTR 1, where A reacts to form B. The product stream of CSTR 1 containing A and B is fed into CSTR 2, where A reacts to form B, B reacts to form C, and C reacts back to form A, inducing a cyclic reaction network.

The process can be modeled as the following delayed differential equation system

$$\begin{aligned}
\frac{dx_{A1}}{dt} &= \frac{x_{A0}}{V_1}F_A(t-\tau_{F_A}) - k_1x_{A1}(t) - \frac{F_{1out}}{V_1}x_{A1}(t)\\
\frac{dx_{B1}}{dt} &= \frac{x_{B0}}{V_1}F_B(t-\tau_{F_B}) + k_1x_{A1}(t) - \frac{F_{1out}}{V_1}x_{B1}(t)\\
\frac{dx_{A2}}{dt} &= \frac{F_{1out}}{V_1}x_{A1}(t-\tau_{F_{1out}}) + k_4x_{C2}(t) - k_2x_{A2}(t) - \frac{F_{2out}}{V_2}x_{A2}(t)\\
\frac{dx_{B2}}{dt} &= \frac{F_{1out}}{V_1}x_{B1}(t-\tau_{F_{1out}}) + k_2x_{A2}(t) - k_3x_{B2}(t) - \frac{F_{2out}}{V_2}x_{B2}(t)\\
\frac{dx_{C2}}{dt} &= k_3x_{B2}(t) - k_4x_{C2}(t) - \frac{F_{2out}}{V_2}x_{C2}(t)
\end{aligned} \tag{6}$$

F_A and F_B are the measured inlet flow rates of pure A and B ($x_{A0} = x_{B0} = 1$) into CSTR 1. The variables x_{A1} and x_{B1} are the concentrations of A and B in CSTR 1, and x_{A2}, x_{B2}, and x_{C2} are the concentrations of A, B, and C in CSTR 2. V_1 and V_2 are the volumes of CSTR 1 and 2, respectively. The coefficient k_1 is the first-order rate constant for the reaction $A \rightarrow B$, while the coefficients k_2, k_3, and k_4 are the first-order rate constants for the reactions from A to B, B to C and C to A, respectively. The time delay constants τ_{F_A}, τ_{F_B}, and $\tau_{F_{1out}}$ represent material transport delay through pipes at both feed inlets of CSTR 1 and the pipe connecting CSTR 1 and 2. Table S1 in Section 2 of the Supporting Information document summarizes the constants used in eq 6.

4.3. Process Simulation and True Causal Graphs. To simulate eq 5 and (6) at steady state, the values of x are updated by solving for $x(t)$

$$f(x(t), x(t-\tau)) + e(t) = 0 \tag{7}$$

where f corresponds to the right-hand-side of eqs 5 and 6, and $e(t) \in \mathbb{R}^N$ is process noise sampled from the uniform distribution with a standard deviation of 0.05. Each process will be simulated for T time points. Concatenating subsequent state vectors $x(t)$ for $t = 0, ..., T$ yields the observation matrix $X \in \mathbb{R}^{N\times T}$. For simplicity, we will also let all $\tau = 1$.

The state vector x in eq 7 corresponds to the set of nodes $\mathcal{V}$. For the first case study, we define the state vector as $x_1 = [x_{A1}, x_{B1}, x_{C1}, x_{A2}, x_{B2}, x_{C2}, x_{D2}, x_{E2}]^T$ and the corresponding set of nodes $\mathcal{V}_1 = \{x_{A1}, x_{B1}, x_{C1}, x_{A2}, x_{B2}, x_{C2}, x_{D2}, x_{E2}\}$. For the second case study, we define the state vector as $x_2 = [F_A, F_B, x_{A1}, x_{B1}, x_{A2}, x_{B2}, x_{C2}]^T$ and the corresponding set of nodes $\mathcal{V}_2 = \{F_A, F_B, x_{A1}, x_{B1}, x_{A2}, x_{B2}, x_{C2}\}$. Note that we included F_A and F_B in x_2 and $\mathcal{V}_2$ as variables, where their dynamics are simply a static noisy process sampled from the uniform distribution with a standard deviation of 0.05 centered around their nominal values F_A^* and F_B^*.

The true causal graph for both case studies can be constructed by identifying the functional dependencies of the states in eq 5 and eq 6. We adopt the multiarrowed graph representation used in.[54] We have the causal relation $v_j \rightarrow v_j \in \vec{\mathcal{E}}$ if $\frac{dx_i}{dt} = f_i(\ldots, x_j(t), \ldots)$, i.e. the dynamics of x_i explicitly depend on x_j. We have the causal relation $v_j \rightarrow\rightarrow v_j \in \vec{\mathcal{E}}$ if $\frac{dx_i}{dt} = f_i(\ldots, x_j(t-\tau), \ldots)$, i.e. the dynamics of x_i explicitly depend on the τ-lagged values of x_j. The true causal graphs for both case studies, $G_1^{true}(\mathcal{V}_1, \vec{\mathcal{E}}_1)$ and $G_2^{true}(\mathcal{V}_2, \vec{\mathcal{E}}_2)$ are shown in Figures 3b and 3d.

5. RESULTS AND DISCUSSION

In this section we demonstrate the application of the aforementioned causal discovery frameworks to the process models described in Section 4. We illustrate how spurious predictions can arise if both instantaneous and lagged causal relations are not considered and challenges that 1) temporal aggregation, 2) temporal subsampling, and 3) the presence of latent confounders introduce into the process of causal discovery and how they can be addressed by exploiting the

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543

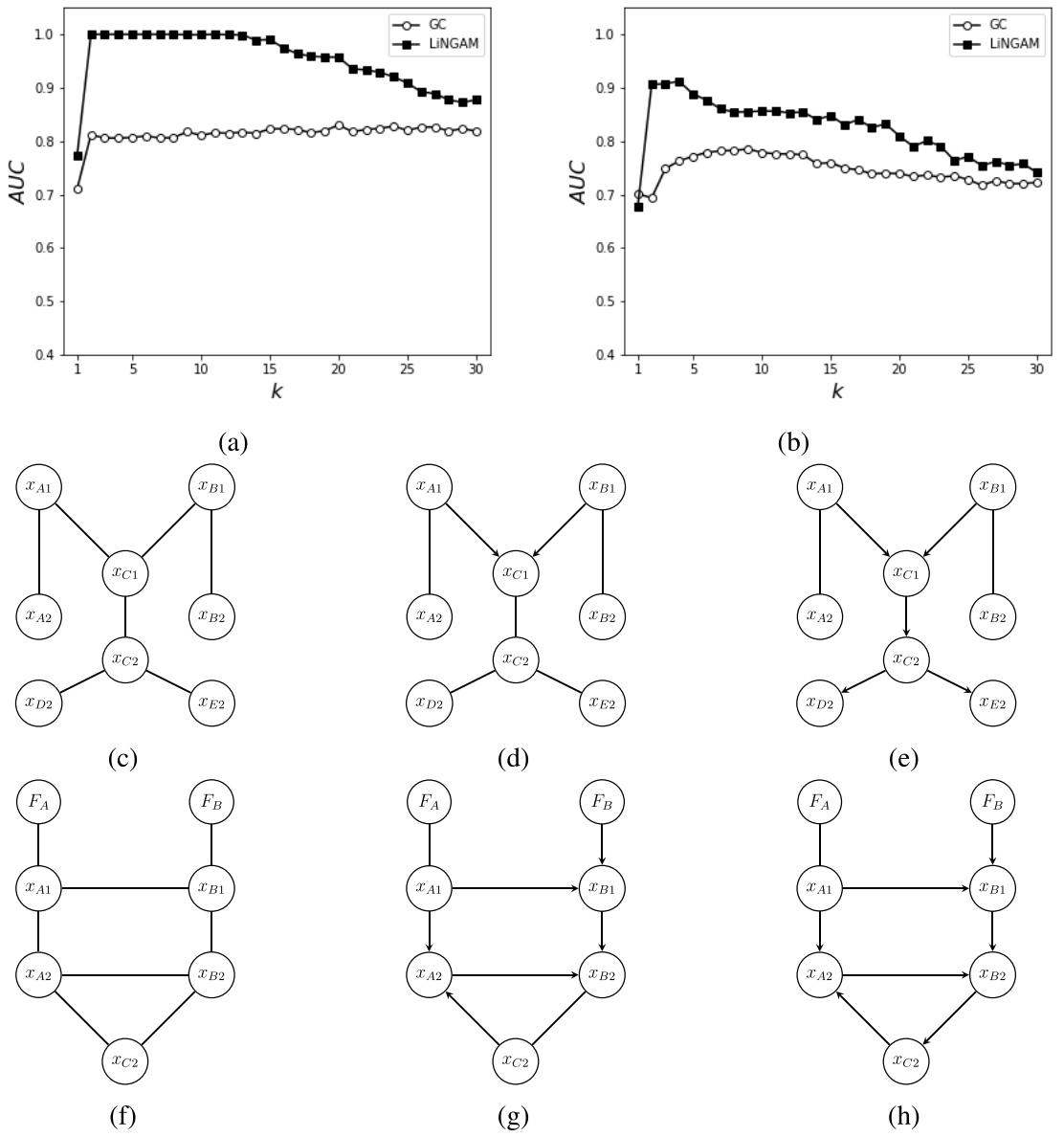


Figure 5. a) AUC vs k curves for GC and LiNGAM when applied to the first case study. b) AUC vs k for GC and LiNGAM when applied to the second case study. c) Predicted skeleton graph of G_1^{true} generated by thresholding a sample partial correlation coefficient inferred from the simulation of the first case study at $k = 30$. d) V-structure graph of the first case study. e) Output CPDAG for the first case study. f) Predicted skeleton graph of G_2^{true} generated by thresholding a sample partial correlation coefficient inferred from the simulation of the second case study at $k = 30$. g) V-structure graph of the second case study. h) Output CPDAG for the second case study.

strengths of different causal discovery frameworks. Furthermore, we will address the general challenges of causal discovery of cyclic processes in combination with these nonideal process observations.

5.1. Simultaneous Consideration of Instantaneous and Lagged Causal Effects. In real processes, different variables can evolve and interact at different time scales.[55] For example, due to the sheer size of a reactor, temperature variation and mixing of the fluid inside the reactor may occur at a much slower rate than the chemical reactions occurring within the vessel.[56] Such processes can lead to different rates of interactions between variables of the system, leading to a mix of lagged and (essentially) instantaneous causal relations. In our case studies, we model the difference in reaction rates and material transport rates by using instantaneous causal relations between reactive components in either CSTR and lagged causal relations between variables in different CSTRs to account for transport delays.

We simulate the corresponding process models to generate time series data with $T = 100,000$ time points and then apply GC, LiNGAM, and VAR-LiNGAM. Figure 4a presents a causal graph predicted by GC for the first case study. We see that GC can only correctly capture lagged causal connections found in the first case study ($x_{A1} \to\to x_{A2}$, $x_{B1} \to\to x_{B2}$, and $x_{C1} \to\to x_{C2}$). Furthermore, it also predicts the spurious lagged causal links $x_{A1} \to\to x_{C2}$ and $x_{B1} \to\to x_{C2}$. These spurious causal predictions are induced by the causal links $x_{A1} \to x_{C1}$ and $x_{B1} \to x_{C1}$. The GC overlooks these instantaneous causal links, resulting in a false

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543

perception of a lagged causal connection between x_{A1} and x_{B2}, and x_{C2}.

Figure 4b presents a causal graph predicted by LiNGAM for the first case study. As expected, LiNGAM can only detect instantaneous causal links ($x_{A1} \to x_{C1}$, $x_{B1} \to x_{C1}$, $x_{C2} \to x_{D2}$, and $x_{C2} \to x_{E2}$). However, LiNGAM also predicts two spurious causal relations ($x_{C2} \to x_{A2}$ and $x_{C2} \to x_{B2}$). These spurious predictions result from the confounding effect of x_{A1} and x_{B1}. In G_1^{true}, the information coming from x_{A1} to both x_{A2} and x_{C2} arrives simultaneously, causing the appearance of an instantaneous causal link between x_{A2} and x_{C2}. This also applies to the other spurious connection $x_{C2} \to x_{B2}$.

These results highlight the problem of accounting for instantaneous and lagged effects separately. Hybrid methods, such as VAR-LiNGAM, can account for instantaneous and lagged effects simultaneously. Figure 4c represents the causal graph predicted by VAR-LiNGAM. We can see that VAR-LiNGAM returns a perfect reconstruction of the true causal structure for the first case study. However, VAR-LiNGAM is limited to identifying processes that contain acyclic instantaneous causal links. Figure 4d represents the causal graph predicted by the VAR-LiNGAM for the second case study, where we can observe the spurious causal prediction $x_{B2} \to x_{A2}$. Due to the instantaneous cycle $x_{A2} \to x_{B2} \to x_{C2} \to x_{A2}$, it becomes more difficult to distinguish the true causal directions as the instantaneous information propagation blends the causal sources and targets of the process. Therefore, similar to LiNGAM, VAR-LiNGAM requires the instantaneous portion of the true causal graph to be acyclic to recover the causal structure accurately. Other hybrid methods face similar limitations.[53,54]

5.2. Causal Discovery of Uniformly Aggregated Observed Data. Industrial historical data are often temporally aggregated for various purposes, such as data storage reduction, noise reduction, and ease of reporting.[57,58] Readings and queries from a network of sensors are spatiotemporally aggregated for ease of analysis and alignment.[59] However, such aggregation muddles temporal sequences and causal relations while also smoothing out fluctuations.[60] This section will address the effects of temporal aggregation on the accuracy of causal discovery frameworks and explore how SEM-based methods can be leveraged to identify causal relations from aggregated process data.

We can represent temporal aggregation as the averages of $k \in \mathbb{Z}$ consecutive, nonoverlapping sections of the true observations. This reduces the total size of the data set for storage but also produces lower-resolution observations:

$$\hat{x}(t) = \frac{1}{k}\sum_{i=1}^{k} x(i + (t-1)k) \tag{8}$$

Here, x is the true observation and $\hat{x}$ is the temporally aggregated observation. This aggregation will impact the performance of temporal causal discovery frameworks, such as the GC, as lagged effects between the variables are diluted with larger k. Gong et al. have proposed a method to identify the causal structure of a VAR-like process given an aggregation order k.[61] In addition to the systematic temporal aggregation of sensor data, temporal aggregation can also be induced by time constants in the process which are faster than the sensor's observation rate. In such a case, the value of k becomes uncertain due to the difference between the true process rate and the observation rate, limiting the reliability of such a framework. However, in,[61] it is also shown that if k is large enough, the data-generating process will resemble instantaneous causal relations found in SEM (eq 3). Thus, a large enough k allows us to obtain potentially reliable causal prediction using SEM-based causal discovery frameworks.

We applied uniform aggregation to 100 simulations of the processes described by eqs 5 and 6 where we consecutively aggregate k time points for $k \in [1, 30]$ ($k = 1$ refers to unaggregated simulations) to obtain a set of time series with $T = 100,000$ time points. We then apply LiNGAM and GC and analyze the effects of aggregation based on the performance of LiNGAM using GC as the reference point. We use *receiver operating characteristic* (ROC) curves to assess the accuracy of both causal discovery frameworks. The ROC curve illustrates the trade-off between the true and false positive rates for all possible threshold values α applied to the statistics described by eq 1. A true positive occurs when a framework correctly predicts a causal edge, and a false positive occurs when it predicts a causal edge while such an edge does not exist given α. The area under the ROC curve (AUC) is a typical metric to quantify the framework's performance, where an area of one indicates an ideal framework. Finally, we will also apply the PC algorithm to a sample of highly aggregated time series ($k = 30$) to illustrate its ability to recover the causal structure of a highly aggregated process.

In Figure 5a, we present the variations of AUC against different values of k for both LiNGAM and GC for the first case study. We can observe that LiNGAM returns perfect predictions of the causal relations for $k \in [2, 13]$. GC gains a slight improvement in its prediction as $k > 1$ due to the "leaking" of the information passed by instantaneous effects through averaging but does not gain significant improvement at larger k. As the process becomes more similar to a SEM process with higher k,[61] LiNGAM is a more reliable alternative than GC when aggregation is present in the observation of an acyclic process, as shown by the larger AUC within this range of k. However, similar to the results of VAR-LiNGAM in section 5.1, LiNGAM is not as reliable when applied to the simulations of the second case study. This is evident in Figure 5b, where, unlike the first case study, the AUC never approaches 1 at any value of k. This is due to the presence of cycles in the process. Note that, in the case of temporal aggregation, lagged cycles collapse into instantaneous cycles at higher k, which can lead to false causal predictions by temporal methods even when the true underlying system contains no instantaneous causal relations.

Within the scope of SEM-based methods, we can also leverage constraint-based algorithms to obtain a reliable prediction, even at large values of k where the process noise becomes closer to Gaussian. The partial correlation can be used to infer conditional independent relations between the aggregated variables to generate skeleton graphs. Similar to eq 1, we test for the presence of a causal influence between v_i and v_j by thresholding the following statistics

$$F_{v_i - v_j|\tilde{Z}} = \underset{\tilde{Z} \in \mathcal{V}\setminus\{v_i, v_j\}}{\text{minimize}}\ \rho(x_i, x_j \mid X_{\tilde{Z}}) \tag{9}$$

where $\rho(x_i, x_j|X_{\tilde{Z}}) = \rho(x_j, x_i|X_{\tilde{Z}})$ is the partial correlation between the vectors x_i and x_j conditioned on a set of vectors $X_{\tilde{Z}}$. However, unlike eq 1, we must search for the set $\tilde{Z} \in \mathcal{V}\setminus\{v_i, v_j\}$ so that $\rho(x_i, x_j|X_{\tilde{Z}})$ is minimized to obtain the measure $F_{v_i - v_j|\tilde{Z}}$. We say the undirected edge $v_i - v_j \in \mathcal{E}$ if $F_{v_i - v_j|\tilde{Z}} > \alpha$ for some α, where the set $\mathcal{E}$ is a set of undirected

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543

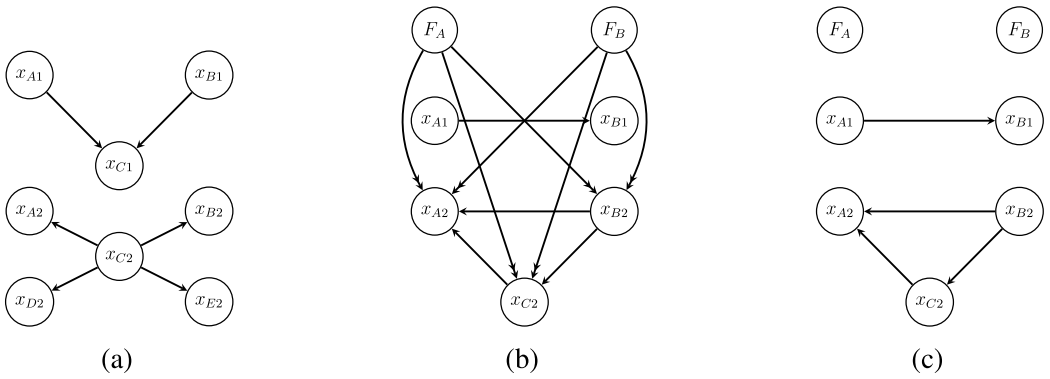


Figure 6. a) Multiarrowed causal graphs predicted by VAR-LiNGAM of the first case study when $k' = 2$. Multiarrowed causal graphs predicted by VAR-LiNGAM of the second case study when b) $k' = 2$ and c) $k' = 3$.

edges corresponding to a skeleton graph $G^{skel}(\mathcal{V}, \mathcal{E})$. Note that every disconnected pair $\{v_i, v_j\}$ entails a set of conditioning variables $\tilde{Z}$, which will be important when orienting the skeleton graph.

Let us consider the application of the PC algorithm to both case studies. We apply eq 9 to the aggregated observation matrix $\hat{X}$ from the simulation results of both case studies corresponding to $k = 30$ and threshold each value of $F_{v_i - v_j|\tilde{Z}}$ using $\alpha = 0.05$ to obtain the predicted skeleton graphs $G_1^{skel}(\mathcal{V}_1, \mathcal{E}_1)$ and $G_2^{skel}(\mathcal{V}_2, \mathcal{E}_2)$ shown in Figures 5c and 5f. The conditioning sets $\hat{Z}$ for each independent/disconnected pair of nodes are also shown in the same figure. We see that the partial correlation returns the correct skeleton graphs relative to G_1^{true} and G_2^{true}.

Consider the orientation process of G_1^{skel}. The orientation of v-structures inferred from the conditional independence relations, listed in Table S3, is shown in Figure 5d, where the only v-structure corresponds to $x_{A1} \rightarrow x_{C1} \leftarrow x_{B1}$ as $x_{C1} \notin \tilde{Z}$ corresponding to the conditionally independent pair x_{A1} and x_{B1} in Table S3. The final output CPDAG of the PC algorithm, denoted as $\hat{G}_1$, shown in Figure 5e was obtained after orienting the remaining edges in Figure 5d by avoiding new v-structures and cycles. We can see that the output CPDAG correctly recovers five causal relations with two remaining undirected edges in the acyclic G_1^{true}, even at $k = 30$ where the reliability of GC and LiNGAM are significantly compromised.

Using similar steps, the orientation of the v-structures in the skeleton using the conditional independence relations listed in Table S4 yields the partially directed graph in Figure 5g, where the inferred v-structures are $F_B \rightarrow x_{B1} \leftarrow x_{A1}$, $x_{A1} \rightarrow x_{A2} \leftarrow x_{C2}$, and $x_{B1} \rightarrow x_{B2} \leftarrow x_{A2}$. The final output CPDAG $\hat{G}_2$ shown in Figure 5h is obtained by orienting $x_{B2} \rightarrow x_{C2}$ to avoid the additional v-structure $x_{B1} \rightarrow x_{B2} \leftarrow x_{C2}$, leaving one edge being undirected $x_{A1} - x_{B1}$.

By relaxing the acyclicity assumption, the PC algorithm is able to recover most of the directed edges in G_2^{true} except for the causal relation $F_A \rightarrow x_{A1}$. This flexibility accommodates cyclic structures, extracting more information from inferred relations and broadening the identification of causal structures beyond DAGs. To guarantee maximum information extraction, constraint-based frameworks that are robust toward cycles utilize PAGs to express causal relations.[47,62,63] Details regarding PAGs and the orientation rules that follow are beyond the scope of this study.

5.3. Causal Discovery of Subsampled Observed Data.

In addition to aggregation, temporal subsampling also produces lower-resolution data and is commonly used in the storage of large process monitoring data sets. Similar to temporal aggregation, the use of subsampling is often driven by practical considerations such as the capabilities of sensors, the capacity of data storage systems, and the computational resources required for analysis. High-frequency observation of a process may generate an overwhelming amount of data, straining both computational capabilities and storage capacity. Hence, subsampling is a pragmatic approach to managing technological and resource constraints. However, it can also obscure important causal relationships that exist at finer temporal resolutions, leading to causal misprediction.

For a set of N observed variables $V = \{v_1, ..., v_N\}$, systematic subsampling of the corresponding time-series variables $X = [x(1), x(2), \ldots x(k'T)] \in \mathbb{R}^{N \times k'T}$ with length $k'T$ at the rate k' is defined as $\hat{X} = [\hat{x}(1), \hat{x}(2), \ldots, \hat{x}(T)] \in \mathbb{R}^{N \times T}$, where $\hat{x}(t) = x(1 + (t-1)k')$ for $t = 1, ..., T$ is the subsampled observation. Assuming that $x(t)$ is generated by a first-order VAR process, $x(t) = Ax(t-1) + e(t)$, the observed process becomes

$$\hat{x}(t+1) = A^{k'}\hat{x}(t) + \sum_{l=0}^{k'-1} A^l e(1 + k't - l) \tag{10}$$

The resulting observation of $\hat{x}(t)$ can induce misleading causal predictions by temporal methods such as GC. Theorem 1 in[64] states that given a large sample size T, the causal relations observed by GC will be given by the graph $G_{k'}(\mathcal{V}, \mathcal{E}_{k'})$, where $v_j \rightarrow v_i \in \mathcal{E}_{k'}$ if $\{A^{k'}\}_{ij} \neq 0$, which contains different directed edges from the true graph $G(\mathcal{V}, \vec{\mathcal{E}})$. Based on this theorem, there are several implications regarding the apparent causal structure. First, the causal structure of $A^{k'}$ for $k' > 1$ encodes ancestral causal relations instead of direct causal relations. Second, following the first implication, if the true adjacency matrix A corresponds to a DAG G with no self-loop (i.e., $A_{ii} = 0$), then $A^{k'}$ will be a zero matrix, leading to an empty graph $G_{k'}$, if k' that is equal to or larger than the longest path in G. Finally, specific structures of G can induce spurious instantaneous connections that can confound methods that account for instantaneous causal connections.[65] These implications complicate the inference of the underlying causal relations in G.

We simulate both case studies to obtain a time series of length $k'T$, where $T = 100,000$ for some k' ($k' = 1$ indicates no

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543

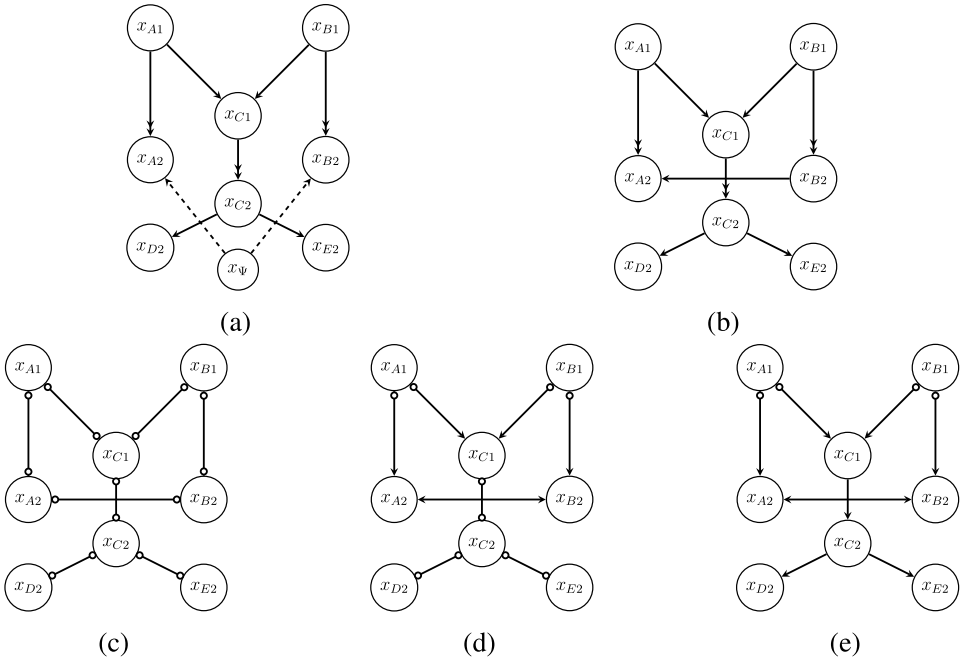


Figure 7. a) True causal graph of the first case study with a latent confounder x_{Ψ} between x_{A2} and x_{B2}. b) A causal graph predicted by GC. c) Skeleton graph of the first case study predicted by the Wiener separation. d) V-structure graph of the FCI algorithm. e) Predicted PAG of the first case study by the FCI algorithm.

subsampling) and apply eq 10 to each time series to obtain a subsampled observation matrix $\hat{X}_{k'}$. We then apply GC and VAR-LiNGAM to each $\hat{X}_{k'}$ and threshold the resulting F values using $\alpha = 0.05$. For conciseness, we will only display VAR-LiNGAM's predicted causal graphs since the results for GC are identical to the lagged causal relations predicted by VAR-LiNGAM.

Consider the results for the first case study. Figure 6a presents a causal graph predicted by VAR-LiNGAM with $k' = 2$. We see that neither VAR-LiNGAM nor GC predicts lagged causal relations. By observing G_1^{true}, we can see that by subsampling at the rate $k' > 1$, all single lag causal relations present in the system are not identifiable, causing the false negative predictions by VAR-LiNGAM and GC. Although the VAR-LiNGAM correctly predicts all the instantaneous causal links in G_1^{true}, it also predicts spurious instantaneous causal links $x_{A1} \rightarrow x_{C1}$ and $x_{A1} \rightarrow x_{C1}$. Note that this graph is identical to the predicted graph generated by LiNGAM when applied to the mixed lag process (Figure 4b). These spurious predictions can be explained using the same reasoning as the spurious predictions found in Figure 4b: The variables x_{A1} and x_{B1} have a confounding effect between x_{A2} and x_{C2}, and x_{B2} and x_{C2}, respectively. Applying hybrid methods to subsampled processes produces the same effect as employing SEM-based methods on the original process, as the lagged effects of the process are either lost due to the subsampling or not accounted for. Higher values of k' would yield the same causal graph in Figure 6a.

Now consider the results for the second case study, where an instantaneous cyclic structure is present. Figure 6b shows the causal graph predicted by VAR-LiNGAM for $k' = 2$. Similar to the results in Figure 4d, the VAR-LiNGAM predicts one spurious edge $x_{A2} \rightarrow x_{B2}$ since the VAR-LiNGAM can only accurately recover processes without instantaneous cycles. Due to the subsampling, the VAR-LiNGAM could not capture the single lag causal relations $F_A \rightarrow x_{A1}$, $F_B \rightarrow x_{B1}$, $x_{A1} \rightarrow x_{A2}$, and $x_{B1} \rightarrow x_{B2}$ in G_2^{true}. In addition, VAR-LiNGAM also falsely predicts lagged causal links from F_A and F_B to every member of the cycle $\{x_{A2}, x_{B2}, x_{C2}\}$ as each member of the cycle is two-time steps away from both F_A and F_B. A higher subsampling rate $k' > 2$ eliminates all lagged causal predictions by VAR-LiNGAM, as seen in the predicted causal graph in Figure 6c.

These results show that causal predictions of subsampled processes represent ancestral relations instead of parental relations. Therefore, unknowingly applying GC and VAR-LiNGAM to acyclic subsampled processes results in ancestral and false instantaneous causal relations that can be easily misinterpreted as direct causal links.

5.4. Causal Discovery in the Presence of Latent Confounders. In dealing with real systems, the complexity and limited knowledge of the process and materials lead to many influential variables contributing to process behavior that can not be directly measured or observed. When attempting to identify the causal structure of such systems, it is crucial to account for the potential existence of unobserved factors to avoid spurious causal predictions. In the causal discovery literature, such factors are known as latent/unobserved variables. In chemical processes, latent variables can take the form of unmeasured intermediate species concentrations in reaction networks, leaks, and exogenous factors such as feed impurities and environmental effects.

As discussed in section 3.2, the FCI algorithm is a viable SEM-based framework in the presence of latent confounders. Together with the notion of Wiener-separation, we can extend the application of the FCI algorithm to processes beyond SEMs. This is particularly useful when analyzing complex dynamical processes with potential latent confounders. To illustrate this, we will apply the FCI to a modified first case study with latent confounding variables using the Wiener separation.

Suppose that a compound Ψ is unknowingly introduced into CSTR 2 (potentially from remaining residue from previous reactor usage or impurities introduced from the inlet feed),

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543

where Ψ undergoes a first-order reaction to form both compounds A and B with rate constants k_4 and k_5, respectively. eq 5 then can be rewritten as

$$\begin{aligned}
\frac{dx_{A1}}{dt} &= \frac{F_{1in}}{V_1}x_{A0} - k_1x_{A1}(t) - \frac{F_{1out}}{V_1}x_{A1}(t)\\
\frac{dx_{B1}}{dt} &= \frac{F_{1in}}{V_1}x_{B0} - k_2x_{B1}(t) - \frac{F_{1out}}{V_1}x_{A1}(t)\\
\frac{dx_{C1}}{dt} &= k_1x_{A1}(t) + k_2x_{B1}(t) - \frac{F_{1out}}{V_1}x_{C1}(t)\\
\frac{dx_{A2}}{dt} &= \frac{F_{1out}}{V_1}x_{A1}(t-\tau_{F_{1out}}) - \frac{F_{2out}}{V_2}x_{A2}(t) + k_4x_\Psi(t-\tau_{k_4})\\
\frac{dx_{B2}}{dt} &= \frac{F_{1out}}{V_1}x_{B1}(t-\tau_{F_{1out}}) - \frac{F_{2out}}{V_2}x_{B2}(t) + k_5x_\Psi(t-\tau_{k_5})\\
\frac{dx_{C2}}{dt} &= \frac{F_{1out}}{V_1}x_{C1}(t-\tau_{F_{1out}}) - k_3x_{C2}(t) - \frac{F_{2out}}{V_2}x_{C2}(t)\\
\frac{dx_{D2}}{dt} &= k_3x_{C2}(t) - \frac{F_{2out}}{V_2}x_{D2}(t)\\
\frac{dx_{E2}}{dt} &= k_3x_{C2}(t) - \frac{F_{2out}}{V_2}x_{E2}(t)\\
\frac{dx_\Psi}{dt} &= C - k_4x_\Psi(t) - k_5x_\Psi(t)
\end{aligned} \tag{11}$$

where the C is a pseudoconstant source of Ψ in CSTR 2. Due to the difference in speed of formation, the formation of A and B from Ψ appears to be delayed at different time lags τ_{k_4} and τ_{k_5}, respectively. In this example, we let $\tau_{k_4} = 2$ and $\tau_{k_5} = 1$, with the other coefficients and time constants remaining the same. The corresponding causal graph for eq 11 is shown in Figure 7a, where x_Ψ is a latent confounder. We then apply GC and Wiener separation to 100 simulations of eq 11, each with $T = 100,000$ time points.

The predicted causal graph for GC is shown in Figure 7b, where we can observe a spurious causal link $x_{B2} \to x_{A2}$ predicted by GC. This spurious link is caused by x_Ψ imparting information earlier to x_{B2} than x_{A2}. Therefore, it is interpreted by GC as a lagged causal connection from x_{B2} to x_{A2} without being able to condition on x_Ψ. Similarly, the skeleton graph predicted by the Wiener separation also contains an undirected edge between x_{A2} and x_{B2}, as shown in Figure 7c. However, by applying the FCI algorithm, we can identify which causal edges correspond to the presence of confounding variables.

The FCI algorithm assigns undirected edges with ∘ on either end of the edges. Similar to the partial correlation, by applying the Wiener separation to process simulation of eq 11, we obtain a list of all conditionally independent pairs listed in Table S5 that is used to generate the skeleton graph shown in Figure 7c by removing edges between pairs that are conditionally independent. Next, by identifying v-structures from $\tilde{Z}$ of every pair of conditionally independent pairs in Table S5, we obtain the collider graph shown in Figure 7d, where $x_{A1} \to x_{C1} \leftarrow x_{B1}$, $x_{A1} \to x_{A2} \leftarrow x_{B2}$, and $x_{A2} \to x_{B2} \leftarrow x_{B1}$ are the only identified v-structures.

The remaining edges are oriented by avoiding additional v-structures to obtain the final output PAG shown in Figure 7e. From the final PAG, we can see that the FCI correctly predicts the location of the latent confounder x_Ψ, indicated by the edge $x_{A2} \leftrightarrow x_{B2}$. Unlike the CPDAG, PAGs allow more freedom in interpreting the information derived from conditional independence relations. If we were to apply the PC algorithm to the data set, the rigidity of a CPDAG would not allow the orientation $x_{A2} \leftrightarrow x_{B2}$ given the same skeleton since it only represents direct parental relationships. However, the generality of PAGs also comes with a cost. The remaining ∘ edge points imply that the FCI cannot tell the ancestral relation between the variables. Therefore, in the equivalence class for PAGs, the edge $x_{A1} \circ\!\!\to x_{A2}$ can either mean that a latent confounder is present in between them ($x_{A1} \leftrightarrow x_{A2}$) or that x_{A1} is a strict ancestor of x_{A2} ($x_{A1} \to x_{A2}$). Nevertheless, we have shown that the detection of latent confounders in dynamic processes is possible by implementing Wiener separation in the FCI algorithm.

6. CONCLUSIONS

In this study, we have introduced broad classes of causal discovery frameworks that can be applied to dynamical processes and addressed several factors that can impact their accuracy. Using two chemical process case studies with both instantaneous and lagged causal relations to emulate different time constants within the system, we have highlighted the importance of simultaneously accounting for instantaneous and lagged causal relations and quantified the performance of the different causal discovery frameworks under multiple nonideal observations: temporally aggregated and subsampled process data, and latent confounding variables. Our results provide several insights into how different frameworks can be leveraged appropriately under different nonideal observations:

- Simultaneously accounting for instantaneous and lagged causal relations is crucial to avoid spurious predictions.
- SEM-based methods are viable alternatives for the causal discovery of temporally aggregated processes.
- Causal discovery of subsampled processes reflects ancestral causal relations instead of direct causal relations.
- Subsampling of process data can result in the appearance of spurious instantaneous causal links, emphasizing the need for caution in interpreting predictions.
- Unobserved confounding variables can lead to misleading topology of the process. Using the Wiener separation, we were able to apply the FCI algorithm to identify unobserved confounders, which is invaluable when considering the complexity and vastness of industrial chemical processes.

A key characteristic of chemical process systems that is not addressed in this work is the often nonlinear and nonstationary data. Most algorithms used in this study are only suitable for linear processes and nonlinear processes whose dynamics are close to a steady state. Additionally, the identifiability of causal relations can present significant challenges in causal discovery. For example, weak causal relations can be easily weeded out even when low threshold values α are used to prune the causal graph. The effects of nonlinearity and nonstationarity, in combination with low-frequency observation, latent confounders, and instances of difficult-to-identify causal relations, will be studied in our future work using more complex benchmark systems to capture more insights on the application of causal discovery on real systems.

■ ASSOCIATED CONTENT

Data Availability Statement

Code is provided upon request.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.iecr.4c01155.

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543

Additional simulation details, algorithms, conditional independence test outputs, and figures (PDF)

AUTHOR INFORMATION

Corresponding Authors

Harman Dewantoro − *Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55414, United States*; Email: dewan055@umn.edu

Alexander Smith − *Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55414, United States*; Email: smi02527@umn.edu

Prodromos Daoutidis − *Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55414, United States*; orcid.org/0000-0003-4803-0404; Email: daout001@umn.edu

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.iecr.4c01155

Author Contributions

Harman Dewantoro, **Alexander Smith**: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing−original draft; **Prodromos Daoutidis**: Conceptualization, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing−review and editing

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Prof. Donatello Materassi and Prof. Murti Salapaka from the Department of Electrical and Computer Engineering at the University of Minnesota for their invaluable comments and insights throughout the study. Financial support from the National Science Foundation—CBET (award number 2313289) is gratefully acknowledged.

ACRONYMS

CPDAG	Complete Partially Direced Acyclic Graph.
CSTR	Continuous Stirred Tank Reactor.
DAG	Directed Acyclic Graph.
FCI	Fast Causal Inference.
GC	Granger Causality.
ICA	Independent Component Analysis.
LiNGAM	Linear Non-Gaussian Acyclic Model.
MEC	Markov Equivalence Class.
PAG	Partial Ancestral Graph.
PC	Peter Clark.
PCMCI	PC-Momentary Conditional Conditional Independence
SEM	Structural Equation Model
SVAR	Structural Vector Autoregression.
VAR	Vector Auto Regression.
VAR-LiNGAM	Vector Autoregressive Linear Non-Gaussian Acyclic Model.

GLOSSARY

causal discovery	Causal discovery is the process of identifying and understanding direct causeand-effect relationships between variables in a data set.
causal graph	A mathematical abstraction used to represent causal relationships between variables with nodes depicting variables and directed edges indicating causal influences.
constraint-based algorithms	A class of causal discovery algorithms that identifies causal relationships by testing for conditional independence between variables of the data.
d-separation	A notion of conditional independence in a graph by assessing the blocking of all paths between all the nodes.
directed edges	Arrows indicating the direction of causal influence from one node to another.
nodes	Element of a graph representing variables of the system. source Origin variable of the causal effect.
stationarity	Refers to the statistical property of a signal with constant mean and variance.
target	The affected variable.
v-structure	A collider structure where there are no edges between the parent nodes.

REFERENCES

(1) Thalmann, S.; Mangler, J.; Schreck, T.; Huemer, C.; Streit, M.; Pauker, F.; Weichhart, G.; Schulte, S.; Kittl, C.; Pollak, C.; Vukovic, M.; Kappel, G.; Gashi, M.; Rinderle-Ma, S.; Suschnigg, J.; Jekic, N.; Lindstaedt, S. Data Analytics for Industrial Process Improvement A Vision Paper. *2018 IEEE 20th Conference on Business Informatics (CBI)*. 2018; pp 92−96.

(2) Corona, F.; Mulas, M.; Baratti, R.; Romagnoli, J. A. On the topological modeling and analysis of industrial process data using the SOM. *Comput. Chem. Eng.* **2010**, *34*, 2022−2032.

(3) Haghighi, R.; Cheah, C. C. Topology-Based Controllability Problem in Network Systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **2017**, *47*, 3077−3088.

(4) Mousavi, S. S.; Haeri, M. Controllability analysis of networks through their topologies. *2016 IEEE 55th Conference on Decision and Control (CDC)*. 2016; pp 4346−4351.

(5) Xiang, L.; Chen, F.; Ren, W.; Chen, G. Advances in Network Controllability *IEEE Circuits and Systems Magazine*2019, *19*, 8−32.

(6) Knegtering, B.; Pasman, H. Safety of the process industries in the 21st century: A changing need of process safety management for a changing industry. *Journal of Loss Prevention in the Process Industries* **2009**, *22*, 162−168.

(7) Choley, J.-Y.; Mhenni, F.; Nguyen, N.; Baklouti, A. Topology-based safety analysis for safety critical CPS. *Procedia computer science* **2016**, *95*, 32−39.

(8) Wu, D.; Zhao, J. Process topology convolutional network model for chemical process fault diagnosis. *Process Safety and Environmental Protection* **2021**, *150*, 93−109.

(9) Dbouk, T. A review about the engineering design of optimal heat transfer systems using topology optimization. *Applied Thermal Engineering* **2017**, *112*, 841−854.

(10) Vuković, M.; Thalmann, S. Causal discovery in manufacturing: A structured literature review. *Journal of Manufacturing and Materials Processing* **2022**, *6*, 10.
(11) Thoben, K.-D.; Wiesner, S.; Wuest, T. "Industrie 4.0" and smart manufacturing-a review of research issues and application examples. *International journal of automation technology* **2017**, *11*, 4−16.
(12) Ljung, L. In *Signal Analysis and Prediction*; Procházka, A., Uhlíř, J., Rayner, P. W. J., Kingsbury, N. G., Eds.; Birkhäuser Boston: Boston, MA, 1998; pp 163−173.
(13) Rivera, D. E.; Lee, H.; Braun, M. W.; Mittelmann, H. D. " Plant-Friendly" system identification: a challenge for the process industries. *IFAC Proceedings Volumes* **2003**, *36*, 891−896.
(14) Favoreel, W.; De Moor, B.; Van Overschee, P. Subspace state space system identification for industrial processes. *Journal of process control* **2000**, *10*, 149−155.
(15) Brunton, S. L.; Proctor, J. L.; Kutz, J. N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 3932−3937.
(16) Ljung, L.; Andersson, C.; Tiels, K.; Schön, T. B. Deep Learning and System Identification. *IFAC-PapersOnLine* **2020**, *53*, 1175−1181. 21st IFAC World Congress
(17) Kamiński, M.; Ding, M.; Truccolo, W. A.; Bressler, S. L. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biological cybernetics* **2001**, *85*, 145−157.
(18) Sokolova, E.; Hoogman, M.; Groot, P.; Claassen, T.; Vasquez, A. A.; Buitelaar, J. K.; Franke, B.; Heskes, T. Causal discovery in an adult ADHD data set suggests indirect link between DAT1 genetic variants and striatal brain activation during reward processing. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **2015**, *168*, 508−515.
(19) Vicente, R.; Wibral, M.; Lindner, M.; Pipa, G. Transfer entropy-a model-free measure of effective connectivity for the neurosciences. *Journal of computational neuroscience* **2011**, *30*, 45−67.
(20) Shen, X.; Ma, S.; Vemuri, P.; Simon, G.; et al. Challenges and opportunities with causal discovery algorithms: application to Alzheimer's pathophysiology. *Sci. Rep.* **2020**, *10*, 2975.
(21) Weichwald, S.; Peters, J. Causality in cognitive neuroscience: concepts, challenges, and distributional robustness. *Journal of Cognitive Neuroscience* **2021**, *33*, 226−247.
(22) Mani, S.; Cooper, G. F. A study in causal discovery from population-based infant birth and death records. *Proceedings of the AMIA Symposium*. 1999; p 315.
(23) Michoel, T.; Zhang, J. D. Causal inference in drug discovery and development. *Drug discovery today* **2023**, *28*, 103737.
(24) Foraita, R.; Friemel, J.; Günther, K.; Behrens, T.; Bullerdiek, J.; Nimzyk, R.; Ahrens, W.; Didelez, V. Causal discovery of gene regulation with incomplete data. *Journal of the Royal Statistical Society Series A: Statistics in Society* **2020**, *183*, 1747−1775.
(25) Kelly, J.; Berzuini, C.; Keavney, B.; Tomaszewski, M.; Guo, H. A review of causal discovery methods for molecular network analysis. *Molecular Genetics & Genomic Medicine* **2022**, *10*, e2055.
(26) Huang, B.; Zhang, K.; Zhang, J.; Ramsey, J.; Sanchez-Romero, R.; Glymour, C.; Schölkopf, B. Causal Discovery from Heterogeneous/Nonstationary Data. *Journal of Machine Learning Research* **2020**, *21*, 1−53.
(27) Marcot, B. G.; Holthausen, R. S.; Raphael, M. G.; Rowland, M. M.; Wisdom, M. J. Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. *Forest ecology and management* **2001**, *153*, 29−42.
(28) Sugihara, G.; May, R.; Ye, H.; Hsieh, C.-h.; Deyle, E.; Fogarty, M.; Munch, S. Detecting causality in complex ecosystems. *science* **2012**, *338*, 496−500.
(29) Ombadi, M.; Nguyen, P.; Sorooshian, S.; Hsu, K.-l. Evaluation of Methods for Causal Discovery in Hydrometeorological Systems. *Water Resour. Res.* **2020**, *56*, No. e2020WR027251.
(30) Rodrigo, V.; Chioua, M.; Hagglund, T.; Hollender, M. Causal analysis for alarm flood reduction. *IFAC-PapersOnLine* **2016**, *49*, 723−728.
(31) Hu, W.; Wang, J.; Chen, T.; Shah, S. L. Cause-effect analysis of industrial alarm variables using transfer entropies. *Control Engineering Practice* **2017**, *64*, 205−214.
(32) Yu, W.; Yang, F. Detection of causality between process variables based on industrial alarm data using transfer entropy. *Entropy* **2015**, *17*, 5868−5887.
(33) Kumari, P.; Bhadriraju, B.; Wang, Q.; Kwon, J. S.-I. A modified Bayesian network to handle cyclic loops in root cause diagnosis of process faults in the chemical process industry. *Journal of Process Control* **2022**, *110*, 84−98.
(34) Chen, H.-S.; Yan, Z.; Yao, Y.; Huang, T.-B.; Wong, Y.-S. Systematic procedure for Grangercausality- based root cause diagnosis of chemical process faults. *Ind. Eng. Chem. Res.* **2018**, *57*, 9500−9512.
(35) Suresh, R.; Sivaram, A.; Venkatasubramanian, V. A hierarchical approach for causal modeling of process systems. *Comput. Chem. Eng.* **2019**, *123*, 170−183.
(36) Cao, L.; Su, J.; Wang, Y.; Cao, Y.; Siang, L. C.; Li, J.; Saddler, J. N.; Gopaluni, B. Causal discovery based on observational data and process knowledge in industrial processes. *Ind. Eng. Chem. Res.* **2022**, *61*, 14272−14283.
(37) Zhu, Q.-X.; Meng, Q.-Q.; Wang, P.-J.; He, Y.-L. Novel causal network modeling method integrating process knowledge with modified transfer entropy: a case study of complex chemical processes. *Ind. Eng. Chem. Res.* **2017**, *56*, 14282−14289.
(38) Yang, S.; Bequette, B. W. Observational process data analytics using causal inference. *AIChE J.* **2023**, *69*, No. e17986.
(39) Gharahbagheri, H.; Imtiaz, S.; Khan, F. Root cause diagnosis of process fault using KPCA and Bayesian network. *Ind. Eng. Chem. Res.* **2017**, *56*, 2054−2070.
(40) Spirtes, P.; Glymour, C.; Scheines, R. *Causation, Prediction, and Search*, 2nd ed.; MIT press, 2000.
(41) Lehmann, E. L.; Romano, J. P.; Casella, G. *Testing statistical hypotheses*; Springer, 1986; Vol. 3.
(42) Granger, C. W. J. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica* **1969**, *37*, 424−438.
(43) Lindner, B.; Auret, L.; Bauer, M.; Groenewald, J. W. Comparative analysis of Granger causality and transfer entropy to present a decision flow for the application of oscillation diagnosis. *Journal of Process Control* **2019**, *79*, 72−84.
(44) Zhang, K.; Peters, J.; Janzing, D.; Schölkopf, B. Kernel-based conditional independence test and application in causal discovery. *27th Conference on Uncertainty in Artificial Intelligence (UAI 2011) (Barcelona)*. 2011.
(45) Kalisch, M.; Bühlman, P. Estimating High-Dimensional Directed Acyclic Graphs with the PC-Algorithm. *Journal of Machine Learning Research* **2007**, *8*, 613.
(46) Colombo, D.; Maathuis, M. H.; Kalisch, M.; Richardson, T. S. Learning high-dimensional directed acyclic graphs with latent and selection variables. *Annals of Statistics* **2012**, *40*, 294−321.
(47) Zhang, J. Causal Reasoning with Ancestral Graphs. *Journal of Machine Learning Research* **2008**, *9*, 1437−1474.
(48) Talukdar, S.; Deka, D.; Doddi, H.; Materassi, D.; Chertkov, M.; Salapaka, M. V. Physics informed topology learning in networks of linear dynamical systems. *Automatica* **2020**, *112*, 108705.
(49) Shimizu, S. LiNGAM: Non-Gaussian methods for estimating causal structures. *Behaviormetrika* **2014**, *41*, 65−98.
(50) Hyvärinen, A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks* **1999**, *10*, 626−634.
(51) Materassi, D.; Salapaka, M. V. Reconstruction of directed acyclic networks of dynamical systems. *2013 American Control Conference*. 2013; pp 4687−4692.
(52) Runge, J. Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets. 2022.

(53) Schiatti, L.; Nollo, G.; Rossato, G.; Faes, L. Extended Granger causality: a new tool to identify the structure of physiological networks. *Physiological measurement* **2015**, *36*, 827.
(54) Dimovska, M.; Materassi, D. Control Theoretic Look at Granger Causality: Extending Topology Reconstruction to Networks With Direct Feedthroughs. *IEEE Transactions on Automatic Control* **2021**, *66*, 699−713.
(55) Fritzson, P. A. *Introduction to modeling and simulation of technical and physical systems with Modelica*; Wiley: Hoboken, N.J, 2011.
(56) Goryachev, A.; Strizhak, P.; Kapral, R. Slow manifold structure and the emergence of mixedmode oscillations. *J. Chem. Phys.* **1997**, *107*, 2881−2889.
(57) Jensen, S. K.; Pedersen, T. B.; Thomsen, C. Time Series Management Systems: A Survey. *IEEE Transactions on Knowledge and Data Engineering* **2017**, *29*, 2581−2600.
(58) Hoffmann, M.; Kotzur, L.; Stolten, D.; Robinius, M. A review on time series aggregation methods for energy system models. *Energies* **2020**, *13*, 641.
(59) Bonnet, P.; Gehrke, J.; Seshadri, P. Lecture Notes in Computer Science; (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) *Lecture Notes in Computer Science*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2001; Vol. *1987*; pp 3−14.
(60) Tiao, G. C. Asymptotic behaviour of temporal aggregates of time series. *Biometrika* **1972**, *59*, 525−531.
(61) Gong, M.; Zhang, K.; Schölkopf, B.; Glymour, C.; Tao, D. Causal discovery from temporally aggregated time series. Uncertainty in artificial intelligence. *Proceedings of the Conference on Uncertainty in Artificial Intelligence*. 2017.
(62) Spirtes, P. Directed Cyclic Graphical Representations of Feedback Models. CoRR 2013, abs/1302.4982.
(63) Richardson, T. A discovery algorithm for directed cyclic graphs. *Proceedings of Conference on Uncertainty and Artificial Intelligence (UAI-95)*; 1996, pp 491−499.
(64) Gong, M.; Zhang, K.; Schoelkopf, B.; Tao, D.; Geiger, P. Discovering temporal causal relations from subsampled data. *International Conference on Machine Learning*. 2015; pp 1898−1906.
(65) Danks, D.; Plis, S. Learning causal structure from undersampled time series. *JMLR: Workshop and Conference Proceedings NIPS Workshop on Causality)*; 2013; pp 1−10.

https://doi.org/10.1021/acs.iecr.4c01155
Ind. Eng. Chem. Res. 2024, 63, 11530−11543