

Discovering Mechanistic Causality from Time Series: A Behavioral-System Approach

Yingzhu Liu[†], Shengyuan Huang[†], Zhongkui Li, Xiaoguang Yang, Wenjun Mei*

Abstract—Identifying “true causality” is a fundamental challenge in complex systems research. Widely adopted methods, like the Granger causality test, capture statistical dependencies between variables rather than genuine driver-response mechanisms. This critical gap stems from the absence of mathematical tools that reliably reconstruct underlying system dynamics from observational time-series data. In this paper, we introduce a new control-based method for causality discovery through the behavior-system theory, which represents dynamical systems via trajectory spaces and has been widely used in data-driven control. Our core contribution is the Behavior-enabled Causality test (the BeCaus test), which transforms causality discovery into solving fictitious control problems. By exploiting the intrinsic asymmetry between system inputs and outputs, the proposed method operationalizes our conceptualization of mechanistic causality: variable X is a cause of Y if X (partially) drives the evolution of Y . We establish conditions for linear time-invariant systems to be causality-discoverable, i.e., conditions for the BeCaus test to distinguish four basic causal structures (independence, full causality, partial causality, and latent-common-cause relation). Notably, our approach accommodates open systems with unobserved inputs. Moreover, an exploratory case study indicates the new method’s potential extensibility to nonlinear systems.

I. INTRODUCTION

A. Background and Motivation

Establishing causal relations is both important and challenging in research on complex systems such as neuroscience [1], earth system science [2], [3], economics [4], and social science [5]. For these systems, controlled lab experiments either are infeasible or do not usually yield reliable results applicable to real-world scenarios. Therefore, inferring causality from observational data is fundamentally important. Ideally, one wishes to extract true causal relations, i.e., genuine driver-response mechanisms, from collected time-series data. However, current widely-adopted methods, such as the regression-based Granger causality test and other methods based on information theory or machine learning, are essentially testing statistical dependencies instead of mechanistic causality. Such limitation is due to the lack of

theoretical tools that reliably represent underlying dynamical systems using offline time-series data.

The behavior-system approach in control theory, which has been widely adopted in data-driven control in recent years, happens to provide such a tool and makes it possible to investigate causality in a more intrinsic and straightforward sense: A variable is a cause of another if the former (partially) drives the dynamic evolution of the latter. Based on the above conceptualization and the behavior-system theory, we propose in this paper a new method that discovers mechanistic causal directions from offline time-series data. This method leads to mathematically tractable results for Linear Time Invariant (LTI) systems, including open systems, i.e., the scenarios when some parts of the system inputs are unobserved. Moreover, exploratory numerical studies the potential of behavior-system-based method in discovering mechanistic causality in nonlinear systems.

B. Literature Review

Regression-based Granger causality test: One prominent branch of model-based causal inference is Granger causality [6]. Granger causality originates from bivariate linear autoregression [7] and evaluates whether the data of one time series improves the forecast of another. This method is mathematically rigorized by implementing a statistical hypothesis test [8]. Granger causality has been further extended to a graphic approach that visualizes causal relations among multiple variables by a directed graph: each time series is treated as a node, and the edges represent Granger-causal relations between variables [9]–[11]. Despite its wide application, a major limitation of the Granger causality test is that it requires the time series to be stationary [12]. Moreover, Granger causality is more of a statistical concept than what is normally understood as the true mechanistic causality.

Other model-free methods: Besides Granger causality tests, alternative frameworks have been proposed based on information theory [13]–[16] or machine learning [17], [18]. Information-theoretic measures, such as transfer entropy [13], [14], mutual information [16], [19] and directed information [15], are used to quantify time-directed information transfer between jointly dependent variables. These measures are particularly useful in analyzing complex systems where the conditions for the Granger causality test do not apply. In particular, transfer entropy is equivalent to Granger causality when the variables involved follow a Gaussian distribution [13]. The development of machine learning provides a new perspective for causal inference. Approaches like neural networks are applied to learn complex,

This work was supported in part by the National Science and Technology Major Project under grant 2022ZD0116401, and in part by the National Natural Science Foundation of China under grants 72201007, 72431001, T2121002, 72131001, U2241214, 62373008, and 723B1001.

Y. Liu, Z. Li, and W. Mei are with the College of Engineering, and the State Key Laboratory for Turbulence and Complex Systems, Peking University, Beijing, China. S. Huang and X. Yang are with the Academy of Mathematics and Systems Science, Chinese Academy of Science, Beijing, China.

[†] Yingzhu Liu and Shengyuan Huang contributed equally to the work.

* The corresponding author is Wenjun Mei (mei@pku.edu.cn)

nonlinear causal relations in high-dimensional data [20]. However, a major challenge of the aforementioned methods is the interpretability of the underlying mechanisms that drive the causal relations. Like Granger causality, causal relations tested by those methods are essentially certain forms of statistical relations instead of mechanistic causality.

Behavior-system theory: One promising way to describe unknown systems via data is rooted in Behavioral systems theory, which is a fundamental tool in data-driven control. This theory offers an alternative non-parametric representation of discrete-time LTI systems [21]–[23]. Unlike traditional system identification techniques, it defines a system as a set of trajectories rather than identifying explicit models. A key result in [21] known as the *fundamental lemma*, provides conditions for the existence of the non-parametric representation and establishes a criterion for determining whether a trajectory belongs to the system. Since then, a range of new algorithms has been proposed for system identification, data-driven simulation, and data-driven (predictive) control [23]–[25]. For more details, readers are referred to the review paper [26]. To the best of our knowledge, the behavior-system theory has not been used for causality discovery.

C. Statement of Contributions

In this paper, we conceptualize causality from a control-theoretic perspective and propose a method that unveils mechanistic causality by solving linear equations constructed from time-series data. Specifically, our contributions are summarized as follows.

Firstly, we propose a plain and straightforward conceptualization of causality: a variable is a cause of another variable if the former (partially) drives the dynamic evolution of the latter, i.e., if the former is a control input and the latter is an output of an underlying dynamical system. Such conceptualization sets a high criterion for causality discovery and ensures that the methods aligned with it detect true mechanistic causal relations rather than statistical correlations.

Secondly, we conduct a rigorous analysis for the scenarios when the underlying dynamical systems are linear and time-invariant. We propose a set of tests that, under certain conditions, can distinguish four different types of causal relations between two vector variables based on their offline time-series data. The four possible relations are: 1) independence, i.e., the two variables are collected from different systems and are mutually independent; 2) unilateral full causality, i.e., one variable is the only input of the underlying system, while the other is an output; 3) unilateral partial causality, i.e., one variable is an input (possibly with the presence of other unknown inputs) and the other variable is an output; 4) latent-common-cause relation, i.e., these two variables are both outputs of an underlying system and the input data is not collected. Particularly, for cases 2) and 3), our method is able to detect the causal direction, i.e., which one is the input and which one is the output. We characterize the conditions for the underlying dynamical systems to be causality-discoverable and provide a rigorous proof.

Thirdly, an exploratory case study indicates that the widely studied data-enabled predictive control (DeePC) [25], which is also based on the behavior-system approach, could be leveraged to infer causal relations between variables in nonlinear systems, which indicates potentially broad and promising applicability of the behavior-system theory in causality discovery.

D. Organization

The remainder of the paper is organized as follows. In Section II, we briefly introduce Granger causality and behavioral systems theory. In Section III, we present the proposed concept of *mechanistic causality* and formally define the problem. Section IV contains the main theoretical results and illustrative examples. We conclude the paper and discuss future directions in Section VI.

II. NOTATIONS, DEFINITIONS, AND PRELIMINARIES

A. Basic Notations

Let \mathbb{R} denote the set of real numbers. We use $\mathbb{Z}_{\geq 0}$ and $\mathbb{Z}_{>0}$ to denote the sets of nonnegative and positive integers, respectively. Denote by $\mathbf{1}_m \in \mathbb{R}^m$ the vector of all ones. We define the index set $\bar{T} = \{1, 2, \dots, T\}$. Denote by $\mathcal{D} = \{\mathcal{D}(t)\}_{t \in \bar{T}} \subset \mathbb{R}^q$ an offline data trajectory of length T , where $\mathcal{D}(t)$ is the value of the trajectory at time t . The notation $\text{col}(u, y)$ denotes vertical concatenation, i.e., $\text{col}(u, y) := [u^\top, y^\top]^\top$. For a matrix or vector H , we use $H_{[k, k+L]}$ to denote the submatrix consisting of rows from the k -th to the $(k+L)$ -th (including the $(k+L)$ -th row).

B. Granger Causality

The Granger causality test is one of the most popular tools for inferring causal relation from time series. Essentially, Granger causality is a statistical concept. To put it simply, it interprets “causality” as whether one time series helps improve the prediction of another time series. The Granger causality test relies on a pre-assumption that the time series are weakly stationary. Otherwise, the time series need to be differenced until they are weakly stationary. A \bar{T} -length time series $\{\theta(t)\}_{t \in \bar{T}}$ is weakly stationary if $\mathbb{E}(\theta(t))$, $\text{Var}(\theta(t))$ are constant for all $t \in \bar{T}$, and $\text{Cov}(\theta(t), \theta(t-k))$ depends only on k but not t . A formal definition of Granger causality is given below.

Definition 1. ([6, Granger Causality]): *Given two weakly stationary time series $\theta = \{\theta(t)\}_{t \in \bar{T}} \subset \mathbb{R}^m$ and $\psi = \{\psi(t)\}_{t \in \bar{T}} \subset \mathbb{R}^p$, denote by $\bar{\theta}_t$ and $\bar{\psi}_t$ the sets of all the current and past values of θ and ψ until time t , respectively. We say that u Granger-causes y if and only if*

$$\text{Var}(y(t+1)|\bar{u}_t, \bar{y}_t) < \text{Var}(y(t+1)|\bar{y}_t)$$

for all $t \in \bar{T}$, where the variations are estimated using optimal linear prediction functions.

As implied by the definition above, Granger causality is based on regression analysis. Therefore, it reflects statistical dependencies rather than any driver-response mechanism.

C. Behavioral System Theory

The behavioral-system theory represents a dynamical system as a set of trajectories in a non-parametric manner. In this paper, we follow the definitions and notations [23].

Definition 2 (Behavior systems). A dynamical system is a 3-tuple $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathcal{B})$ where $\mathbb{Z}_{\geq 0}$ is the discrete-time axis, \mathbb{W} is a signal space, and $\mathcal{B} \subseteq \mathbb{W}^{\mathbb{Z}_{\geq 0}}$ is the set of all the possible trajectories of the signal, referred to as the behavior.

- (i). $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathcal{B})$ is linear if \mathbb{W} is a vector space and \mathcal{B} is a linear subspace of $\mathbb{W}^{\mathbb{Z}_{\geq 0}}$.
- (ii). $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathcal{B})$ is time invariant if $\mathcal{B} \subseteq \sigma\mathcal{B}$ where $\sigma: \mathbb{W}^{\mathbb{Z}_{\geq 0}} \rightarrow \mathbb{W}^{\mathbb{Z}_{\geq 0}}$ is the forward time shift defined by $(\sigma w)(t) = w(t+1)$ and $\sigma\mathcal{B} = \{\sigma w \mid w \in \mathcal{B}\}$.

We denote the class of systems $(\mathbb{Z}_{\geq 0}, \mathbb{R}^q, \mathcal{B})$ depicted by Definition 2 as \mathcal{L}^q . Next, we define trajectories truncated to a window of length T by a set

$$\mathcal{B}_T = \{w \in (\mathbb{R}^q)^T \mid \exists v \in \mathcal{B} \text{ s.t. } w(t) = v(t), 1 \leq t \leq T\}.$$

For a linear time-invariant system \mathcal{B} , given any input/output partitioning $w := \text{col}(u, y)$, the input/state/output (i/s/o) description is given by

$$\mathcal{B}(A, B, C, D) := \left\{ \text{col}(u, y) \in (\mathbb{R}^q)^{\mathbb{Z}_{\geq 0}} \mid \exists x \in (\mathbb{R}^n)^{\mathbb{Z}_{\geq 0}} \right. \\ \left. \text{s.t. } \sigma x = Ax + Bu, y = Cx + Du \right\}.$$

A *minimal representation* is referred to as the i/o/s representation with the smallest state dimension, and its order is denoted by $\mathbf{n}(\mathcal{B})$. Define the dimension of the input of \mathcal{B} by $\mathbf{k}(\mathcal{B})$, the dimension of the output by $\mathbf{p}(\mathcal{B})$. We define the observability matrix by $\mathcal{O}_\tau(C, A) := \text{col}(C, CA, \dots, CA^{\tau-1})$ and the Toeplitz matrix

$$\mathcal{T}_\tau(A, B, C, D) := \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ CA^{\tau-2}B & \cdots & CB & D \end{bmatrix}.$$

The lag of a system $\mathcal{B} \in \mathcal{L}^q$, denoted by $\ell(\mathcal{B})$, is defined as the smallest $\ell \in \mathbb{Z}_{>0}$ such that the observability matrix $\mathcal{O}_\ell(A, C)$ has full rank, i.e., $\text{rank}(\mathcal{O}_\ell(A, C)) = \mathbf{n}(\mathcal{B})$. The set of all the systems, with q -dimension signal space, order n , k -dimension input space, and lag l , is denoted by

$$\partial\mathcal{L}_{k,\ell}^{q,n} = \{\mathcal{B} \in \mathcal{L}^q \mid \mathbf{k}(\mathcal{B}) = k, \mathbf{p}(\mathcal{B}) = q - k, \\ \mathbf{n}(\mathcal{B}) = n, \ell(\mathcal{B}) = \ell\}.$$

A depth- L Hankel matrix of a time series $w \in (\mathbb{R}^q)^T$ is

$$\mathcal{H}_L(w) := \begin{pmatrix} w(1) & \cdots & w(T-L+1) \\ \vdots & \ddots & \vdots \\ w(L) & \cdots & w(T) \end{pmatrix}. \quad (1)$$

Below, we cite two useful lemmas needed in our analysis.

Lemma 1. ([27, Corollary 19]): *Let the T -length offline data w_d be generated by $\mathcal{B} \in \partial\mathcal{L}_{k,\ell}^{q,n}$. Then, $\text{image}\mathcal{H}_L(w_d)$ equals \mathcal{B}_L , for $L > \ell$ if and only if $\text{rank}(\mathcal{H}_L(w_d)) = mL + n$.*

Lemma 2. ([24, Lemma 1]): *Let $\mathcal{B} \in \partial\mathcal{L}_{k,\ell}^{q,n}$ and $\mathcal{B}(A, B, C, D)$ be a minimal i/s/o representation. Given an initial system data $\text{col}(u_{\text{ini}}, y_{\text{ini}}) \in \mathcal{B}_{T_{\text{ini}}}$, let $T_{\text{ini}}, T_{\text{f}} \in \mathbb{Z}_{>0}$ with $T_{\text{ini}} \geq \ell$ and $\text{col}(u_{\text{ini}}, u_{\text{f}}, y_{\text{ini}}, y_{\text{f}}) \in \mathcal{B}_{T_{\text{ini}}+T_{\text{f}}}$. Then there exists a unique $x_{\text{ini}} \in \mathbb{R}^n$ such that*

$$y_{\text{f}} = \mathcal{O}_{T_{\text{f}}}(A, C)x_{\text{ini}} + \mathcal{T}_{T_{\text{f}}}(A, B, C, D)u_{\text{f}}. \quad (2)$$

III. CONCEPTUALIZATION OF CAUSALITY AND PROBLEM STATEMENT

A. Conceptualization of Causality

With the behavior-system theory, we are able to conceptualize and test causality in a more straightforward and intrinsic way. Instead of statistical dependencies, we care about the genuine causal relations, interpreted as whether one variable (partially) drives the dynamic evolution of another variable. This aligns with the classical input-output relation in control systems and is formalized as follows.

Definition 3. (Mechanistic Causality): A vector variable θ is *mechanistically causal* to another vector variable ψ if there exist functions f and g , a latent state variable x and possibly a latent input v , such that the joint trajectory (θ, ψ) satisfies the following dynamical system with the input $u(t) = \theta(t)$ and the output $y(t) = \psi(t)$:

$$\begin{cases} x(t+1) = f(x(t); u(t), v(t)), \\ y(t) = g(x(t); u(t), v(t)). \end{cases} \quad (3)$$

The above definition implicitly assumes that the causal effect of the input on the output is instant and deterministic, without any time delay or noise. Causality discovery in delayed or stochastic systems is of great practical importance. New mathematical tools are needed to solve these problems. In this paper, we focus on the prototype problem: deterministic LTI systems with no time delay.

B. Problem Formulation

In this paper, we study how to discover mechanistic causality specified in Definition 3, given the time series of two vector variables generated from an underlying LTI system $\mathcal{B}(A, B, C, D) \in \partial\mathcal{L}_{k,l}^{q,n}$:

$$\begin{cases} x(t+1) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \end{cases} \quad (4)$$

where $u(t) = \text{col}(u_1(t), u_2(t))$, $y(t) = \text{col}(y_1(t), y_2(t))$, and

$$B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}, C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}.$$

All the aforementioned matrices and vectors are of proper dimensions. The matrices A, B, C, D , the input dimension k , the output dimension $q-k$, and the order n are all unknown.

The observation data collected are two T -length time series $\theta_d \in (\mathbb{R}^m)^T$ and $\psi_d \in (\mathbb{R}^p)^T$ generated from the above system, where m and p are the dimensions of the two variables θ and ψ respectively. The mechanistic causal relation between θ and ψ are determined by their roles in the

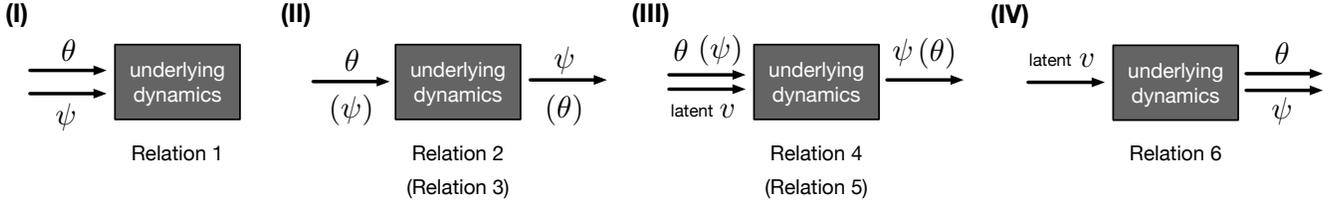


Fig. 1: Four possible causal relations between variables θ and ψ , and the associated six possible relations. **(I)** independence; **(II)** full causality; **(III)** partial causality; **(IV)** latent-common-cause relation.

above underlying system, categorized into the following four basic causal structures with six possible relations, visually illustrated by Fig. 1.

(I) Independence: The variables θ and ψ are mutually independent. This case can be incorporated into our framework by letting

$$\theta(t) = u_1(t), \quad \psi(t) = u_2(t)$$

i.e., they are two free inputs. This is Relation 1.

(II) Full causality: one variable drives the evolution of another. This structure could be considered as the scenarios when θ is the system input and ψ is the system output, or the other way around. That is,

$$\begin{aligned} &\text{either } \theta \rightarrow \psi, \text{ (Relation 2)} \\ &\quad \text{i.e., } \theta(t) = u(t), \psi(t) = y(t), \\ &\text{or } \psi \rightarrow \theta, \text{ (Relation 3)} \\ &\quad \text{i.e., } \psi(t) = u(t), \theta(t) = y(t). \end{aligned}$$

(III) Partial causality: one of the variables is a system input and the other is a system output, with the presence of other latent input not collected in the data. That is, there exists a latent input v such that

$$\begin{aligned} &\text{either } \theta, v \rightarrow \psi, \text{ (Relation 4) i.e.,} \\ &\quad \theta(t) = u_1(t), v(t) = u_2(t), \psi(t) = y(t), \\ &\text{or } \psi, v \rightarrow \theta, \text{ (Relation 5) i.e.,} \\ &\quad \psi(t) = u_1(t), v(t) = u_2(t), \theta(t) = y(t). \end{aligned}$$

(IV) Latent-common-cause relation: A common latent input v drives the evolution of both θ and ψ . That is

$$\begin{aligned} &v \rightarrow \theta, \psi, \text{ (Relation 6)} \\ &\quad \text{i.e., } v(t) = u(t), \theta(t) = y_1(t), \psi(t) = y_2(t). \end{aligned}$$

Our goal is to distinguish the above six possible relations based on the collected data $w_d = \text{col}(\theta_d, \psi_d)$. Intuitively, whether this is feasible depends on how the time series are generated. For an extreme example, if the system input is constantly zero, then it is impossible to infer causal relations correctly from the collected data. In this paper, we impose some requirements for the quality of the time series w_d , referred to as *identifiable time series*. Such requirements are widely adopted in data-driven control [22], [25], [28].

Definition 4. (Identifiable Time Series): For a T -length time series $w_d = \text{col}(\theta_d, \psi_d) \in (\mathbb{R}^{m+p})^T$ collected from an

unknown discrete-time LTI system $\mathcal{B} \in \partial \mathcal{L}_{k,\ell}^{q,n}$ according to the four causal structures in Fig 1, w_d is an identifiable time series if the observed data (θ_d, ψ_d) and the uncollected latent signal v_d satisfy:

- (i) If (I), (III), or (IV) occurs, there exists $T_{\text{ini}} > \ell$ such that $\text{rank}(\mathcal{H}_{T_{\text{ini}}+T_f}(\bar{w}_d)) = k(T_{\text{ini}} + T_f) + n$. Here $\bar{w}_d = \text{col}(\theta_d, \psi_d, v_d) \in \mathcal{B}_T$ and v_d is the time series of the hidden variable v , which is not collected and thus unknown.
- (ii) If (II) occurs, there exists $T_{\text{ini}} > \ell$ such that $\text{rank}(\mathcal{H}_{T_{\text{ini}}+T_f}(w_d)) = k(T_{\text{ini}}+T_f)+n$, where $w_d \in \mathcal{B}_T$.

Remark 1. The rank condition in Definition 4 is satisfied with probability one, when the input signal is sampled i.i.d. from a non-degenerate (i.e., nontrivial) continuous distribution and the time horizon T is sufficiently large.

IV. CAUSAL INFERENCE FOR DETERMINISTIC LTI SYSTEMS

In this section, we propose a behavioral-system approach that transforms causal discovery into solving a fictitious control problem. Suppose we have collected an identifiable T -length offline trajectory $w_d = (\theta_d, \psi_d)$ generated from an underlying system $\mathcal{B} \in \partial \mathcal{L}_{k,\ell}^{q,n}$, i.e., $w_d \in \mathcal{B}_T$. Here the variables θ and ψ obey one the six possible causal relations specified in Section III.B. Suppose we know the underlying system's lag or an upper bound of it. Pick from w_d a piece of T_{ini} -length trajectory $w_{\text{ini}} = \text{col}(\theta_{\text{ini}}, \psi_{\text{ini}}) \in \mathcal{B}_{T_{\text{ini}}}$ with $T_{\text{ini}} > \ell$. According to Lemma 1, any T_f -length ($T_f \in \mathbb{N}_+$) time series (θ_f, ψ_f) is a legal trajectory following w_{ini} if there exists $g \in \mathbb{R}^{T-T_{\text{ini}}-T_f+1}$ such that

$$\begin{bmatrix} \Theta_p \\ \Psi_p \\ \Theta_f \\ \Psi_f \end{bmatrix} g = \begin{bmatrix} \theta_{\text{ini}} \\ \psi_{\text{ini}} \\ \theta_f \\ \psi_f \end{bmatrix}, \quad (5)$$

where

$$\begin{pmatrix} \Theta_p \\ \Theta_f \end{pmatrix} := \mathcal{H}_{T_{\text{ini}}+T_f}(\theta_d), \quad \begin{pmatrix} \Psi_p \\ \Psi_f \end{pmatrix} := \mathcal{H}_{T_{\text{ini}}+T_f}(\psi_d). \quad (6)$$

We know that the inputs of a system, which are free variables, should behave intrinsically differently than the outputs, which are dependent variables. Since equations (5) and (6) correctly represent the underlying system in a finite-time horizon, the input-output asymmetry must be reflected

in these equations and can thus be leveraged to distinguish causal relations. This idea leads to the following Behavior-enabled Causality test (BeCaus test).

Definition 5 (BeCaus Test). Consider a T -length identifiable time series $w_d = \text{col}(\theta_d, \psi_d) \in (\mathbb{R}^{m+p})^T$. Let $\Theta_p, \Theta_f, \Psi_p, \Psi_f$ be given by equation (6) and let $\theta_{\text{ini}}, \psi_{\text{ini}}$ be the first T_{ini} -length series of θ_d, ψ_d , with $T_{\text{ini}} > \ell$. Set $T_f = 2$ and denote by $\theta_f = \text{col}(\theta_f(1), \theta_f(2)) \in (\mathbb{R}^m)^2$ and $\psi_f = \text{col}(\psi_f(1), \psi_f(2)) \in (\mathbb{R}^p)^2$ the two-step future trajectory. We test whether the following conditions hold:

- ① For any θ_f , equation (5) yields a unique solution ψ_f .
- ② For any ψ_f , equation (5) yields a unique solution θ_f .
- ③ There exist $\theta_f(1), \psi_f(1)$ and $\psi_f(2)$ such that equation

$$\begin{bmatrix} \Theta_p \\ \Psi_p \\ \Theta_{f[1,m]} \\ \Psi_{f[1,m]} \end{bmatrix} g = \begin{bmatrix} \theta_{\text{ini}} \\ \psi_{\text{ini}} \\ \theta_f(1) \\ \psi_f(1) \end{bmatrix}, \quad (7)$$

with g as the unknown variable, admits at least one solution, while equation (5), with g and $\theta_f(2)$ as the unknown variables, admits no solution;

- ④ There exist $\theta_f(1), \psi_f(1)$ and $\theta_f(2)$ such that equation (7), with g as the unknown variable, admits at least one solution, while equation (5), with g and $\psi_f(2)$ as the unknown variables, admits no solution.

Based on the results of the above tests, the causal relations between θ and ψ is determined as follows:

TABLE I: Causal relations inferred by the BeCaus test. Here \checkmark means ‘‘True’’, while \times means ‘‘False’’.

①	②	③	④	Structure	Causal relation
\times	\times	\times	\times	(I)	Relation 1 (no relation)
\checkmark	\times	\checkmark	\times	(II)	Relation 2 ($\theta \rightarrow \psi$)
\times	\checkmark	\times	\checkmark	(II)	Relation 3 ($\psi \rightarrow \theta$)
\times	\times	\checkmark	\times	(III)	Relation 4 ($\theta, (v) \rightarrow \psi$)
\times	\times	\times	\checkmark	(III)	Relation 5 ($\psi, (v) \rightarrow \theta$)
\times	\times	\checkmark	\checkmark	(IV)	Relation 6 ($(v) \rightarrow \theta, \psi$)

Now we investigate for what systems the BeCaus test with identifiable time series always leads to correct results. Such systems are referred to as *causality-discoverable systems*. Below, we present sufficient conditions for a system to be causality-discoverable.

Theorem 3. (Causality-Discoverable Systems): Given any identifiable time series $w_d = \text{col}(\theta_d, \psi_d) \in (\mathbb{R}^{m+p})^T$ generated by an underlying LTI system (4) according to one of the six scenarios specified in Section III. B, the BeCaus test in Definition 5 always leads to a correct result if the matrices C and D in (4) satisfies $D_{11} \neq 0$, $D_{22} \neq 0$, and neither $[C_1, D_{11}, D_{12}]$ nor $[C_2, D_{21}, D_{22}]$ has a full row rank;

Proof. We prove this theorem by showing that each of the six possible causal relations must lead to the corresponding test results in Table I.

We start with **Structure (II), Relation 2** ($\theta \rightarrow \psi$): The following fact will be repeatedly used: With any identifiable time series w_d , according to Lemma 1, the set of (θ_f, ψ_f) satisfying equation (5) is equal to the set of (θ_f, ψ_f) satisfying

$$\begin{aligned} \psi_f(1) &= Cx_{\text{ini}} + D\theta_f(1) \\ \psi_f(2) &= CAx_{\text{ini}} + CB\theta_f(1) + D\theta_f(2), \end{aligned} \quad (8)$$

where x_{ini} , according to Lemma 2, is uniquely determined by w_{ini} , since $T_{\text{ini}} > \ell$. Now we go through each test.

Test ①: In equation (8), any $\theta_f \in (\mathbb{R}^m)^2$ uniquely determines a $\psi_f \in (\mathbb{R}^p)^2$. Therefore, equation (5) yields a unique solution ψ_f for any given θ_f . That is, the result of Test ① must be ‘‘ \checkmark ’’.

Test ②: Since neither $[C_1, D_{11}, D_{12}]$ nor $[C_2, D_{21}, D_{22}]$ has a full row rank, the matrix D does not have a full row rank and thus the image space $\mathcal{R}(D) \subsetneq \mathbb{R}^p$. Given any x_{ini} uniquely determined by w_{ini} , let $\psi_f(1) = Cx_{\text{ini}} + w$, where $w \in \mathbb{R}^p \setminus \{\mathcal{R}(D)\}$. Then there does not exist any ψ_f satisfying equation (8), and thereby no ψ_f satisfies equation (5). Therefore, Test ② must return the result ‘‘ \times ’’.

Test ③: For any given $\theta_f(1) \in \mathbb{R}^m$, $\psi_f(1) = Cx_{\text{ini}} + D\theta_f(1)$, satisfies equation (8) where x_{ini} is uniquely determined by w_{ini} . Therefore, $(\theta_f(1), \psi_f(1))$ also satisfies equation (7), which is a part of equation (5). Let

$$\psi_f(2) = CAx_{\text{ini}} + CB\theta_f(1) + w, \text{ with } w \in \mathbb{R}^p \setminus \mathcal{D}.$$

Since $\mathcal{R}(D) \subsetneq \mathbb{R}^p$, there does not exist any $\theta_f(2)$ satisfying equation (8). Therefore, no $\theta_f(2)$ satisfies equation (5) either. Namely, Test ③ must return the result ‘‘ \checkmark ’’.

Test ④: Since equation (7) is a part of equation (5) with the equations for $(\theta_f(2), \psi_f(2))$ removed, any $\theta_f(1)$ and $\psi_f(1)$ satisfying equation (7) also satisfies equation (8). Moreover any $\theta_f(2) \in \mathbb{R}^m$ uniquely determines a $\psi_f(2) \in \mathbb{R}^p$ in equation (8). Therefore, given any $\theta_f(1)$ and $\psi_f(1)$ satisfying equation (7), and any $\theta_f(2) \in \mathbb{R}^m$, there exists a unique $\psi_f(2)$ satisfying equation (5). That is, Test ④ returns the result ‘‘ \times ’’.

Structure (II), Relation 3 ($\psi \rightarrow \theta$): The proof is the same as that for Relation 2, except that θ and ψ are interchanged.

Structures (I), (III) and (IV): One feature these structures have in common is that there exists a hidden input or output variable v , which generates a time series v_d such that $\bar{w}_d = \text{col}(\theta_d, \psi_d, v_d)$ satisfies the identifiable conditions, but v_d is not included in the collected data $w_d = (\theta_d, \psi_d)$. Let v_{ini} be the first T_{ini} -length series of v_d . Any future 2-length trajectory (θ_f, ψ_f, v_f) following the time series \bar{w}_{ini} satisfies not only equation (5) but also

$$\begin{bmatrix} V_p \\ V_f \end{bmatrix} g = \begin{bmatrix} v_{\text{ini}} \\ v_f \end{bmatrix} \quad (9)$$

In equations (5) and (9), θ_f, ψ_f, v_f , and g are the unknown variables. Since $T_{\text{ini}} > \ell$, $\theta_{\text{ini}}, \psi_{\text{ini}}$ and v_{ini} together uniquely determine x_{ini} . However, since the data of v_{ini} is not collected and thus unknown, there could be multiple x_{ini} ’s compatible with $(\theta_{\text{ini}}, \psi_{\text{ini}})$. For simplicity of notations, let

$$D_{*1} = \text{col}(D_{11}, D_{21}), \quad D_{*2} = \text{col}(D_{12}, D_{22}),$$

$$D_{1*} = [D_{11}, D_{12}], \quad D_{2*} = [D_{21}, D_{22}].$$

Structure (III), Relation 4 ($\theta, (v) \rightarrow \psi$): According to Lemma 1, the set of (θ_f, ψ_f) satisfying equation (5) is equal to that satisfying

$$\begin{aligned} \psi_f(1) &= Cx_{ini} + D_{*1}\theta_f(1) + D_{*2}v_f(1), \\ \psi_f(2) &= CAx_{ini} + CB_1\theta_f(1) + CB_2v_f(1) \\ &\quad + D_{*1}\theta_f(2) + D_{*2}\psi_f(2), \end{aligned} \quad (10)$$

where v_f is a free variable and x_{ini} , by Lemma 2, is compatible with but not uniquely determined by $(\theta_{ini}, \psi_{ini})$, since v_{ini} is unknown. Now we analyze each test.

Test ①: Since $D_{22} \neq 0$ and $v_f(1)$ is a free variable, any given $\theta_f(1)$ alone does not uniquely determine $\psi_f(1)$ in equation (10) and thereby does not in equation (5) either. Therefore, Test ① returns the result “ \mathbf{X} ”.

Test ②: Since neither $[C_1, D_{1*}]$ nor $[C_2, D_{2*}]$ has a full row-rank, neither does $[C, D_{*1}, D_{*2}]$. As a result, $\mathcal{R}([C, D_{*1}, D_{*2}]) \subsetneq \mathbb{R}^p$. For any $\psi_f(1) \in \mathbb{R}^p \setminus \mathcal{R}([C, D_{*1}, D_{*2}])$, there does not exist any $\theta_f(1)$ and $\psi_f(1)$ satisfying equation (10), and thereby no $\theta_f(1)$ satisfies equation (5). That is, Test ② must return the result “ \mathbf{X} ”.

Test ③: For any $\theta_f(1)$ and $v_f(1)$, let $\psi_f(1)$ be given by equation (10). Then $(\theta_f(1), \psi_f(1))$ satisfy equation 7, which is a part of equation (5). Moreover, since neither $[C_1, D_{11}, D_{12}]$ nor $[C_2, D_{21}, D_{22}]$ has a full row-rank, $[C, D_{*1}, D_{*2}]$ does not have a full row-rank either. As a result, $[CA, CB_1, CB_2, D_{*1}, D_{*2}]$ does not have a full row-rank and thus $\mathcal{R}([CA, CB_1, CB_2, CD_{*1}, CD_{*2}]) \subsetneq \mathbb{R}^p$. Pick a $\psi_f(2) \in \mathbb{R}^p \setminus \mathcal{R}([CA, CB_1, CB_2, CD_{*1}, CD_{*2}])$. Then no $\psi_f(2)$ satisfies equation (10) and thus no $\psi_f(2)$ satisfies equation (5). Therefore, Test ③ must return “ \checkmark ”.

Test ④: We have shown that any pair $(\theta_f(1), \psi_f(1))$ given by equation (10) also satisfy equation (5) and thus equation (7). Moreover, for any such $(\theta_f(1), \psi_f(1))$, any $\theta_f(2)$, and any $v_f(2)$, $\psi_f(2)$ given by equation (5) always satisfies equation (5) and thus equation (7). That is, Test ④ must return the result “ \mathbf{X} ”.

Structure (III), Relation 5 ($\psi, (v) \rightarrow \theta$): The proof is the same as that for Relation 4 except that the symbols θ and ψ are interchanged.

Structure (IV), Relation 6 ($(v) \rightarrow \theta, \psi$): According to Lemma 1, the set of (θ_f, ψ_f) satisfying equation (5) is equal to that satisfying

$$\begin{aligned} \theta_f(1) &= C_1x_{ini} + D_{1*}v_f(1), \\ \psi_f(1) &= C_2x_{ini} + D_{2*}v_f(1), \\ \theta_f(2) &= C_1Ax_{ini} + CBv_f(1) + D_{1*}v_f(2), \\ \psi_f(2) &= C_2Ax_{ini} + CBv_f(1) + D_{2*}v_f(2), \end{aligned} \quad (11)$$

where v_f is a free variable and x_{ini} , by Lemma 2, is compatible with but not uniquely determined by $(\theta_{ini}, \psi_{ini})$ since v_{ini} is unknown.

Test ①: Since $[C_1, D_{1*}]$ does not have a full row rank, we have $\mathcal{R}([C_1, D_{1*}]) \subsetneq \mathbb{R}^m$. As a result, for any given $\theta_f(1) \in \mathbb{R}^m \setminus \mathcal{R}([C_1, D_{1*}])$, there does not exist any $\theta_f(1)$ satisfying equation (11) and thereby $\psi_f(1)$ does not exist

either. Such a pair $(\theta_f(1), \psi_f(1))$ will not satisfy equation (5). Therefore, Test ① will return the result “ \mathbf{X} ”.

Test ②: Due to the symmetry between θ and ψ in Structure IV, the proof for Test ② is the same as the proof for Test ①, except that the symbols θ and ψ are interchanged.

Test ③: For any $v_f(1)$, by Lemma 1, $\theta_f(1)$ and $\psi_f(1)$ given by equation (11) satisfy equation (5). Moreover, since $[C_1, D_{1*}]$ does not have a full row rank, $[C_1A, CB, D_{1*}]$ does not have a full row rank either, and thereby $\mathcal{R}([C_1A, CB, D_{1*}]) \subsetneq \mathbb{R}^m$. As a result, for any given $\theta_f(2) \in \mathbb{R}^m \setminus \mathcal{R}([C_1A, CB, D_{1*}])$, no v_f satisfies equation (11). Therefore, such $(\theta_f(1), \psi_f(1), \theta_f(2))$ is not a feasible trajectory and does not satisfy equation (5). That is, Test ③ must return the result “ \mathbf{X} ”.

Test ④: The proof is the same as the proof for Test ③, except that the symbols θ and ψ are interchanged.

Structure (I), Relation 1 ($\theta, \psi \rightarrow (v)$): In this structure, since θ_f and ψ_f are free variables, Test ①-④ will all return the result “ \mathbf{X} ”. This concludes the proof. ■

Remark 2. One might find a “loophole” in the BeCaus test: Suppose the underlying dynamics is Structure I. Since, in this case, both θ and ψ are free variables, what if the underlying system accidentally generates an offline trajectory (θ_d, ψ_d) that coincides with some trajectory generated by Structure II, III, or IV? In fact, the identifiable-data condition ensures that, if the offline data (θ_d, ψ_d) coincides with the data generated by Structure II, III or IV and satisfies their corresponding identifiability condition, then it must not satisfy the identifiability condition for Structure I and is thus precluded. To put it in another way, if $\theta(t)$ and $\psi(t)$ are randomly generated according to some i.i.d distributions, then almost surely the collected data (θ_d, ψ_d) almost surely will not coincide with the data generated by Structure II, III or IV.

Remark 3. The conditions in Theorem 3 have clear control-theoretic interpretations. The condition that $[C_1, D_{11}, D_{12}]$ and $[C_2, D_{21}, D_{22}]$ are not full row rank implies that each output is not controllable in one period, while $D_{11}, D_{22} \neq 0$ indicates that the inputs affect outputs without delay. In face, the proof of Theorem 3 reveals that the required conditions for each model can be slightly weaker than the unified condition in Theorem 3. In particular, the conclusions Table I hold if the following conditions are satisfied:

- 1) If Structure (II) occurs, D is not full row rank;
- 2) If Structure (III) occurs, $[C, D_u, D_v]$ is not full row rank, $D_v \neq 0$;
- 3) If Structure (IV) occurs, $[C_1, D_\theta]$ and $[C_2, D_\psi]$ are not full row rank, $D_\theta, D_\psi \neq 0$. Matrices D_u, D_v, D_θ , and D_ψ are defined in the proof.

V. COMPARISONS WITH GRANGER CAUSALITY AND FURTHER DISCUSSIONS

In this section, we compare the proposed BeCaus test with the Granger causality test and then discussion a tentative extension of our approach to nonlinear systems.

A. Comparing with Granger causality test via examples

In this subsection, we compare the Granger causality test with the proposed BeCaus tests via some simple examples of LTI systems. In the following examples, the state variable of the underlying systems is set to be two-dimensional, with the initial state $x(0) = [1, 0]^\top$. The length of the collected offline time series is $T = 50$, all satisfying the identifiable-data conditions. It turns out that the Granger test returns incorrect results in all the examples except Example 1, while the results of the BeCaus test are correct in all the examples.

Example 1 (Structure (I)). *In this example, θ and ψ are both scalar free variables. The entries in the T -length sequence θ_d are independently randomly generated from the uniform distribution on $(-1, 1)$, while the entries of ψ_d are independently randomly generated from the uniform distribution on $(-10, 10)$. We verify that both θ_d and ψ_d are stationary by using the Augmented Dickey-Fuller (ADF) test. Both the Granger and BeCaus tests indicate that there is no causal relation between the two variables; see Fig. 2(A).*

Example 2 (Structure (II), $\theta \rightarrow \psi$). *In this example, we construct the following LTI system:*

$$\begin{aligned} x(t+1) &= \begin{bmatrix} 1 & -0.5 \\ 0.5 & 1 \end{bmatrix} x(t) + \begin{bmatrix} -0.5 \\ 2 \end{bmatrix} \theta(t), \\ \psi(t) &= \begin{bmatrix} 2 & -2 \end{bmatrix} x(t). \end{aligned}$$

The offline data θ_d is independently randomly generated from a uniform distribution over $[0, 1]$, and the corresponding output ψ_d is obtained. Both θ_d and ψ_d are stationary time series, so the direct Granger test is applicable. It turns out that the BeCaus test returns the correct result, while the Granger causality test fails, see Fig. 2(B).

Example 3 (Structure (III), $\theta, (v) \rightarrow \psi$). *In this example, the underlying system is*

$$\begin{aligned} x(t+1) &= \begin{bmatrix} 1.5 & -0.5 \\ 0.5 & 0.8 \end{bmatrix} x(t) + \begin{bmatrix} -0.5 \\ 2 \end{bmatrix} \theta(t) + \begin{bmatrix} 1.5 \\ -2 \end{bmatrix} v(t), \\ \begin{bmatrix} \psi_1(t) \\ \psi_2(t) \end{bmatrix} &= \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} x(t) + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \theta(t) + \begin{bmatrix} 2 \\ 1 \end{bmatrix} v(t). \end{aligned}$$

Here the input series θ_d and v_d are independently randomly generated from uniform distributions on $U[-1, 1]$ and $U[-10, 10]$ respectively. The corresponding output series ψ_d is collected. The Granger test yields an incorrect result that there is no causal relation between θ and either ψ_1 or ψ_2 . The result of the BeCaus test is correct, see Fig. 2(C).

Example 4 (Structure (IV), $(v) \rightarrow \theta, \psi$). *In this example, the underlying system is*

$$x(t+1) = \begin{bmatrix} 0.5 & -0.5 \\ 0.5 & 0.5 \end{bmatrix} x(t) + \begin{bmatrix} -0.5 & 3 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix},$$

$$\begin{bmatrix} \theta_1(t) \\ \theta_2(t) \\ \psi_1(t) \\ \psi_2(t) \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 0 & 0 \\ 2 & 0.5 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}.$$

In this system, $\theta_2(t)$ and $\psi_2(t)$ are constantly zero. Only the time series of θ_1 and ψ_1 are collected, denoted by θ_d and ψ_d respectively. The latent inputs v_1 and v_2 are independently randomly generated from the uniform distribution on $[0, 1]$. The Granger test indicates that θ_1 and ψ_1 causes each other, which is incorrect. The BeCaus test returns the correct result, see Fig.2(D).

B. An exploratory case study for nonlinear systems

In this subsection, we briefly explore the potential of applying the behavioral-system theory to causality discovery for nonlinear systems. Consider the following underlying system:

$$x(t+1) = \tanh(Ax(t) + B\theta(t)), \quad \psi(t) = Cx(t),$$

where A , B , and C are taken from the simulation setup in Section V of [29]. Given the offline data (θ_d, ψ_d) , we distinguish the causal relations between θ and ψ by solving the following two fictitious optimal control problems:

$$\min_{\theta_f, \psi_f, g} \|\psi_f(2) - r\mathbf{1}_m\|^2 + \|\theta_f\|^2 + \|g\|_1, \quad (12)$$

$$\min_{\theta_f, \psi_f, g} \|\theta_f(2) - r\mathbf{1}_p\|^2 + \|\psi_f\|^2 + \|g\|_1, \quad (13)$$

with the constraints given by equation (5). Here r is a reference signal set to be a large value, say 1000. Intuitively, driving the input signal to match a target value after two steps may cause large variations between consecutive steps, while the output signal typically exhibits smoother transitions due to its dependence on previous outputs. Consistent with such intuition, solving (12) yields $\|\psi^*(2)\|/\|\psi^*(1)\| = 2.432$, while solving (13) gives $\|\theta^*(2)\|/\|\theta^*(1)\| = 61.425$. The significant difference between these two indices suggests that θ is more likely to be the input variable. This exploratory case study indicates that the intrinsic input-output asymmetry is still preserved in equation (5), which is no longer a precise representation of the underlying nonlinear dynamics though. Such asymmetry could still be leveraged for causality discovery.

VI. CONCLUSION AND DISCUSSION

This paper investigates causal discovery from time series using a behavioral systems perspective. We introduce the concept of *mechanistic causality*, which captures the underlying structural dependencies between variables, rather than relying solely on predictive statistics. Focusing on deterministic linear time-invariant (LTI) systems, we analyze four representative causal structures and propose the BeCaus test that distinguish different causal directions without explicit model identification. It differs from the classic Granger causality test in the sense of how causality are conceptualized. Moreover, in deterministic LTI systems with identifiable data and without time delay, the BeCaus test outperforms

(A) Example 1

True causal relation		$\theta, \psi \rightarrow (v)$ (Relation 1: independence)					
Test	Granger test		BeCaus test				
	$\theta \rightarrow \psi$	$\psi \rightarrow \theta$	①	②	③	④	
	p value = 0.27	p value = 0.24	x	x	x	x	
Result	No causality between θ and ψ		$\theta, \psi \rightarrow (v)$ (Relation 1)				
Is the result correct?	YES		YES				

(B) Example 2

True causal relation		$\theta \rightarrow \psi$ (Relation 2: full causality)					
Test	Granger test		BeCaus test				
	$\theta \rightarrow \psi$	$\psi \rightarrow \theta$	①	②	③	④	
	p value = 0.25	p value = 0.51	√	x	√	x	
Result	No causality between θ and ψ		$\theta \rightarrow \psi$ (Relation 2)				
Is the result correct?	NO		YES				

(C) Example 3

True causal relation		$\theta, (v) \rightarrow \psi$ (Relation 4: partial causality)							
Test	Granger test				BeCaus test				
	$\theta \rightarrow \psi(1)$	$\psi(1) \rightarrow \theta$	$\theta \rightarrow \psi(2)$	$\psi(2) \rightarrow \theta$	①	②	③	④	
	p = 0.40	p = 0.73	p = 0.40	p = 0.73	x	x	√	x	
Result	No causality between θ and ψ				$\theta, (v) \rightarrow \psi$ (Relation 4)				
Is the result correct?	NO				YES				

(D) Example 4

True causal relation		$(v) \rightarrow \theta, \psi$ (Relation 6: latent-common-cause relation)					
Test	Granger test		BeCaus test				
	$\theta_1 \rightarrow \psi_1$	$\psi_1 \rightarrow \theta_1$	①	②	③	④	
	p value = 0.00	p value = 0.00	x	x	√	√	
Result	θ and ψ Granger cause each other.		$(v) \rightarrow \theta, \psi$ (Relation 6)				
Is the result correct?	NO		YES				

Fig. 2: Detailed test results for Examples 1 4.

the Granger causality test, since the former is always correct while the latter often yield spurious results.

To the best of our knowledge, this work is the first to bridge causal discovery with behavioral systems theory. This interaction offers a novel perspective on causality and holds strong potential to enrich existing theoretical frameworks. Several promising extensions lie beyond the scope of deterministic LTI systems, including noisy, nonlinear, and time-delay open systems.

REFERENCES

- [1] T. O. Bergmann and G. Hartwigsen, "Inferring causality from noninvasive brain stimulation in cognitive neuroscience," *Journal of Cognitive Neuroscience*, vol. 33, no. 2, pp. 195–225, 2021.
- [2] L. A. T. Cox and D. A. Popken, "Has reducing fine particulate matter and ozone caused reduced mortality rates in the united states?" *Annals of Epidemiology*, vol. 25, no. 3, pp. 162–173, 2015.
- [3] J. Runge, S. Bathiany, E. Bollt, G. Camps-Valls, D. Coumou, E. Deyle, C. Glymour, M. Kretschmer, M. D. Mahecha, J. Muñoz-Marí, E. H. van Nes, J. Peters, R. Quax, M. Reichstein, M. Scheffer, B. Schölkopf, P. Spirtes, G. Sugihara, J. Sun, K. Zhang, and J. Zscheischler, "Inferring causation from time series in earth system sciences," *Nature Communications*, vol. 10, no. 1, p. 2553, 2019.
- [4] C. A. Sims, "Money, income, and causality," *The American Economic Review*, vol. 62, no. 4, pp. 540–552, 1972.
- [5] G. W. Imbens, "Causal inference in the social sciences," *Annual Review of Statistics and Its Application*, vol. 11, no. 1, pp. 123–152, 2024.
- [6] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, pp. 424–438, 1969.
- [7] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*. New York : Springer, 2005.
- [8] T. J. Sargent, "A classical macroeconomic model for the united states," *Journal of Political Economy*, vol. 84, no. 2, pp. 207–237, 1976.
- [9] S. Basu, A. Shojaie, and G. Michailidis, "Network Granger causality with inherent grouping structure," *Journal of Machine Learning Research*, vol. 16, no. 13, pp. 417–453, 2015.
- [10] M. Eichler, "Graphical modelling of multivariate time series," *Probability Theory and Related Fields*, vol. 153, no. 1-2, pp. 233–268, 2012.
- [11] M. Dimovska and D. Materassi, "A control theoretic look at Granger causality: Extending topology reconstruction to networks with direct feedthroughs," *IEEE Transactions on Automatic Control*, vol. 66, no. 2, pp. 699–713, 2021.
- [12] C. Granger, "Some recent development in a concept of causality," *Journal of Econometrics*, vol. 39, no. 1-2, pp. 199–211, 1988.
- [13] L. Barnett, A. B. Barrett, and A. K. Seth, "Granger causality and transfer entropy are equivalent for gaussian variables," *Physical Review Letters*, vol. 103, no. 23, p. 238701, 2009.
- [14] R. Vicente, M. Wibral, M. Lindner, and G. Pipa, "Transfer entropy – a model-free measure of effective connectivity for the neurosciences," *Journal of Computational Neuroscience*, vol. 30, no. 1, pp. 45–67, 2011.
- [15] P.-O. Amblard and O. J. J. Michel, "On directed information theory and Granger causality graphs," *Journal of Computational Neuroscience*, vol. 30, no. 1, pp. 7–16, 2011.
- [16] K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka, and J. Bhattacharya, "Causality detection based on information-theoretic approaches in time series analysis," *Physics Reports*, vol. 441, no. 1, pp. 1–46, 2007.
- [17] J. Peters, D. Janzing, and B. Schölkopf, *Elements of Causal Inference: Foundations and Learning Algorithms*. The MIT Press, 2017.
- [18] R. Guo, L. Cheng, J. Li, P. R. Hahn, and H. Liu, "A survey of learning causality with data," *ACM Computing Surveys*, vol. 53, no. 4, pp. 1–37, 2021.
- [19] V. Solo, "On causality and mutual information," in *47th IEEE Conference on Decision and Control*, 2008, pp. 4939–4944.
- [20] A. Tank, I. Covert, N. Foti, A. Shojaie, and E. B. Fox, "Neural granger causality," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 8, pp. 1–1, 2021.
- [21] J. C. Willems, "From time series to linear system—part i. finite dimensional linear time invariant systems," *Automatica*, vol. 22, no. 5, pp. 561–580, 1986.
- [22] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.
- [23] I. Markovsky, *Exact and Approximate Modeling of Linear Systems: A Behavioral Approach*. Society for Industrial and Applied Mathematics, 2006, no. 11.
- [24] I. Markovsky and P. Rapisarda, "Data-driven simulation and control," *International Journal of Control*, vol. 81, no. 12, pp. 1946–1959, 2008.
- [25] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the deepc," in *18th European Control Conference (ECC)*, 2019, pp. 307–312.
- [26] Ivan Markovsky and F. Dörfler, "Behavioral systems theory in data-driven analysis, signal processing, and control," *Annual Reviews in Control*, vol. 52, pp. 42–64, 2021.
- [27] I. Markovsky and F. Dörfler, "Identifiability in the behavioral setting," *IEEE Transactions on Automatic Control*, vol. 68, no. 3, pp. 1667 – 1677, 2022.
- [28] I. Markovsky and F. Dorfler, "Identifiability in the behavioral setting," *IEEE Transactions on Automatic Control*, vol. 68, no. 3, pp. 1667–1677, 2023.
- [29] F. Fiedler and S. Lucia, "On the relationship between data-enabled predictive control and subspace predictive control," in *2021 European Control Conference (ECC)*, Jun. 2021, pp. 222–229.