

Causal Discovery in Symmetric Dynamic Systems with Convergent Cross Mapping

Yiting Duan^{a,*}, Jack Yang^b, Ming Yin^c, Yi Guo^a

^a*School of Computer, Data and Mathematical Sciences, Western Sydney University, Parramatta, NSW, 2150, Australia*

^b*School of Computing and Information Technology, University of Wollongong, Wollongong, 2500, Australia*

^c*School of Semiconductor Science and Technology, South China Normal University, Guangzhou, 510631, China*

Abstract

This paper systematically discusses how the inherent properties of chaotic attractors influence the results of discovering causal relationships from time series using convergent cross mapping, particularly how convergent cross mapping misleads the bidirectional causal relationship as unidirectional when the chaotic attractor exhibits symmetry. We propose a novel method based on the k-means clustering method to address the challenges when the measurement function defines the even parity, which may cause this issue. This method is demonstrated to recover the symmetry of the latent chaotic attractor and discover the correct causal links between time series without introducing information from other variables. We validate the accuracy of this method using time series derived from numerical simulations and a real-world system for which convergent cross mapping may conclude erroneous results.

Keywords: Convergence Cross Mapping; Symmetric Chaos; Attractor Manifold; Causal Discovery; Chaotic Dynamic System

Discovering nonlinear causal interactions in complex chaotic systems is essential for uncovering the underlying physical phenomena. Convergent cross mapping is the crucial tool for this causal discovery which is based on embedding the time series into a shadow manifold that is topologically equivariant to the original attractor through delay embedding. However, despite the problem of selecting embedding parameters, the accuracy of the obtained causal relationships is also highly related to the intrinsic properties of the original attractor, especially when the chaotic attractor exhibits rotation symmetry, such as Lorenz63 and Burke & Shaw. Under this circumstance, the shadow manifold reconstructed from the variable that defines even parity may lose the original symmetry, causing the convergent cross mapping to mislead bidirectional causal links as unidirectional. In this work, we illustrate this as the reconstructing mapping mods out the symmetry, thereby no longer being

*Corresponding author

Email addresses: Y.Duan@westernsydney.edu.au (Yiting Duan), jiey@uow.edu.au (Jack Yang), m.yin@scnu.edu.cn (Ming Yin), y.guo@westernsydney.edu.au (Yi Guo)

a one-to-one mapping. We propose a new framework based on k-means clustering to solve this problem without introducing any information from another variable. The efficiency of this method is demonstrated using various time series derived from numerical simulations and a real-world system.

1. Introduction

Identifying the causation between signals from complex systems has garnered significant interest and is essential in ecology, epidemiology, and climatology, where causal relations are inferred from observational signals. Granger causality, as outlined by [1], is a prevalent method for detecting causal links in signals based on the principle of predictability and has been widely used to analyze economic growth and neuroscience for effective brain connectivity [2, 3]. However, the effectiveness of Granger causality is contingent on the separability of the system, making it inadequate for signals from nonlinear or deterministic physical systems, particularly systems with attractors characterized by non-separability and moderate coupling.

Addressing the limitations of Granger causality, Sugihara [4] proposed the Convergent Cross Mapping (CCM) method. This method leverages the principles of chaotic dynamics to detect the causal relationship between signals from nonlinear, nonseparable, and moderately coupled dynamic systems. The key idea of CCM for detecting the causal relationship between two signals is determining whether the historical states of one signal can reliably predict the states of another by checking the existence of the mapping between two attractors reconstructed from two signals. Recently, CCM has been successfully applied in several fields, including climatology [5], neuroscience [6], and ecology [7].

The application of CCM relies heavily on the embedding technique, mainly the delay-coordinate mapping, which serves as a diffeomorphism between the reconstructed manifold, also termed the shadow manifold $\mathcal{M}$, and the original attractor manifold. Ideally, this diffeomorphism ensures that the dynamics of the original system are faithfully captured in the shadow manifold. However, the delay-coordinate mapping can not reach a diffeomorphism in practical scenarios. This discrepancy can arise due to various external factors, including noise, missing values, and improper selection of reconstruction parameters, e.g., the inappropriate lags for the delay-coordinate mapping. Furthermore, we point out that even in cases where the signals are clean and sufficiently lengthy, intrinsic properties of the chaotic system, such as the observability of the variables and inherent symmetries, may also influence the embedding quality. Despite the significance of these factors, few studies have thoroughly examined the impact of such inherent factors on CCM results.

Moreover, special attention should be given to cases when signals from the attractor exhibit rotation symmetry since the shadow manifold reconstructed from the signals of the symmetric invariant variable of this attractor may lose the symmetry such that the delay-coordinate mapping is no longer one-to-one.

For example, the shadow manifold reconstructed from the z -coordinate projection of the Lorenz63 system does not exhibit the same butterfly attractor as the original ones, as noted in counterexamples provided in [8, 9] and the supplementary material of [4]. Although such time series arising from symmetric attractors are abundant in real scenarios, such as output signals from chaotic circuits, the previous works have yet offer a practical solution to address this issue.

In this work, we make the following contributions:

1. We systematically discuss how the inherent properties of chaotic systems, particularly those with symmetric characteristics, influence the results of CCM and derive a general conclusion.
2. We propose a new framework based on the k-means clustering method to address the challenges when the measurement function defines the even parity, which may cause the misleading result caused by CCM.
3. We validate the correctness and efficiency of our method via several examples.

The rest of this article is organized as follows. Section 2 introduces CCM and discusses related work in this field. Section 3 describes the challenge of CCM for several counterexamples. In section 4, we explain the reason by systematically discussing the influence of symmetry and propose our framework. Section 5 assesses the proposed method in several cases. Finally, conclusions are discussed in Section 6.

2. Background

In this section, we provide a self-contained introduction to the CCM methodology, specifically focusing on its ability to identify causal relationships between two variables X and Y , generated from two deterministic dynamic systems.

2.1. Attractor Reconstruction

In practical applications, signals from dynamic systems with attractors are often shown as time series of real numbers generated from observations. The goal is to reconstruct the state space of the unknown attractor for analysis. We introduce both discrete-time and continuous-time dynamic systems.

A *discrete-time dynamic system* is represented in the form $\mathbf{x}_{t+1} = \phi(\mathbf{x}_t)$, where the state $\mathbf{x}_t \in M$ is defined at some time $t \in \mathbb{Z}$ on a compact n -dimensional manifold M where $\phi : M \rightarrow M$ is a diffeomorphism, also called the iteration map, describing the evolution of the system with time. A *continuous-time dynamic system* is a set of first-order ordinary differential equations which is represented in the form $\dot{\mathbf{x}} = v(\mathbf{x})$, where the state $\mathbf{x} \in M$ is defined at all time $t \in \mathbb{R}$, M is a smooth compact n -dimensional manifold and v is smooth ($\mathcal{C}^2$) vector field on M . The unique solution $\psi(\mathbf{x}, t) : M \times \mathbb{R} \rightarrow M$ is the *flow* generated by the vector field v , denoted as $\psi_t(\mathbf{x}) = \psi(\mathbf{x}, t)$ satisfying $\psi_{t_1} \circ \psi_{t_2} = \psi_{t_1+t_2}$ for

$t \in \mathbb{R}$. Give a certain initial condition $\mathbf{x}_0 \in M$, the solution to the dynamic system is the *trajectory*, which is denoted as $\{\psi(\mathbf{x}_t)\}$, for $t = 0, 1, 2, \dots$ or a curve $\{\psi_t(\mathbf{x}_0)\}_{t \geq 0}$. Moreover, if the vector field v or the iteration map ϕ is time-independent, the dynamic system is an *autonomous dynamic system*. In this paper, our analysis is confined to autonomous systems. This focus is justified because a non-autonomous dynamical system can be transformed into an autonomous system by introducing the time variable as an additional dimension in the phase space. For example, the non-autonomous system $\dot{\mathbf{x}} = v(\mathbf{x}, t)$ is equivalent to the autonomous system in one higher dimension $\dot{\mathbf{x}} = v(\mathbf{x}, s)$, with $\dot{s} = 1$ where $s = t$.

It is important to note that although the system is inherently continuous, practical constraints necessitate that measurements are taken at a regular sampling interval T . Given this sampling rate, the discrete flow denoted as $\xi_T : M \rightarrow M$, can be defined and characterized by the equations $\xi_T(\mathbf{x}_t) = \mathbf{x}_{t+T}$ and $\xi_T^{-1}(\mathbf{x}_t) = \mathbf{x}_{t-T}$ which associate each point state $\mathbf{x} \in M$ the vector $\xi_T(\mathbf{x})$. The sampling rate T represents an integer multiple of the iteration step for discrete dynamic systems.

In the theoretical framework of CCM, signals are causally linked if they share the common attractor manifold $\mathbb{A}$. The *attractor* $\mathbb{A}$ is defined as a closed subspace on a smooth compact manifold M with $\dim(\mathbb{A}) \leq n$ which satisfies the following three axioms [10]:

1. $\mathbb{A}$ attracts an open set of initial conditions: There is an open set U containing $\mathbb{A}$ such that if the initial state $\mathbf{x}_0 \in U$, then

$$\lim_{t \rightarrow \infty} \text{dist}(\mathbf{x}_t, \mathbb{A}) = 0$$

2. Invariant set: If $\mathbf{x}_{t_0} \in \mathbb{A}$, then $\mathbf{x}_t \in \mathbb{A}$ for all $t \geq t_0$
3. Minimal: No proper subset of $\mathbb{A}$ also satisfies these conditions.

The first two axioms ensure that points within the "basin of attraction" are drawn towards $\mathbb{A}$, while the third establishes that every part of $\mathbb{A}$ is crucial. The existence of attractors has been studied for several decades, and it is established that every smooth, compact dynamical system possesses at least one attractor [11]. Without loss of generality, we assume that the initial state $\mathbf{x}_0$ is contained within the attractor, ensuring that the trajectory remains within the attractor $\mathbb{A}$ during the periods of interest, such that $\mathbf{x}_t \in \mathbb{A} \subset \mathbb{R}^N$ for $t \geq 0$.

Direct observation of the full state of the dynamic system is often infeasible. Instead, we have access to observations via a real-valued measurement function $h : \mathbb{A} \rightarrow \mathbb{R}$, producing the signals $\{s_i\}_{i \in \mathbb{N}} = \{h(\mathbf{x}_{i,T})\}_i$, for $i = 0, 1, 2, \dots$, where $\mathbf{x}_0$ is the initial state. A *homeomorphism* between two manifolds M_1 and M_2 is a continuous bijection $f : M_1 \rightarrow M_2$, where its inverse function $f^{-1} : M_2 \rightarrow M_1$ is also continuous. Moreover, if the homeomorphism and its inverse are smooth, it is a diffeomorphism. An *embedding* is a diffeomorphism from a manifold M_1 into another manifold M_2 , defined as $f : M_1 \rightarrow f(M_1) \subset M_2$. An important point is that embeddings are always injective and without self-intersections. Moreover, our goal is to find an embedding to reconstruct the attractor $\mathbb{A}$ from the signal $\{s_i\}_i$. Given

the certain measurement function h , the following theorem forms the theoretical foundation for attractor reconstruction.

Theorem 1 (Takens). *[12] Let M be an n -dimensional smooth manifold. If v is a vector field on M with flow ψ_t and h is a measurement function on M , then for generic choices of v and h , the differential mapping $F_{h,m} : M \rightarrow \mathbb{R}^m$ of the continuous dynamic system into $\mathbb{R}^m$ is given by:*

$$F_{h,m}(\mathbf{x}) = (h(\mathbf{x}), \left. \frac{d}{dt} \right|_0 h(\psi_t(\mathbf{x})), \dots, \left. \frac{d^{m-1}}{dt^{m-1}} \right|_0 h(\psi_t(\mathbf{x}))) \quad (2.1)$$

which is an embedding when $m = 2n + 1$, where m is the embedding dimension, $\left. \frac{d}{dt} \right|_0$ means the derivatives are evaluated at $t = 0$ and the flow ψ satisfies

$$\left. \frac{d}{dt} \right|_0 \psi_t(\mathbf{x}) = v(\psi_0(\mathbf{x})) \quad (2.2)$$

for every time $t \in \mathbb{R}$.

The above theorem also holds for discrete dynamic systems with a diffeomorphism ψ on a compact n -dimensional manifold M and a measurement function h , for which the embedding is defined as Equation (2.3), where the value of the lag value τ is an integer multiple of the iteration size. The generic in this theory means that the differential mapping $F_{h,m}$ is an open and dense embedding in the set of all mappings under the measurement function h and the flow ψ_t . The best way to understand this is regarding this theorem as a generalization of the Weak Whitney Embedding Theorem [13].

Theorem 2 (Weak Whitney Embedding). *Every n -dimensional manifold M embeds in $\mathbb{R}^{2n+1}$.*

This theorem states that any manifold M can be embedded in $\mathbb{R}^{2n+1}$ without self-intersections given an arbitrary mapping. Whitney proves that the optimal linear bound for the minimum embedding dimension is $2n$. Takens theorem demonstrates that the differential mapping (2.1) satisfies this condition, embedding the compact n -dimensional manifold M into the reconstructed space $\mathbb{R}^{2n+1}$, even when considering finite discrete samples.

For practical use, discrete versions of the differential mapping (2.1) are required when working with signals $\{s_i\}_i$ generated by the discrete flow ξ_T with a specific sampling interval T . The most common approach is the delay-coordinate mapping $F_{h,\tau,m}(\mathbf{x}_{i,\tau}) : M \rightarrow \mathbb{R}^m$, which is defined as:

$$F_{h,\tau,m}(\mathbf{x}_{i,\tau}) = \begin{bmatrix} h(\mathbf{x}_{i,\tau}) \\ h(\mathbf{x}_{(i-1)\cdot\tau}) \\ \vdots \\ h(\mathbf{x}_{(i-m+1)\cdot\tau}) \end{bmatrix} = \begin{bmatrix} h(\mathbf{x}_{i,\tau}) \\ h(\xi_\tau^{-1}(\mathbf{x}_{i,\tau})) \\ \vdots \\ h(\xi_\tau^{-m+1}(\mathbf{x}_{i,\tau})) \end{bmatrix} \quad (2.3)$$

where the parameter $\tau = k \cdot T$, for $k \in \mathbb{Z}$ is the lag value, and m is the embedding dimension. Theoretically, for minimal time delay τ , a linear combination of coordinates can approximate the derivative

such that the delay-coordinate mapping plays the same role as the differential mapping. The well-defined differential mapping is suitable for analytical purposes. This paper explores the properties of shadow manifolds reconstructed through differential mappings while implementing experiments using delay-coordinate mappings.

In practical scenarios, the sampling rate T of signals often cannot be small enough to accurately approximate the differential and higher-order differentials at the given point. However, by selecting an appropriate lag value τ , the delay-coordinate mapping method can obtain the same result in reconstructing the shadow manifold using a sufficiently small τ . Since chaotic dynamic systems consist of highly nonlinear and coupled variables, the signal obtained from the projection function, which serves as the measurement function, has the potential to recover information from other dimensions. The differential mapping method works by separating coupled information and projecting the observed data—via differentiation—in the direction of maximum linear independence, thereby isolating information about variables that are not directly observed. The critical part lies in accurately recovering information from the unknown dimensions using the observed data.

Thus, selecting the lag value τ plays a critical role in reconstructing the shadow manifold. If the lag value is suitable, the delay-coordinate mapping $\mathbf{F}_{h,\tau,n}(\mathbf{x}(t))$ is equivalent to the differential mapping $\mathbf{F}_{h,n}(\mathbf{x})$ under an affine transformation [14], and plays as a diffeomorphism between the shadow manifold and the original attractor. However, the selection of the lag value is not only restricted by external factors like the sampling rate T but also its intrinsic properties. For a continuous-time dynamic system with discrete flow, if τ is too small, the resulting vectors may be highly linear dependent and redundant, leading to a "squeezed" shadow manifold. Conversely, if τ is excessively large, the new coordinates may become essentially unrelated [15], causing the shadow manifold to collapse. Based on the above analysis, we can observe that as τ increases, the shadow manifold undergoes a "stretch-and-fold" process, as depicted in Fig. 1. For convenience, we omit the sampling interval T for τ such that the number of τ shown in this paper refers to the k in the definition, indicating the number of times the sampling interval T , for example, $\tau = 5$ means $\tau = 5T$, where T is the sampling rate or the iteration steps for the discrete dynamic system.

Although the selection of lag value for delay-coordinate mapping is an open problem, several works have been done in this field [16, 17]. The most widely-used method to choose the suitable lag value τ is the mutual information method [18]. The basic idea is to calculate the mutual information between the system's observed values at different lag values and the original observed data and then select the first lag value at which the mutual information transitions from decreasing to increasing as the optimal τ since this represents the lag value that contains sufficient new information while still maintaining some correlation with the original data. This information-based method is theoretically intuitive, but the resulting values often do not correspond to the points at which the shadow manifold is fully stretched before collapsing.

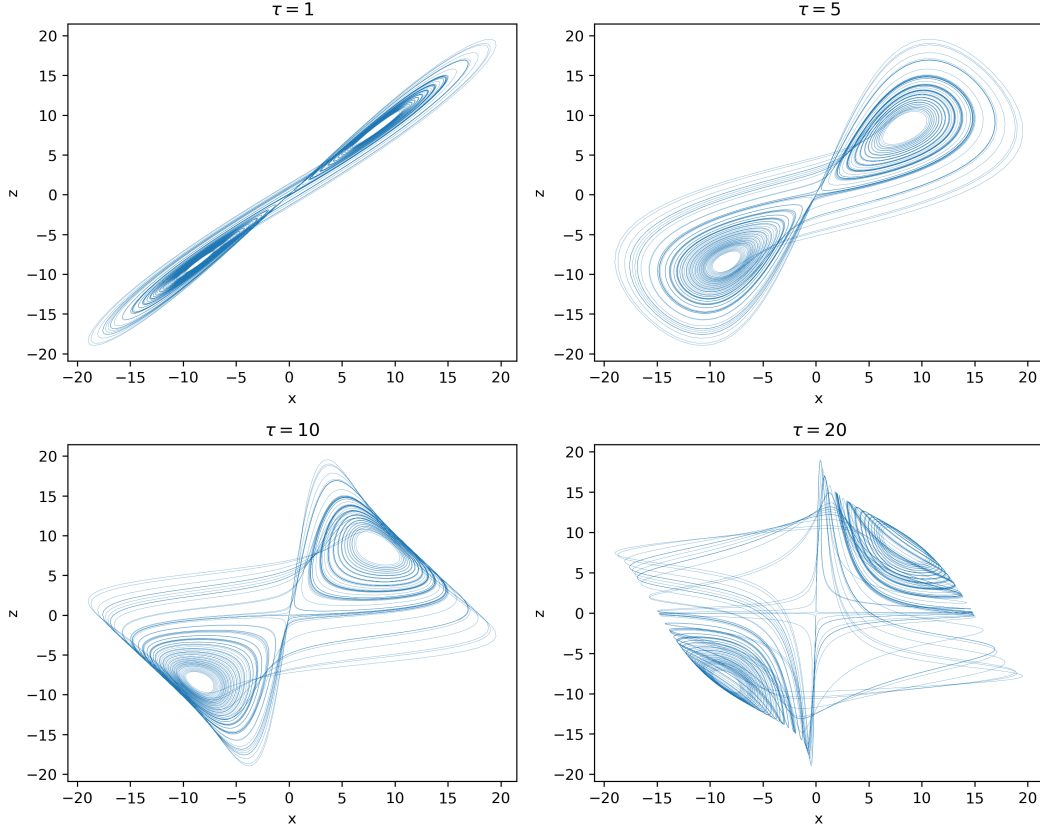


Figure 1: Sequential figures representing the changes in reconstructed shadow manifold $\mathcal{M}_x$ of Lorenz63 system with increasing lag value τ .

Furthermore, in cases where mutual information monotonically decreases with increasing lag value, this method does not work.

According to Takens' theorem, a dynamic system can be embedded in an Euclidean space without self-intersection through any mapping. However, it should be noted that the ideal dimension for embedding an n -dimensional dynamic system is n . In many cases, the optimal dimension for the shadow manifold generated by differential embedding is also n . However, embedding an n -dimensional dynamic system into a higher dimension (greater than n) does not significantly affect the results of CCM, which is consistent with the results of tests using false nearest neighbors [19], as the redundant information introduced by the extra dimensions does not provide extra information about the original dynamic system and, therefore, does not impact the prediction. Therefore, when the dimension of the original dynamic system is known, the differential embedding method can be directly used to obtain a shadow manifold of the same dimension. In cases where information about the dimensions of the original dynamic system is lacking, false nearest neighbors are a practical approach for dimension selection.

Notation	Meaning
$\mathbf{x}_t$	The state of dynamic system
M	Compact n -dimensional smooth manifold
$\mathcal{M}$	Reconstructed shadow manifold
h	Measurement function
ϕ	Iteration map for discrete-time dynamic system
$\mathbf{x}_t$	State at time t , where $\mathbf{x}_t \in M$
ψ_t	Flow generated by the vector field v
v	Smooth vector field on M
$\psi(\mathbf{x}_t)$	Trajectory of the system
ξ_T	Discrete flow with sampling rate T
$\mathbb{A}$	Attractor
$F_{h,m}$	The differential mapping
τ	Lag value for the delay-coordinate mapping $F_{h,\tau,m}$
$F_{h,\tau,m}$	The delay-coordinate mapping

Table 1: A list of mathematical notations in Section 2.

2.2. Convergent Cross Mapping

Here, we demonstrate how Takens' embedding theorem can be applied to detect causal relationships. Given two-time series variables, X and Y , which originate from two deterministic dynamic systems by using projection functions. Here we say X causes Y means the evolution of $\mathbf{y}_t$ depends on $\mathbf{x}_t$. The solid mathematical framework of CCM based on Takens' theorem is developed by Cummins et al. [20] by interpreting causal relationship and bidirectional causal relationship as projection and homeomorphism between two shadow manifolds separately. More precisely, if X unidirectionally causes Y (meaning Y does not cause X), then the joint dynamic system with an attractor $\mathbb{A}_{xy}$ can be constructed by concatenating X and Y as follows:

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = \begin{bmatrix} \psi_x(x_t) \\ \psi_y(x_t, y_t) \end{bmatrix} = \psi_{xy}(x_t, y_t). \quad (2.4)$$

The attractor $\mathbb{A}_x$ is then only a subset of the joint attractor $\mathbb{A}_{xy}$, which is identical to $\mathbb{A}_y$ and there is a noninjective projection $\pi_{yx} : \mathbb{A}_y \rightarrow \mathbb{A}_x$ between them. Consequently, there is also a noninjective map $\Pi_{yx} : \mathcal{M}_y \rightarrow \mathcal{M}_x$ from the shadow manifold $\mathcal{M}_y$ to $\mathcal{M}_x$ and the induced noninjective map $\tilde{\Pi}_{y,x}$. In other words, we can reconstruct the shadow manifold $\mathcal{M}_x$ from $\mathcal{M}_y$, while the inverse is not guaranteed. Furthermore, if X causes Y and Y also causes X , then the attractors $\mathbb{A}_x$ and $\mathbb{A}_y$ are equal to the joint attractor $\mathbb{A}_{xy}$. In this case, the shadow manifold $\mathcal{M}_x$ and $\mathcal{M}_y$ are diffeomorphic to the joint attractor. There exists a homeomorphism $f : \mathcal{M}_x \rightarrow \mathcal{M}_y$ between two shadow manifolds and an induced

homeomorphism f' between $\mathbb{A}_x$ and $\mathbb{A}_{xy}$. Consequently, each shadow manifold can be reconstructed from the other. The corresponding relationship of mapping is shown in Fig. 2.

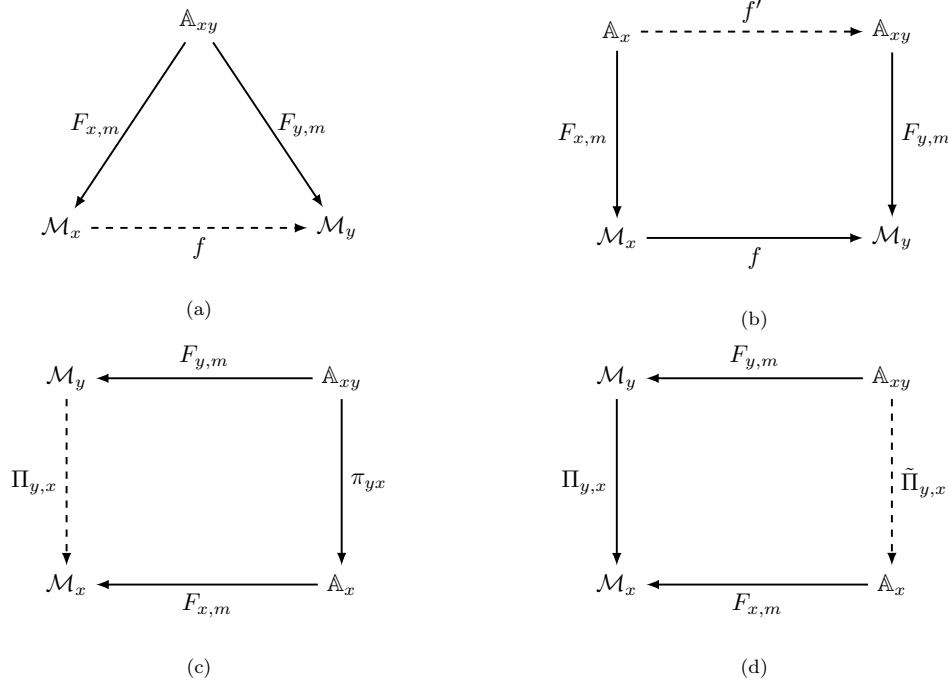


Figure 2: First line: Bidirectional causal relationship $X \Leftrightarrow Y$. (a). Induced homeomorphism f between $\mathcal{M}_x$ and $\mathcal{M}_y$. (b). Induced homeomorphism f' between $\mathbb{A}_x$ and $\mathbb{A}_{xy}$. Last line: Unidirectional causal relationship $X \Rightarrow Y$ when f exists. (c). Induced noninjective projection $\Pi_{y,x}$ between $\mathcal{M}_y$ and $\mathcal{M}_x$. (d). Induced noninjective projection $\tilde{\Pi}_{y,x}$ when $\Pi_{y,x}$ exists.

Based on the above illustration, Sugihara et al. proposed the CCM algorithm to infer causal links between two signals. It aims to do this by using k -nearest-neighbor regression on one shadow manifold to make predictions of another and evaluating the forecasting skill via the Pearson correlation coefficient. Given two time series $\{x_1, x_2, \dots, x_l\}$ and $\{y_1, y_2, \dots, y_l\}$ from two time series variables X and Y with finite length l , CCM tests the causal relationship from $X \Rightarrow Y$ in the following three steps:

1. Reconstruct two shadow manifolds $\mathcal{M}_x$ and $\mathcal{M}_y$ using the delay-coordinate mappings $F_{h,\tau,n}$.
2. For each point $\mathbf{y}_t$ in the shadow manifold $\mathcal{M}_y$, collect its $n+1$ nearest points $\mathbf{y}_{t_i}, i = 1, 2, \dots, n+1$. Estimate $\hat{x}_t|\mathcal{M}_y$ using locally weighted averages of the corresponding x_{t_i} values from the original time series X , calculated as:

$$\hat{x}_t|\mathcal{M}_y = \sum_{i=1}^{n+1} w_i x_{t_i} \quad (2.5)$$

where the weights w_i depends on the distance between $\mathbf{y}_t$ and its i -th closest neighbor in $\mathcal{M}_y$, computed as:

$$w_i = \frac{u_i}{\sum_{j=1}^{n+1} u_j} \quad (2.6)$$

with u_i :

$$u_i = \exp\left\{\frac{-d[\mathbf{y}_t, \mathbf{y}_{t_1}]}{d[\mathbf{y}_t, \mathbf{y}_{t_1}]}\right\} \quad (2.7)$$

where $d[\mathbf{y}_t, \mathbf{y}_{t_1}]$ represents the Euclidean distance between two vectors $\mathbf{y}_t$ and $\mathbf{y}_{t_1}$. Similarly, cross mapping in the other direction is defined by swapping variables.

3. Calculate the Pearson correlation coefficient between $\hat{x}_t|\mathcal{M}_y$ and x_t , and its absolute value is the corresponding forecast skill ρ_{xy} . Repeat 1 and 2 sequentially as the library length l increases, and observe whether the value of ρ_{xy} converges to a nonzero value.

If X causes Y , then as the library length l increases, the shadow manifold becomes denser, allowing the estimator $\hat{x}_t|\mathcal{M}_y$ converge to x_t . Consequently, the Pearson correlation between the ground truth and the predictor gradually converges. In other words, as we increase the length of the signal if the predictability of newly added points aligns with the previous points, it suggests the presence of causal relationship rather than a mere statistical coincidence. The Pearson correlation ρ value may converge to a high plateau; otherwise, the cause-and-effect relationship does not exist.

Moreover, it is important to note that in cases of strong coupling, for instance, when X significantly influences Y , it may exhibit behaviors that are not independent of X . In such scenarios, the causal links may appear reversible, leading to an erroneous interpretation of unidirectional causal relationship as bidirectional. This phenomenon is known as the problem of generalized synchrony. Further research on detecting generalized synchrony within the CCM framework is discussed in [21].

2.3. Related Work about CCM

Several factors, such as noisy data, irregular or sporadic sampling, and improper selection of embedding parameters, significantly influence the correctness of CCM. Here, we introduce some related work as refinements of CCM. A key limitation of CCM is its sensitivity to missing data, necessitating long, uninterrupted data for regular sampling [22]. Latent-CCM has been developed to address this in repeated short, sporadic time series. It leverages a Neural ODE model to learn the inherent dynamics of the data, a method that provides more state space information than the multi-spatial CCM [23]. The TD-CCM method, employing denoising techniques like wavelet methods and empirical mode decomposition, has been proposed in noisy data and trending noise signals.

Furthermore, Feng [24] introduced a Bayesian version of CCM, using a deep Gaussian process based on Cross Mapping Smoothness (CMS) [25]. This approach shifts from the traditional online k-nearest-neighborhood approximation to a deep Gaussian process, underlining the correlation between map smoothness and the strength of time series relationships. It also includes a method for dimension reduction of the shadow manifold using the Gaussian Process Latent Model. The CMS aims to detect causal relations

in short-term time series by equating map smoothness with causal detection, using a Radial Basis Function Network for training. While theoretically sound, this method struggles with differentiating whether the neural network’s fitting ability or significant training errors contribute to the model’s performance, making model selection crucial. Moreover, CMS’s effectiveness for short-term data may be limited due to potential overfitting issues, as smoothness is a necessary but insufficient condition for cross mapping.

Finally, Bulter [9] highlights the importance of auto-predictability and recurrence as prerequisites for employing CCM effectively. Autopredictability ensures the deterministic nature of the system for reliable prediction based on its history, while recurrence validates the presence of neighborhoods on the attractor manifold. Fulfilling these criteria is crucial for successful state space construction, and additional experiments [26, 27] are recommended to test the statistical significance of CCM results.

3. Problem Formulation

For a complex system with interacting variables connected, there are two kinds of causal relationships: direct and indirect. A prototypical example of an indirect causal link is the variable Y within the Lorenz63 system, which acts as an interacting variable, introducing indirect causal links between X and Z . CCM can indiscriminately detect such indirect causal relationship due to the causation transitivity [28, 29]. Although CCM cannot distinguish those two types of links, [28] provides the Partial Cross Mapping method to eliminate indirect causal influence. Based on this, we do not discuss whether the causal relationship detected by CCM is indirect or direct in this paper. This section shows that CCM may not provide the correct causal links for the Lorenz63 system. Notably, this counterexample is not a statistical anomaly. We have identified similar counterexamples originating from various systems; for example, the Chen & Ueta dynamic system [30] wherein CCM fails to detect cross maps introduced by the interacting variable Y .

The classical chaotic Lorenz63 system is:

$$\begin{aligned}\frac{dx}{dt} &= \sigma(y - x) \\ \frac{dy}{dt} &= \rho x - y - xz \\ \frac{dz}{dt} &= xy - \beta z\end{aligned}\tag{3.1}$$

where the parameters are defined as $\sigma = 10, \rho = 28$, and $\beta = \frac{8}{3}$ with the initial condition $\mathbf{x}_0 = [1, 1, 1]$. In the Lorenz63 system, X influences both Y and Z , Y affects both X and Z , and Z affects Y and X . However, as shown in Fig. 3, CCM mistakes bidirectional causal relationships $Y \Leftrightarrow Z$ and $X \Leftrightarrow Z$ as unidirectional $Z \Rightarrow Y$ and $Z \Rightarrow X$. Multiple experiments have also been carried out to avoid the influence of parameters to test the influence of the value of τ and n shown in Fig. 4.

The fluctuation of lag values in a reasonable range may only fluctuate the convergent score of CCM

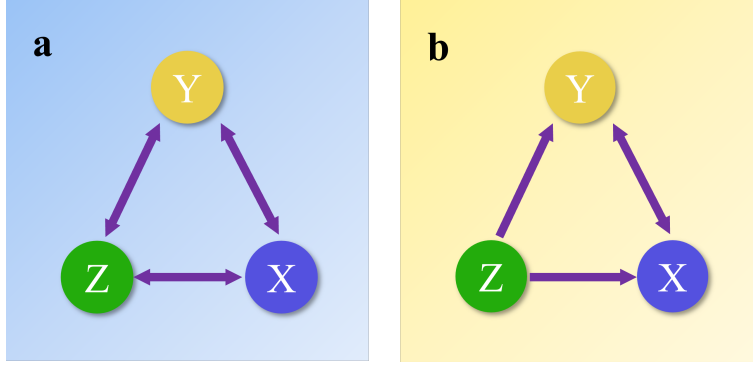


Figure 3: True causal relationships versus CCM's result. a. Graphical model of causal relations within the Lorenz63 system. b. Graphical model of causal relations obtained by CCM. The Lorenz63 system is simulated using the fourth-order Runge-Kutta method, where the temporal domain is $t \in [0, 100]$, the time step for discretization is $\Delta t = 0.01$ which could be regarded as the sampling rate T for continuous-time systems, the initial condition is $\mathbf{x}_0 = [1, 1, 1]$, and $\tau = 9$, $n = 3$ for CCM.

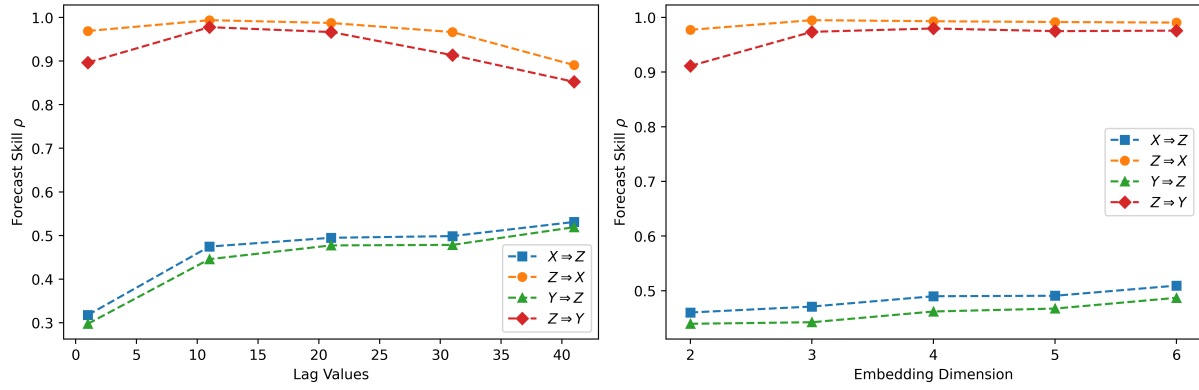


Figure 4: Influence of embedding parameters. Left. The influence of lag value τ when $n = 3$. Right. The influence of embedding dimension n when $\tau = 9$.

in a small range, thus persisting in the same casual direction. However, the condition of the embedding dimension is different. The score is almost stable when the embedding dimension reaches the suitable number ($n = 3$ in this case) because information obtained by increasing the dimension is redundant and never provides additional information for improving predicting skills. The failure in the Lorenz63 system poses a significant query: Is this a statistical anomaly? Further investigation should be done to find more similar counterexamples. Moreover, such counterexamples should be distinct from the Lorenz63 system, neither transformable into it nor derivable from it through linear or nonlinear coordinate transformations (i.e., diffeomorphisms).

Besides the Lorenz63 system, we also discover numerous counterexamples for which CCM fails to detect the correct causal links. This section focuses primarily on three examples: the Chen & Ueta system, the Burke & Shaw system [31], and the three-scroll chaotic attractor [32]. Generally speaking, all three systems are Lorenz-like systems since they share some common properties with the Lorenz63

dynamic system [33], for example, similar butterfly attractors like the Lorenz63 attractor. However, they are topologically different and nonequivalent to the Lorenz63 system. Additional examples are elaborated in Section 5.

The first example is the Chen & Ueta dynamic system; its mathematical representation is formulated as follows:

$$\begin{aligned}\frac{dx}{dt} &= \alpha(y - x) \\ \frac{dy}{dt} &= (\gamma - \alpha)x - xz + \gamma y \\ \frac{dz}{dt} &= xy - \beta z\end{aligned}\tag{3.2}$$

where the initial condition is set as $x_0 = [-10, 0, 37]$. This system, characterized by its chaotic nature and multiple attractors, is shown in Fig. 5, with parameters $\alpha = 35$, $\beta = 3$, and $\gamma = 28$. Despite its

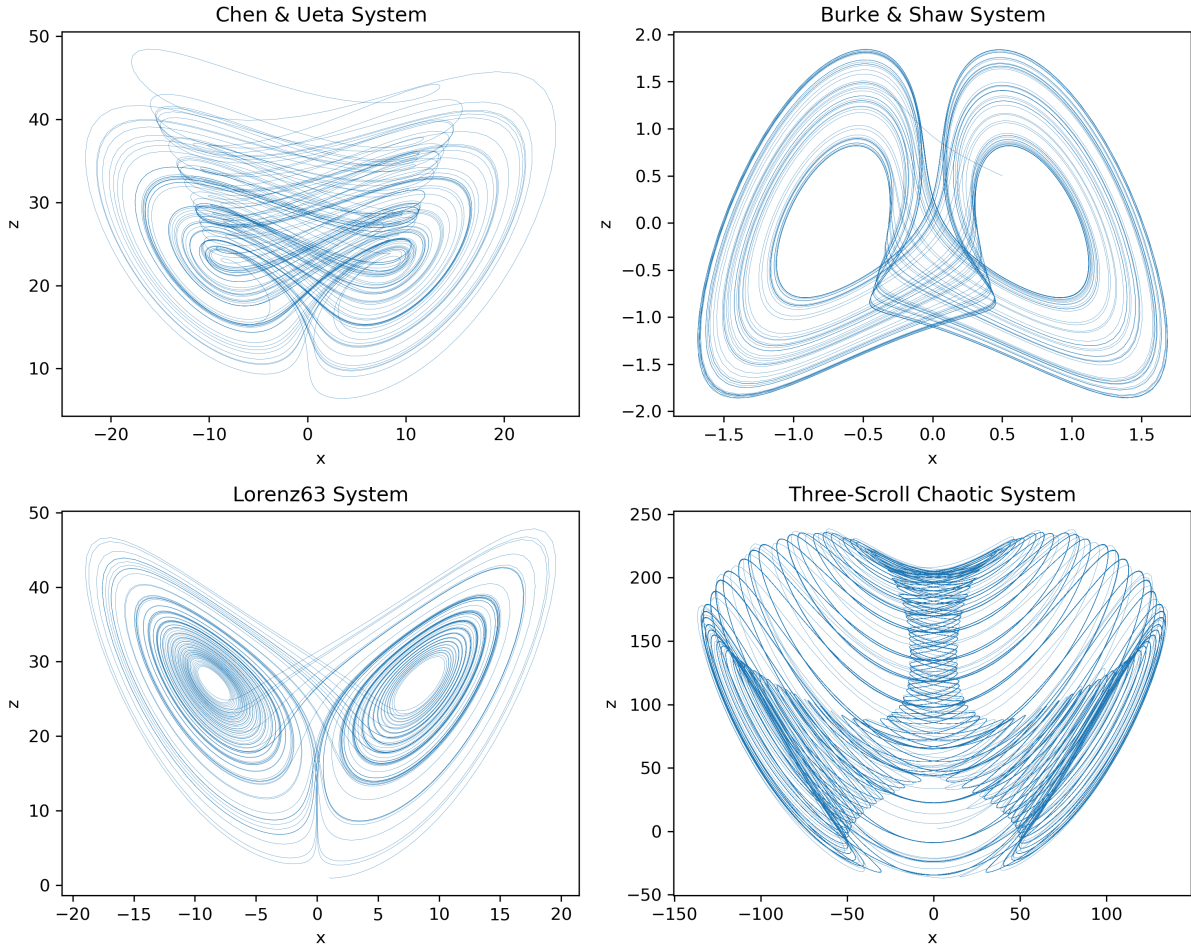


Figure 5: x-z plane projections of attractors. a. Chen & Ueta system, b. Lorenz63 system, c. Burke & Shaw system, d. three-scroll chaotic system.

mathematical similarity to the Lorenz63 system, the Chen & Ueta system exhibits topological differences, primarily in the number of equilibria: three in this case, as opposed to two in the Lorenz63. The causal

structure of this system is the same as the Lorenz63. Another Lorenz-like system, the Burke & Shaw system, is described as follows:

$$\begin{aligned}\frac{dx}{dt} &= -\alpha(x + y) \\ \frac{dy}{dt} &= -y - \alpha xz \\ \frac{dz}{dt} &= \beta + \alpha xy\end{aligned}\tag{3.3}$$

with parameters $\alpha = -10.0$, $\beta = 4.272$ and initial state $\mathbf{x}_0 = [0.5, 0.5, 0.5]$. The attractor produced by equation (3.3) is also topologically different from the one of the Lorenz63 system [34].

We simulated both systems using the fourth-order Runge-Kutta method. For the Chen & Ueta system, the temporal domain is $t \in [0, 50]$, and $\Delta t = 0.005$. Here, we choose a smaller Δt because this system has a larger vector field, and a small step size allows for better illustration and simulation. The lag value and embedding dimension we optimized of the delay coordinate map as $\tau = 20$ and $n = 3$. For the Burke & Shaw system, the trajectories are simulated in the temporal domain $t \in [0, 100]$ with $\Delta t = 0.01$, and the embedding parameters we optimized are $\tau = 10$ and $n = 3$.

The last counterexample is the three-scroll chaotic system:

$$\begin{aligned}\frac{dx}{dt} &= 40(y - x) + 0.16xz \\ \frac{dy}{dt} &= 55x + 20y - xz \\ \frac{dz}{dt} &= -0.65x^2 + xy + \frac{11}{6}z\end{aligned}\tag{3.4}$$

with the initial point $\mathbf{x}_0 = [2, 2, 2]$, temporal domain $t = [0, 150]$, $\Delta t = 0.0015$, and $\tau = 20$, $n = 3$. This dynamic system generates a toroidal attractor bounded by a genus-3 torus, which is shown in Fig. 5. Six causal links exist within this chaotic system as $X \Leftrightarrow Y$, $X \Leftrightarrow Z$, and $Y \Leftrightarrow Z$. Causal results of these three systems obtained through CCM are listed in Table 2, where we can see that CCM fails to provide the correct result. These instances suggest that the failure observed in the Lorenz63 is not merely an anomaly but indicative of an underlying commonality that warrants further investigation.

4. Explanation and Proposed Solution

In the previous section, we proposed that CCM may provide incorrect causal links for several cases. Here, we explain the reason and extend it to general cases. We point out that for the n -fold dynamic system, which is equivariant under a nontrivial cyclic group $\mathbb{C}_n$ of order n , the differential mapping of the invariant variable is a n -to-one mapping rather than an embedding. Consequently, when there is a bidirectional causation between invariant and non-invariant variables, there is no longer a homeomorphism between two shadow manifolds, misleading a wrong causal relationship between two variables. To address this problem, we propose a practical method for the case when the dynamic system is equivariant under the order two cyclic group $\mathbb{C}_2$.

4.1. Preliminary

Here, we briefly introduce some basic definitions and properties that will be used later. *Group theory* is the mathematical tool to study symmetry. For a finite group G , the *order* of the group refers to the number of elements it contains, denoted by $|G|$. A *group action* " $\cdot$ " of a group G on $\mathbb{R}^n$ is defined as a map $G \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ which satisfies: (a). $e \cdot x = x$ for all $x \in \mathbb{R}^n$, where e is the identity element of the group G and (b). $(gh) \cdot x = g \cdot (h \cdot x)$ for $x \in \mathbb{R}^n$ and $g, h \in G$. Here $g \cdot x$ means the image of x under the action of $g \in G$.

A dynamic system $\dot{\mathbf{x}} = v(\mathbf{x})$ with flow $\psi_t(\mathbf{x})$ is $|G|$ -fold *symmetric* or *G-equivariant* if

$$g \cdot \dot{\mathbf{x}} = g \cdot v(\mathbf{x}) = v(g \cdot \mathbf{x}). \quad (4.1)$$

for every $g \in G$ and $\mathbf{x} \in \mathbb{R}^n$. For every $g \in G$, the *fundamental domain* $\mathcal{D}_g \in \mathbb{R}^n$ under the group G is a representative region such that the action of G tessellates the entire space without overlaps. Formally, it satisfies:

$$g_1 \cdot \mathcal{D}_{g_1} \cap g_2 \cdot \mathcal{D}_{g_2} \neq \emptyset \Rightarrow g_1 = g_2 \quad (4.2)$$

for $g_1, g_2 \in G$. For example, the fundamental domain of the Lorenz63 system can be chosen as $\mathcal{D}_{R_z} = \{(x, y, z) | x \geq 0\}$. If G is a reflection group, then the dynamic system is *reflection symmetric*. If G is the cyclic group $\mathbb{C}_n$, then the dynamic system is *rotation symmetric*. If the dynamic system is equivariant under the inversion group $G = \{e, P\}$ where P denotes the inversion operation as $P : (x_1, x_2, \dots, x_n) \rightarrow (-x_1, -x_2, \dots, -x_n)$, then the dynamic system is *inversion symmetric*.

There are several crucial factors we should notice for a symmetric attractor. Firstly, the group G should satisfy certain algebraic conditions when the reflection planes exist such that the G -equivariant dynamic system may generate a symmetric attractor [35, 36]. Moreover, not all symmetric attractors are connected because of the existence of the reflection plane of the symmetry group. The system gives a typical example of this [37]:

$$\begin{aligned} \dot{x} &= x - xy \\ \dot{y} &= z \\ \dot{z} &= -y - az + x^2 \end{aligned} \quad (4.3)$$

which is G -equivariant under the group $G = \{e, \sigma_x\}$, where $\sigma_x : (x, y, z) \rightarrow (-x, y, z)$, producing the so-called "kissing" attractor shown in Fig. 6. It is evident that the symmetric pair of strange attractors nearly touch each other on opposite sides of the reflection plane $x = 0$, as $\dot{x} = 0$ in this plane. Consequently, a trajectory originating from an initial point in one part of its disconnected symmetric attractor may not exhibit symmetry behavior. Based on the above analysis, when the symmetry group has a reflection plane, as in such cases, the corresponding trajectory may not show the expected symmetry. In this paper, we focus on investigating the impact of the symmetry of rotation symmetric systems with no reflection plane on detecting causal relationships by CCM.

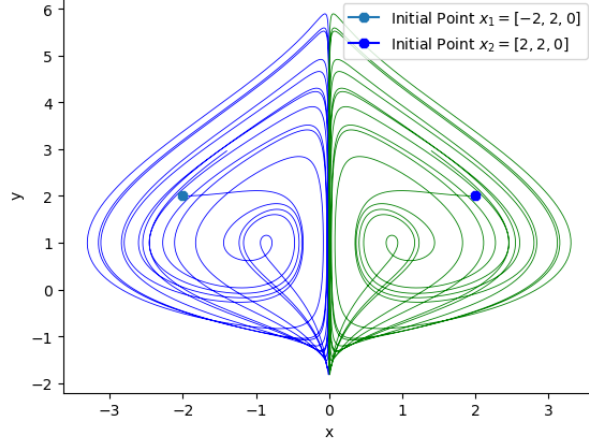


Figure 6: A symmetric pair of attractor "kissing" of the reflection equivariant system, when $a = 0.7$ and two initial values $\mathbf{x}_1 = [-2, 2, 0]$ and $\mathbf{x}_2 = [2, 2, 0]$. Two parts of trajectories can become arbitrarily close but never cross the plane $x = 0$.

The following are two useful results for differential mapping. For $g \in G$, if the function u satisfying:

$$u(g \cdot \mathbf{x}) = \pm u(\mathbf{x}), \quad (4.4)$$

we say that u defines the (*even or odd*) *parity* under g . For a G -equivariant dynamic system $\dot{\mathbf{x}} = v(\mathbf{x})$ with flow $\psi_t(\mathbf{x})$, by the definition of the derivative, we have:

$$\left. \frac{d}{dt} \right|_0 u[\psi_t(\mathbf{x})] = \lim_{t \rightarrow 0} \frac{u[\psi_t(\mathbf{x})] - u(\mathbf{x})}{t} = \lim_{t \rightarrow 0} \frac{u[\mathbf{x} + tv(\mathbf{x})] - u(\mathbf{x})}{t} \quad (4.5)$$

where $\left. \frac{d}{dt} \right|_0$ denotes that the derivative is evaluated at time $t = 0$, and $\psi_t(\mathbf{x})$ represents the flow of the system. By using the property in (4.1), and transforming $\mathbf{x}$ to $g \cdot \mathbf{x}$, we obtain:

$$\begin{aligned} \left. \frac{d}{dt} \right|_0 u[\psi_t(g \cdot \mathbf{x})] &= \lim_{t \rightarrow 0} \frac{u[\psi_t(g \cdot \mathbf{x})] - u(g \cdot \mathbf{x})}{t} \\ &= \lim_{t \rightarrow 0} \frac{u[g \cdot \mathbf{x} + tv(g \cdot \mathbf{x})] - u(g \cdot \mathbf{x})}{t} \\ &= \lim_{t \rightarrow 0} \frac{u[g \cdot (\mathbf{x} + tv(\mathbf{x}))] - u(g \cdot \mathbf{x})}{t} \\ &= \left. \frac{d}{dt} \right|_0 u[g \cdot \psi_t(\mathbf{x})] \end{aligned} \quad (4.6)$$

Thus, it implies that the time derivative of u also preserves the same parity. By mathematical induction, we can conclude that the n -th derivative of u inheres the same parity.

Through the differential mapping $F_{h,n}$ defined as (2.1), the shadow manifold is constructed from consecutive derivatives. The corresponding vector field of the shadow manifold is:

$$V^i = \left. \frac{d}{dt} \right|_0 F_{h,n}^i(\psi_t(\mathbf{x})) \quad (4.7)$$

where $F_{h,n}^i$ and V^i denote the i -th component of $F_{h,n}(\mathbf{x})$ and V separately, and we have the following rules:

$$\left. \frac{d}{dt} \right|_0 F_{h,n}^2(\psi_t(\mathbf{x})) = F_{h,n}^3(\mathbf{x}). \quad (4.8)$$

Since $V^1 = F^2$, by mathematical induction, we have the general rule:

$$V^i(\mathbf{x}) = \frac{d}{dt}\bigg|_0 F_{h,n}^i(\psi_t(\mathbf{x})) = F_{h,n}^{i+1}, \quad i < n \quad (4.9)$$

Thus, the n -dim shadow manifold always has the analytic canonical form:

$$\begin{aligned} V^1 &= \frac{d}{dt}\bigg|_0 F_{h,n}^1 = F_{h,n}^2, \\ V^2 &= \frac{d}{dt}\bigg|_0 F_{h,n}^2 = F_{h,n}^3, \\ &\vdots \\ V^{n-1} &= \frac{d}{dt}\bigg|_0 F_{h,n}^{n-1} = F_{h,n}^n, \\ V^n &= \frac{d}{dt}\bigg|_0 F_{h,n}^n = f(F_{h,n}^1, \dots, F_{h,n}^n). \end{aligned} \quad (4.10)$$

for some function f , which is the only unknown term.

The following theorem explains the preservation of symmetry in the reconstructed system when using differential mapping to reconstruct a symmetric system from a single observation function [38].

Theorem 3 (Cross). *A differential reconstruction of any nonlinear dynamic system preserves, at most, a two-fold symmetry.*

Specifically, there are only two possibilities for reconstructed shadow manifolds: either (a). The shadow manifold has no symmetry, and (b). The shadow manifold is equivariant under the inversion group $G = \{e, P\}$.

4.2. Explanation

We begin with a few key observations. First, all three dynamic systems are two-fold symmetric, specifically three-dimensional ones that rotate around the z -axis. In such systems, the differential mapping $F_{z,n}$ generates two symmetric shadow manifolds, $\mathcal{M}_x$ and $\mathcal{M}_y$, along with a nonsymmetric one, $\mathcal{M}_z$. Second, all trajectories of these shadow manifolds are connected. These observations lead us to investigate the properties of symmetric systems under differential mappings. Here, we give the following remark as an explanation for the general case.

Proposition 1 . *If the dynamic system $\dot{\mathbf{x}} = v(\mathbf{x}) \in \mathbb{R}^n$ is symmetric under the order-two cyclic group $\mathbb{C}_2 = \{e, R_{x_n}(\pi)\}$, where $R_{x_n}(\pi) : (x_1, x_2, \dots, x_n) \rightarrow (-x_1, -x_2, \dots, x_n)$, CCM may mislead the bidirectional causal relationship $x_n \Leftrightarrow x_i$ between the invariant variable x_n and $x_i, i \neq n$ as unidirectional relationship $x_n \Rightarrow x_i$.*

Proof. It is sufficient to show that there is a noninjection mapping $\Pi_{x_i x_n} : \mathcal{M}_{x_i} \rightarrow \mathcal{M}_{x_n}$ between $\mathcal{M}_{x_i}$ and $\mathcal{M}_n$. Since the dynamic system is two-fold symmetric under $\mathbb{C}_2$, then when the measurement function

h is taken to be the x_n -coordinate projection, it defines even parity as $h(R_{x_n}(\pi) \cdot \mathbf{x}) = h(\mathbf{x})$. Using the results (4.6) and (4.9), the differential mapping $F_{x_i,n}$ and the new vector field also inherit the same even parity as h . Therefore, the $F_{x_i,n}$ is noninjective and is two-to-one at the nicest case. Because of the bidirectional causal relationship $x_n \Leftrightarrow x_i$, we can find the induced homeomorphism $f : \mathbb{A}_{x_n} \rightarrow \mathbb{A}_{x_i x_n}$ and $f' : \mathbb{A}_{x_i x_n} \rightarrow \mathbb{A}_{x_n}$. Then the homeomorphism $\tilde{f} : \mathcal{M}_{x_i} \rightarrow \mathbb{A}_{x_n}$ can be obtained by $\tilde{f} : f' \circ F_{x_i,n}^{-1}$, since $F_{x_i,n}$ is a diffeomorphism. So the map

$$\begin{aligned}\Pi_{x_i x_n} &:= F_{x_i,n} \circ \tilde{f} \\ &= F_{x_i,n} \circ f' \circ F_{x_i,n}^{-1}\end{aligned}\tag{4.11}$$

constructed from Fig. 7 is the desired noninjective mapping since $F_{x_i,n}$ is noninjective. ■

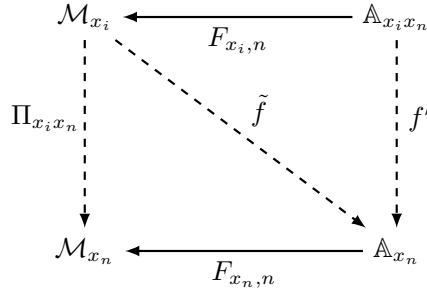


Figure 7: Induced projection $\Pi_{x_i x_n} : \mathcal{M}_{x_i} \rightarrow \mathcal{M}_{x_n}$, homeomorphism $f' : \mathbb{A}_{x_i x_n} \rightarrow \mathbb{A}_{x_n}$ and $\tilde{f} : \mathcal{M}_{x_i} \rightarrow \mathbb{A}_{x_n}$.

Next, we use the Lorenz63 and the Burke & Shaw systems to show concrete examples. Both systems are two-fold symmetric system under the cyclic group $\mathbb{C}_2 = \{e, R_z(\pi)\}$. For the Lorenz63 system, when the measurement function is taken to be the x -coordinate projection which defines odd parity under the action of $R_z(\pi)$, the corresponding differential mapping $F_{x,3} : (x, y, z) \rightarrow (u, v, w)$ is as follows:

$$\begin{aligned}u &= x \\ v &= \sigma(y - x) \\ w &= \sigma[(\rho + \sigma)x - (\sigma + 1)y - xz]\end{aligned}\tag{4.12}$$

where we use (u, v, w) to denote new coordinate to avoid confusion. The new vector field of the x -induced shadow manifold also referred to as the induced Lorenz system, is:

$$\begin{aligned}\dot{u} &= v \\ \dot{v} &= w \\ \dot{w} &= \beta\sigma(\rho - 1)u - \beta(\sigma + 1)v - (1 + \beta + \sigma)w - u^2v - \sigma u^3 + \frac{v}{u}(w + (1 + \sigma)v)\end{aligned}\tag{4.13}$$

where the domain is $\{(u, v, w) \in \mathbb{R}^3 | u \neq 0\}$.

The same process applies to obtain the invariant system generated by $F_{z,3}$ for the Burke & Shaw

equation (3.3) as:

$$\begin{aligned}\dot{u} &= -(\alpha + 1)u - \alpha(1 - w)v + (1 - \alpha)\rho \\ \dot{v} &= \alpha(1 - w)u - \alpha(1 + w)\rho - (\alpha + 1)v \\ \dot{w} &= \frac{\alpha}{2}v + \beta\end{aligned}\tag{4.14}$$

where $\rho = \sqrt{u^2 + v^2}$. The invariant image of Lorenz63 and Burke & Shaw is shown in Fig 8.

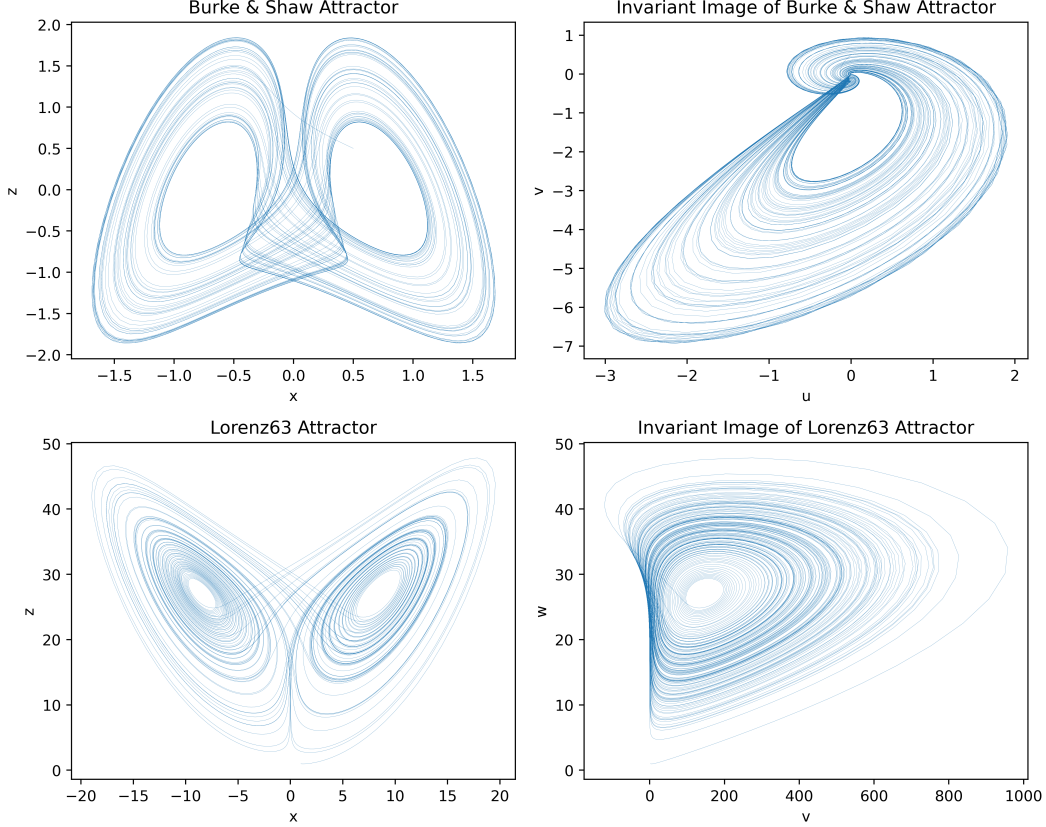


Figure 8: Left: Strange attractors generated by the Lorenz63 system (3.1) and Burke & Shaw system (4.14). Right: Invariant images of the Lorenz63 and Burke & Shaw systems.

We have systematically examined the structure of two-fold rotation symmetric systems in $\mathbb{R}^3$, the reconstruction of their shadow manifolds, and the characteristics of the differential mapping $F_{h,3}$. This analysis explains why CCM fails to provide a correct causal graph consistent with other counterexamples. Specifically, the variable that defines even parity and remains invariant under the representation of the order-two group lacks sufficient information about the attractor's symmetry, making it unable to distinguish between different fundamental domains in the phase space.

4.3. Proposed Method

The previous analysis demonstrates that CCM may misinterpret the bidirectional causal relationship as unidirectional when the underlying dynamic system exhibits two-fold symmetry under the cyclic group

$\mathbb{C}_2$. The key to resolving this issue lies in reconstructing a one-to-one homeomorphism between two shadow manifolds, enabling CCM to provide correct causal relationship. An intuitive approach is to recover the rotational symmetry of $\mathcal{M}_z$, modded by the delay coordinate map, thereby restoring the one-to-one mapping. Existing methods achieve this by incorporating symmetry information from another variable, for example, using (y, z) or (x, z) to reconstruct the Lorenz63 system. However, when applying CCM to detect causal relationships, it is crucial to avoid introducing information from other variables of shadow manifold reconstruction. In other words, we cannot use the shadow manifold reconstructed from the set (x, z) to detect the causal relationship between X and Z . We find another way to recover the one-to-one mapping without introducing additional information about another variable and propose the segment convergent cross mapping (sCCM) method to address this issue.

Instead of recovering the modded-out symmetry to reconstruct the one-to-one mapping, we propose another way by projecting the inversion-symmetric shadow manifold (i.e. $\mathcal{M}_x$ of Lorenz63) into its fundamental domain $\mathcal{D}_P$ and its reflected counterpart $P \cdot \mathcal{D}_P$. This segmentation results in two equivalent time sets:

$$[t_1] = \{t \in T | \mathbf{x}_t \in D_{R_z}\}, \quad [t_2] = \{t \in T | \mathbf{x}_t \in R_z \cdot D_{R_z}\}. \quad (4.15)$$

Moreover, two equivalent sub-manifolds, $\mathcal{M}_{x|[t_1]}$ and $\mathcal{M}_{x|[t_2]}$, depending on which fundamental domain the trajectory is in at time t . Next, we split the invariant shadow manifold $\mathcal{M}_z$ into two sub-manifolds $\mathcal{M}_{z|[t_1]}$ and $\mathcal{M}_{z|[t_2]}$, corresponding to the time sets $[t_1]$ and $[t_2]$. In other words, the segmentation of $\mathcal{M}_{z|[t_i]}$ depends on the indices of $\mathcal{M}_{x|[t_i]}$, the final step is applying CCM on those two pairs of sub-manifolds for detecting the causal relationship. Through this segmentation process, we eliminate the influence of symmetry, which makes the reconstruction mapping two-to-one (under the nicest case) during the reconstruction process. Moreover, we preserve the local geometry structure of each sub-manifold during the segmentation process. If the bidirectional causal relationship exists, that is the one-to-one mapping $\tilde{f} : \mathcal{M}_{x|[t_i]} \rightarrow \mathcal{M}_{z|[t_i]}$ exists, then the forecasting skill score ρ obtained on two separate experiments should both high and closed to one. Here, we calculate the average as the final score as the final score.

In a practical scenario, without any analytic information about the dynamic system, the first step is to check the recurrence property of the dynamic system via the recurrence plot or the method shown in [9], which is a necessary condition for the usage of CCM and details are discussed in Appendix Appendix B. Next, it is necessary to check the symmetry property of the time series by checking whether the shadow manifolds pose inversion symmetry, which can be achieved by using the indicator shown in [39]. If one poses inversion symmetry while the other is non-symmetric, our method should be applied under this circumstance to get the correct causal relationship.

Here, we use the k-means clustering method as the implementation for the segmentation since k-means performs well on datasets with symmetric properties [40]. For the Lorenz63 system, the pipeline is illustrated in Fig. 9. Specifically, we first segment the symmetric shadow manifold $\mathcal{M}_x$ into two

sub-manifolds, $\mathcal{M}_{x|t_1}$ and $\mathcal{M}_{x|t_2}$, using k-means clustering methods ($k=2$). Subsequently, the indices derived from $\mathcal{M}_{x|t_i}$ are applied to segment $\mathcal{M}_z$ into $\mathcal{M}_{z|t_1}$ and $\mathcal{M}_{z|t_2}$.

Algorithm 1 Segment Convergent Cross Mapping for two-fold symmetric system under $\mathbb{C}_2$

Input: Give two data sets $\{h_x\}$ and $\{h_z\}$ generated from the coordinate projections, and embedding parameters τ and n

Output: The causal relationship between X and Z

- 1: Test the recurrence property for the dynamic system.
 - 2: Use delay-coordinate mappings $F_{x,\tau,3}$ and $F_{z,\tau,3}$ to embed those two data sets into two shadow manifolds $\mathcal{M}_x$, and $\mathcal{M}_z$.
 - 3: **if** One is inversion symmetric $\mathcal{M}_x$ and the other is non-symmetric $\mathcal{M}_z$ **then**
 - 4: Use k-means clustering method to partition the symmetric shadow manifold $\mathcal{M}_x$ into two parts, $\mathcal{M}_{x|t_1}$ and $\mathcal{M}_{x|t_2}$.
 - 5: Segment $\mathcal{M}_z$ into $\mathcal{M}_{z|t_1}$ and $\mathcal{M}_{z|t_2}$ according to the time index derived from $\mathcal{M}_{x|t_i}$.
 - 6: Do CCM for $(\mathcal{M}_{x|t_1}, \mathcal{M}_{z|t_1})$; $(\mathcal{M}_{x|t_2}, \mathcal{M}_{z|t_2})$ and obtain $\rho_{x_1 z_1}$, $\rho_{z_1 x_1}$, $\rho_{x_2 z_2}$, $\rho_{z_2 x_2}$
 - 7: Combine the results as $\rho_{xz} = \frac{\rho_{x_1 z_1} + \rho_{x_2 z_2}}{2}$ and $\rho_{zx} = \frac{\rho_{z_1 x_1} + \rho_{z_2 x_2}}{2}$
 - 8: **else**
 - 9: Do CCM on $\mathcal{M}_x$ and $\mathcal{M}_z$.
 - 10: **end if**
-

Another point to note is that the shadow manifold is when the length of the signal is limited, the number of trajectories lying in the fundamental domain $\mathcal{D}_P$ and its reflection counterpart $P \cdot \mathcal{D}_P$ of the inversion symmetric shadow manifold may differ; that is, one domain may be denser than the other. This can be clearly observed in Fig. 9, where $\mathcal{M}_{x|t_2}$ looks denser than $\mathcal{M}_{x|t_1}$. This difference is due to the asymmetric periodic motions of the Lorenz63 system; however, as the number of observations increases, both $\mathcal{M}_{x|t_i}$ and $\mathcal{M}_{z|t_i}$ would become denser to the same extent.

For attractors in higher-dimensional space than three, the complexity far exceeds that of three-dimensional attractors, and the observability dramatically affects the effectiveness of CCM since the quality of reconstructed shadow manifolds faces substantial challenges due to the intricate complicated dynamics. For instance, the system is often unobservable from a single variable, and multiple variables are usually required to obtain an effective embedding [41]. More details are discussed in Appendix Appendix A.

5. Experiments

In this section, we test the efficiency of our method through several numerical simulation results and one real-world case. Moreover, we also discuss symmetric systems in the higher dimension.

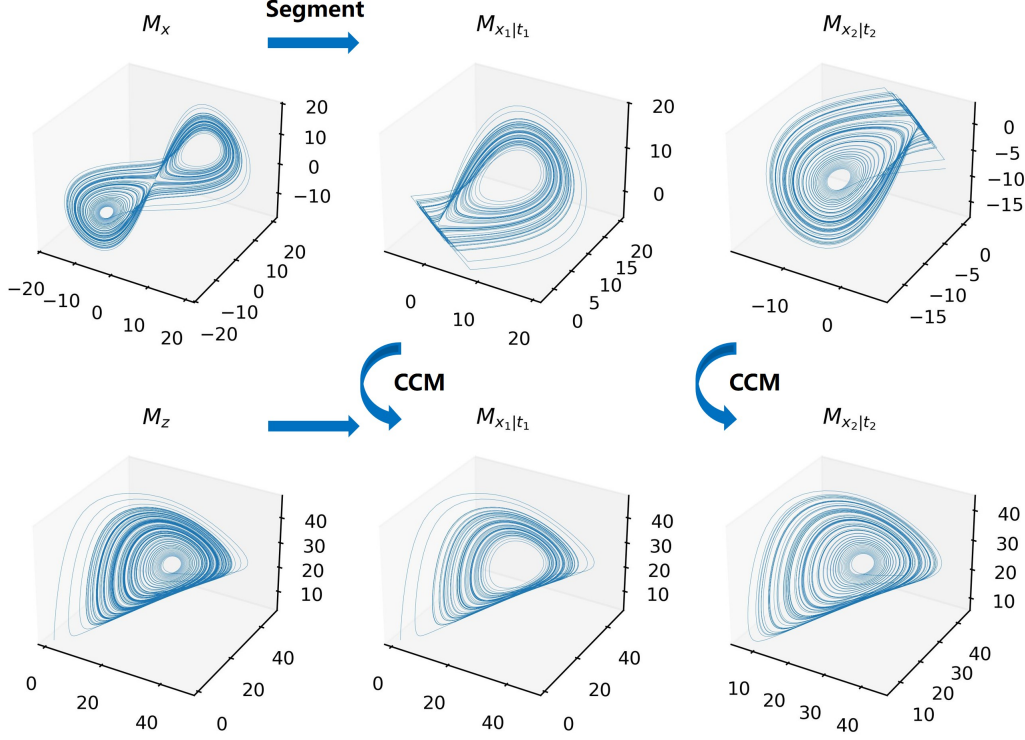


Figure 9: A schematic depicting the details for implementation of sCCM between $\mathcal{M}_x$ and $\mathcal{M}_z$ for the Lorenz63.

5.1. Case 1: Rotation symmetric systems under order-two group $\mathbb{C}_2$ in $\mathbb{R}^3$

We implement our method on several $\mathbb{C}_2$ symmetric systems in $\mathbb{R}^3$, including the counterexamples mentioned in Section 3 and causal relationships between X and Z are shown in Table 2. Similarly, the correct relationship between Y and Z can be obtained using the same steps. The Pearson correlation coefficients $\rho_{x_1 z_1}$, $\rho_{x_2 z_2}$, $\rho_{z_1 x_1}$, $\rho_{z_2 x_2}$ all converge to nonzero values close to 1 as the length of the library increases, indicating a bidirectional causal relationship between X and Z . We can observe that CCM misleads the bidirectional causal relationship between two signals as unidirectional due to the influence of modded-out symmetry, while sCCM provides the correct causal link.

Moreover, we test the robustness of sCCM by varying noise levels in the Lorenz63 system. Gaussian noise $n \sim \mathcal{N}(\mu, \sigma)$ is added to the observations and run our experiments for ten times. In our tests, μ is set to zero, and we explore different noise levels with standard deviations σ at 0.1, 0.5, and 1. The trajectories are generated as described in Section 3, ensuring consistency in other experiments. Results are shown in Table 3. The findings indicate that our method maintains considerable effectiveness even under moderate noise levels, underscoring its robustness. This conclusion is consistent with the other examples and scenarios in this study section.

Dynamic systems	CCM	Segment CCM
Lorenz63	$\rho_{xz} = 0.471, \rho_{zx} = 0.995$	$\rho_{xz} = 0.992, \rho_{zx} = 0.997$
Chen & Ueta	$\rho_{xz} = 0.244, \rho_{zx} = 0.997$	$\rho_{xz} = 0.965, \rho_{zx} = 0.997$
Burke & Shaw	$\rho_{xz} = 0.283, \rho_{zx} = 0.999$	$\rho_{xz} = 0.989, \rho_{zx} = 0.999$
Three scroll chaotic attractor	$\rho_{xz} = 0.131, \rho_{zx} = 0.928$	$\rho_{xz} = 0.869, \rho_{zx} = 0.926$
Wang [42]	$\rho_{xz} = 0.013, \rho_{zx} = 0.999$	$\rho_{xz} = 0.994, \rho_{zx} = 0.999$
Shimizu & Morioka [43]	$\rho_{xz} = 0.535, \rho_{zx} = 0.983$	$\rho_{xz} = 0.965, \rho_{zx} = 0.982$
Rucklidge [44]	$\rho_{xz} = 0.392, \rho_{zx} = 0.973$	$\rho_{xz} = 0.939, \rho_{zx} = 0.973$
Sprott B [45]	$\rho_{xz} = 0.034, \rho_{zx} = 0.999$	$\rho_{xz} = 0.995, \rho_{zx} = 0.999$
Sprott C [45]	$\rho_{xz} = 0.326, \rho_{zx} = 0.995$	$\rho_{xz} = 0.970, \rho_{zx} = 0.996$
Rikitake [46]	$\rho_{xz} = 0.582, \rho_{zx} = 0.991$	$\rho_{xz} = 0.988, \rho_{zx} = 0.987$
Liu & Yang [47]	$\rho_{xz} = 0.358, \rho_{zx} = 0.989$	$\rho_{xz} = 0.983, \rho_{zx} = 0.967$
Chongxin <i>et al.</i> [48]	$\rho_{xz} = 0.185, \rho_{zx} = 0.998$	$\rho_{xz} = 0.968, \rho_{zx} = 0.999$
Lü, Chen & Cheng [49]	$\rho_{xz} = 0.035, \rho_{zx} = 0.986$	$\rho_{xz} = 0.946, \rho_{zx} = 0.986$

Table 2: Results for a range of $\mathbb{C}_2$ symmetric systems in $\mathbb{R}^3$.

5.2. Case 2: Rotation Symmetric Systems under order-two group $\mathbb{C}_2$ in High Dimension

In this part, we apply our method to three high-dimensional, two-fold rotation symmetric systems, including two four-dimensional and one five-dimensional system. These high-dimensional dynamic systems are visualized by projecting them onto lower-dimensional planes. Observability is critical in determining whether a shadow manifold can be successfully reconstructed in such cases. Consequently, poor observability can lead to incorrect results from CCM.

The first case is a four-dimensional dissipative Lorenz-like dynamic system, representing a self-excited oscillatory modular circuit with the following mathematical formulation [50]:

$$\begin{aligned}
\frac{dx}{dt} &= a(y - x) \\
\frac{dy}{dt} &= xz + w \\
\frac{dz}{dt} &= b - xy \\
\frac{dw}{dt} &= yz - cw
\end{aligned} \tag{5.1}$$

where parameters are set as $a = 6, b = 11, c = 5$, with the initial value $(x, y, z, w) = (10, 10, 0, 0)$. We simulate this system using the fourth-order Runge-Kutta method over the temporal domain $t = [0, 100.0]$ with a time step of $\Delta t = 0.01$. This system is rotation symmetric under the order-two group $\mathbb{C}_2 = \{e, R_z(\pi)\}$, where $R_z : (x, y, z, w) \rightarrow (-x, -y, z, -w)$, producing two distinct types of attractors: four-wing and two-wing attractors which are shown in Fig. 10.

noise level (σ)	CCM	Segment CCM
0.1	$\rho_{xz} = 0.0993(0.010), \rho_{zx} = 0.991(0.0003)$ $\rho_{yz} = 0.140(0.0126), \rho_{zy} = 0.961(0.0014)$	$\rho_{xz} = 0.988(0.0004), \rho_{zx} = 0.992(0.0002)$ $\rho_{yz} = 0.990(0.0005), \rho_{zy} = 0.965(0.0014)$
0.5	$\rho_{xz} = 0.0183(0.0135), \rho_{zx} = 0.9665(0.0012)$ $\rho_{yz} = 0.0430(0.0276), \rho_{zy} = 0.934(0.0013)$	$\rho_{xz} = 0.973(0.0007), \rho_{zx} = 0.962(0.0014)$ $\rho_{yz} = 0.976(0.0010), \rho_{zy} = 0.939(0.0012)$
1	$\rho_{xz} = 0.0438(0.0213), \rho_{zx} = 0.931(0.0019)$ $\rho_{yz} = 0.0263(0.0176), \rho_{zy} = 0.8904(0.0039)$	$\rho_{xz} = 0.942(0.0018), \rho_{zx} = 0.934(0.0022)$ $\rho_{yz} = 0.949(0.0012), \rho_{zy} = 0.892(0.0003)$

Table 3: Performance of Segment CCM under different noise levels. We show the mean over ten runs, and the values in parentheses represent the corresponding standard deviation.

From equation (5.1), we infer bidirectional causal links between variables: $X \Leftrightarrow Z$, $Y \Leftrightarrow Z$, and $W \Leftrightarrow Z$. However, CCM results show unidirectional causal links $Z \Rightarrow X$, $Z \Rightarrow Y$, and $Z \Rightarrow W$, because of the nonsymmetric shadow manifold $\mathcal{M}_z$. By applying sCCM, we successfully recover the correct causalities, and results are shown in Table 4.

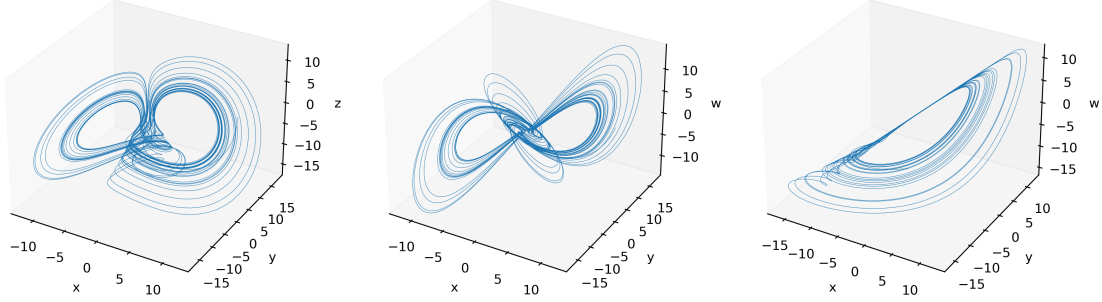


Figure 10: From left to right: Two-wing butterfly chaotic attractors in the (x, y, z) plane, the four-wing butterfly chaotic attractors in the (x, y, w) plane, and the (x, y, w) projection of the reconstructed shadow manifold $\mathcal{M}_z$, where $\tau = 50$, $n = 4$.

Another four-dimensional Lorenz-like hyperchaotic dynamic system is given by [51]:

$$\begin{aligned}
\frac{dx}{dt} &= 10(y - x) \\
\frac{dy}{dt} &= 28x - y - xz + w \\
\frac{dz}{dt} &= -\frac{8}{3}z + xy \\
\frac{dw}{dt} &= -k_1x - k_2y
\end{aligned} \tag{5.2}$$

where $k_1 = -9.3$, and $k_2 = -5$. This dynamic system is constructed by adding a linear controller to the second equation of the Lorenz63 system. This dynamic system is rotational symmetric under the order-two group $\mathbb{C}_2 = \{e, R_z\}$, where $R_z : (x, y, z, w) \Rightarrow (-x, -y, z, w)$. We simulate this dynamic system in the temporal domain $t = [0, 50]$ with $\Delta t = 0.005$, and parameters for the embedding are $\tau = 15$, $n = 4$.

Similarly, all shadow manifolds except $\mathcal{M}_z$ are symmetric. Consequently, CCM incorrectly identifies bidirectional causal links $X \Leftrightarrow Z$, $Y \Leftrightarrow Z$, and $W \Leftrightarrow Z$ as unidirectional. True causal links are correctly detected by applying sCCM, and results are shown in Table 4.

The last case is a five-dimensional hyperchaotic system. Similarly, this system is obtained by adding a nonlinear quadratic controller to the first equation and a linear controller to the second equation of the modified Lorenz63 system. Moreover, this system can be realized via an electronic system [52]. It is mathematical formulation is:

$$\begin{aligned}\frac{dx}{dt} &= 10(y - x) + u \\ \frac{dy}{dt} &= 28x - y - xz - v \\ \frac{dz}{dt} &= -\frac{8}{3}z + xy \\ \frac{du}{dt} &= -xz + k_1u \\ \frac{dv}{dt} &= k_2y\end{aligned}\tag{5.3}$$

where $k_1 = 1$ and $k_2 = 30$. This system is simulated in the temporal domain $t = [0, 100]$ with $\Delta t = 0.01$, and embedding parameters are $\tau = 60$, $n = 5$. This hyperchaotic system is rotation symmetric under the order-two group $\mathbb{C}_2 = \{e, R_z(\pi)\}$, where $R_z : (x, y, z, u, w) \Rightarrow (-x, -y, z, -u, -w)$ as shown in Fig. 11, which exhibits a genus-1 attractor in the (v, z) plane.

Similarly, the nonsymmetric shadow manifold $\mathcal{M}_z$ causes CCM to mislead bidirectional causal links $X \Leftrightarrow Z$, $Y \Leftrightarrow Z$, and $W \Leftrightarrow Z$ as $Z \Rightarrow X$, $Z \Rightarrow Y$, and $Z \Rightarrow W$. However, true causal links are correctly identified using our framework. Results are shown in Table 4.

Equations	CCM	Segment CCM
4D dissipative system (5.1)	$Z \Rightarrow X$ ($\rho_{xz} = 0.547, \rho_{zx} = 0.994$)	$X \Leftrightarrow Z$ ($\rho_{xz} = 0.967, \rho_{zx} = 0.973$)
	$Z \Rightarrow Y$ ($\rho_{yz} = 0.474, \rho_{zy} = 0.990$)	$Y \Leftrightarrow Z$ ($\rho_{yz} = 0.971, \rho_{zy} = 0.931$)
	$Z \Rightarrow W$ ($\rho_{wz} = 0.273, \rho_{zw} = 0.978$)	$W \Leftrightarrow Z$ ($\rho_{wz} = 0.950, \rho_{zw} = 0.904$)
4D Lorenz-like system (5.2)	$Z \Rightarrow X$ ($\rho_{xz} = 0.597, \rho_{zx} = 0.988$)	$X \Leftrightarrow Z$ ($\rho_{xz} = 0.973, \rho_{zx} = 0.964$)
	$Z \Rightarrow Y$ ($\rho_{yz} = 0.583, \rho_{zy} = 0.962$)	$Y \Leftrightarrow Z$ ($\rho_{yz} = 0.966, \rho_{zy} = 0.905$)
	$Z \Rightarrow W$ ($\rho_{wz} = 0.609, \rho_{zw} = 0.954$)	$W \Leftrightarrow Z$ ($\rho_{wz} = 0.912, \rho_{zw} = 0.861$)
5D Lorenz-like system (5.3)	$Z \Rightarrow X$ ($\rho_{xz} = 0.151, \rho_{zx} = 0.999$)	$X \Leftrightarrow Z$ ($\rho_{xz} = 0.994, \rho_{zx} = 0.999$)
	$Z \Rightarrow Y$ ($\rho_{yz} = 0.194, \rho_{zy} = 0.999$)	$Y \Leftrightarrow Z$ ($\rho_{yz} = 0.985, \rho_{zy} = 0.986$)
	$Z \Rightarrow U$ ($\rho_{uz} = 0.098, \rho_{zu} = 0.999$)	$U \Leftrightarrow Z$ ($\rho_{uz} = 0.991, \rho_{zu} = 0.999$)
	$Z \Rightarrow W$ ($\rho_{wz} = 0.105, \rho_{zw} = 0.999$)	$W \Leftrightarrow Z$ ($\rho_{wz} = 0.993, \rho_{zw} = 0.999$)

Table 4: Performance of Segment CCM for high dimensional rotational symmetric systems under $\mathbb{C}_2$.

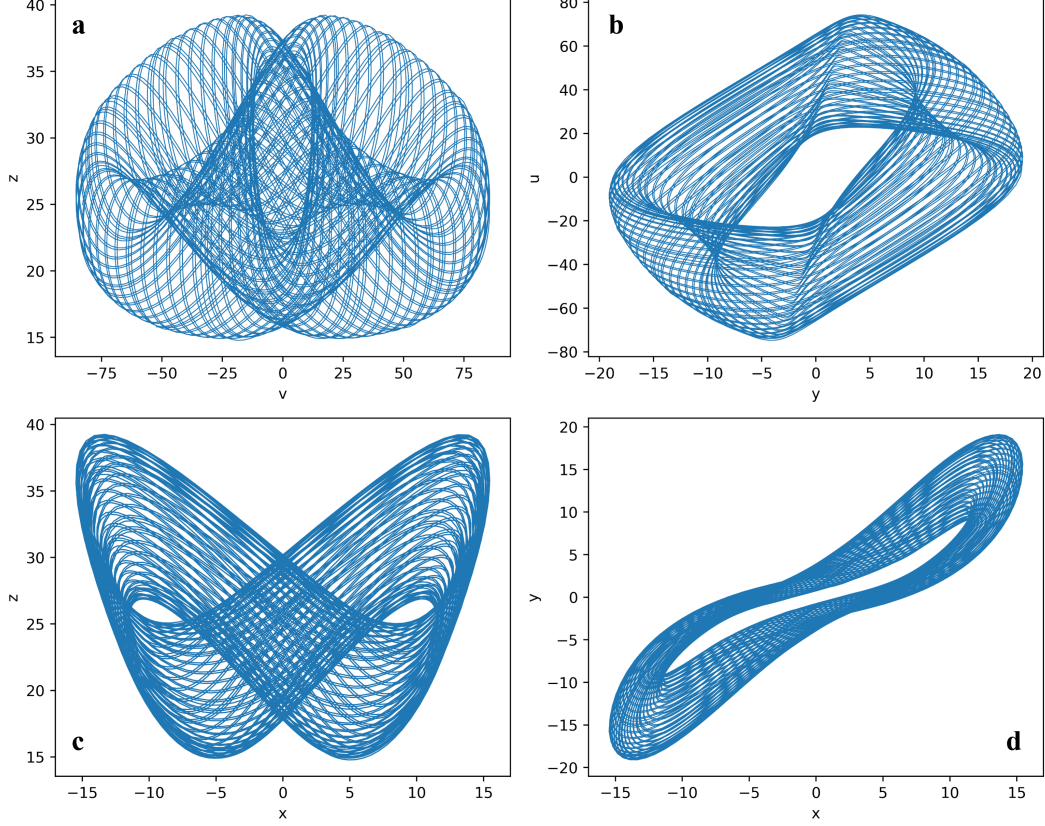


Figure 11: Projections of the attractor generated by equation (5.3) a. (v,z) plane projection. b. (y,u) plane projection. c. (x,z) plane projection. d. (x,y) plane projection.

5.3. Application in Real-world system

To evaluate our the performance of our framework with respect to a real-world system, we built a nonlinear chaotic circuit for which implementation requires two multipliers with multiple op-ams or using two multipliers with a couple of resistors and capacitors. The signals recorded from these two channels are referred to as X and Z . As an illustration, the following symmetric chaotic system can be realized using the AD633 multiplier and resistor-capacitor coupling for current control, which is provided in Fig . 12, where the set parameters are as follows: $C_1 = C_2 = 1nF$, $C_3 = 10nF$, $R_1 = 10k\Omega$, $R_2 = 1k\Omega$, $R_3 = 200\Omega$, $R_4 = 2k\Omega$, $V_{cc} = 4V$ [53]. The input voltage V_{cc} powers the circuit, ensuring the proper operation of the system. The initial conditions of the system are typically set by the initial voltages across the capacitors and the configuration of resistors in the system. We chose four pairs of initial conditions for X , Y and Z . In the experimental setup, there is a bidirectional causal link between X and Z . We run our sCCM method and CCM for five times and the results are plotted in Fig . 13. Interestingly, our algorithm detect the correct causal direction, identifying a bidirectional causal link between X and Z , whereas CCM incorrectly inferred unidirectional coupling.

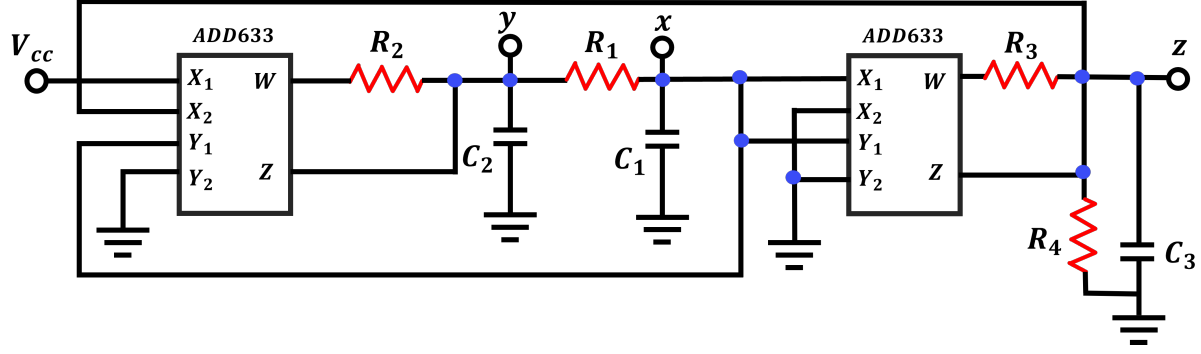


Figure 12: Circuit diagram used for generating the real-world data.

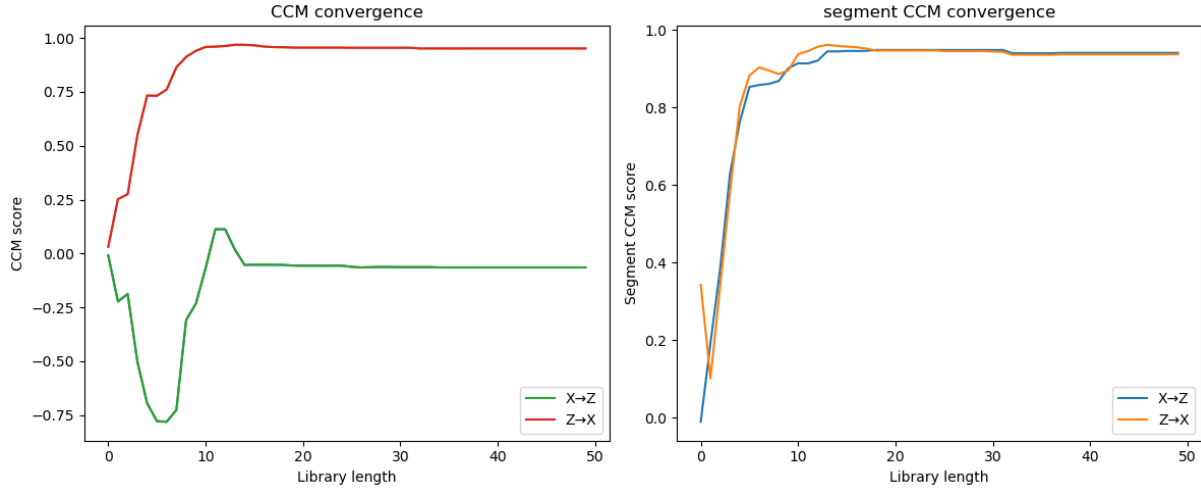


Figure 13: performance of Segment CCM on the circuit case.

5.4. Further discussion: High dimensional four-fold rotational Laser system

For symmetric attractors with n -fold symmetry ($n > 2$), our method can still correct the misleading results produced by CCM due to the variable that defines even parity. We use the five dimensional Laser system [54], which take into account the detuning parameter δ between the normalized steady state laser frequency and the molecular resonance frequency. This is a five dimensional, four-fold rotation symmetric system, and its mathematical expression is:

$$\begin{aligned}
 \frac{dx_1}{dt} &= -\sigma(x_1 + \delta x_2 - y_1) \\
 \frac{dx_2}{dt} &= -\sigma(x_2 - \delta x_1 - y_2) \\
 \frac{dy_1}{dt} &= Rx_1 - y_1 + \delta y_2 - x_1 z \\
 \frac{dy_2}{dt} &= Rx_2 - \delta y_1 - y_2 - x_2 z \\
 \frac{dz}{dt} &= -\gamma z + x_1 y_1 + x_2 y_2
 \end{aligned} \tag{5.4}$$

where R , σ , and γ have the same physical meaning. This dynamic system is rotation symmetric under the following matrix of the representation:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}. \quad (5.5)$$

Here, we simulate this dynamic system with parameters $\sigma = 2$, $\delta = 0.1$, $R = 24$, $\gamma = 0.25$ in the temporal domain $t = [0, 250]$, $\Delta t = 0.025$, initial point $\mathbf{x}_0 = [4, 0, 1, 0, 30]$, and embedding parameters are $\tau = 30$ and $n = 5$. We both implement CCM and our framework on this system and the results are shown in the following Table 4. We find that our framework detect the correct causal relationship $Y_2 \Leftrightarrow Z$, whereas the CCM shows the misleading bidirectional causal inference.

CCM	sCCM
$Z \Rightarrow Y_2$ ($\rho_{y_2 z} = 0.133$, $\rho_{zy_2} = 0.956$)	$Z \Leftrightarrow Y_2$ ($\rho_{y_2 z} = 0.825$, $\rho_{zy_2} = 0.944$)

Table 5: Performance of Segment CCM.

6. Conclusion

This paper explores the complexities of detecting causal relations in nonlinear dynamic systems using CCM, explicitly focusing on the addressing the challenges when the measurement function defines the even parity, which may cause the misleading result caused by CCM. The result reveals that shadow manifolds reconstructed by the differential/delay-coordinate mapping from the measurement function which defines even parity, lacking symmetry information, can potentially mislead the causal relation detection in CCM since the delay-coordinate mapping is no longer a one-to-one mapping. We proposed a new framework based on the k-means clustering method, partitioning the symmetric shadow manifolds into two sub-manifolds, thus transforming the two-to-one mapping into two one-to-one for solving this problem. We evaluate the efficiency of our method on various numerical simulation cases, as well as a real-world application involving a nonlinear circuit, to verify its accuracy.

Several areas warrant further exploration. First, the result of CCM is highly dependent on the performance of the delay-coordinate mapping. As we discussed, for high-dimensional dynamic systems, one-dimensional observations generally provide poor observability of the original attractor, necessitating the combination of observations from multiple variables for sufficient embedding. However, this approach may introduce confounding information, complicating the detection of causal relationships. Moreover, the statistical response of the algorithm under the perturbation of additive noise would be our future research.

Acknowledgments

The author acknowledge that no external funding or support was received for this research.

References

- [1] C. W. Granger, Investigating causal relations by econometric models and cross-spectral methods, *Econometrica: journal of the Econometric Society* (1969) 424–438.
- [2] F. A. Kitole, L. J. Msoma, J. K. Sesabo, Navigating the economic landscape: a comprehensive analysis of government spending, economic growth, and poverty reduction nexus in tanzania, *Applied Economics Letters* (2024) 1–5.
- [3] L. N. Ross, D. S. Bassett, Causation in neuroscience: Keeping mechanism meaningful, *Nature Reviews Neuroscience* 25 (2) (2024) 81–90.
- [4] G. Sugihara, R. May, H. Ye, C.-h. Hsieh, E. Deyle, M. Fogarty, S. Munch, Detecting causality in complex ecosystems, *science* 338 (6106) (2012) 496–500.
- [5] Y. Qing, S. Wang, Z.-L. Yang, P. Gentine, Soil moisture- atmosphere feedbacks have triggered the shifts from drought to pluvial conditions since 1980, *Communications Earth & Environment* 4 (1) (2023) 254.
- [6] S. Avvaru, K. K. Parhi, Effective brain connectivity extraction by frequency-domain convergent cross-mapping (fdccm) and its application in parkinson’s disease classification, *IEEE Transactions on Biomedical Engineering* 70 (8) (2023) 2475–2485.
- [7] B. Gao, J. Yang, Z. Chen, G. Sugihara, M. Li, A. Stein, M.-P. Kwan, J. Wang, Causal inference from cross-sectional earth system data with geographical convergent cross mapping, *nature communications* 14 (1) (2023) 5875.
- [8] A. E. Yuan, W. Shou, Data-driven causal analysis of observational time series: A synthesis, *BioRxiv* (2020) 2020–08.
- [9] K. Butler, G. Feng, P. M. Djurić, On causal discovery with convergent cross mapping, *IEEE Transactions on Signal Processing* (2023).
- [10] S. H. Strogatz, *Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering*, CRC press, 2018.
- [11] J. Milnor, On the concept of attractor, *Communications in Mathematical Physics* 99 (1985) 177–195.

- [12] F. Takens, Detecting strange attractors in turbulence, in: *Dynamical Systems and Turbulence, Warwick 1980: proceedings of a symposium held at the University of Warwick 1979/80*, Springer, 2006, pp. 366–381.
- [13] H. Whitney, The self-intersections of a smooth n -manifold in $2n$ -space, *Annals of Mathematics* 45 (2) (1944) 220–246.
- [14] T. D. Tsankov, A. Nishtala, R. Gilmore, Embeddings of a strange attractor into $\mathbb{R}^3$, *Physical Review E* 69 (5) (2004) 056215.
- [15] J. Nichols, J. Nichols, Attractor reconstruction for non-linear systems: a methodological note, *Mathematical Biosciences* 171 (1) (2001) 21–32.
- [16] E. Tan, S. Algar, D. Corrêa, M. Small, T. Stemler, D. Walker, Selecting embedding delays: An overview of embedding techniques and a new method using persistent homology, *Chaos: An Interdisciplinary Journal of Nonlinear Science* 33 (3) (2023).
- [17] R. Martin, C. Greve, C. Huerta, A. Wong, J. Koo, D. Eckhardt, A robust time-delay selection criterion applied to convergent cross mapping, *Chaos: An Interdisciplinary Journal of Nonlinear Science* 34 (9) (2024).
- [18] H. Kim, R. Eykholt, J. Salas, Nonlinear dynamics, delay times, and embedding windows, *Physica D: Nonlinear Phenomena* 127 (1-2) (1999) 48–60.
- [19] C. Rhodes, M. Morari, False-nearest-neighbors algorithm and noise-corrupted time series, *Physical Review E* 55 (5) (1997) 6162.
- [20] B. Cummins, T. Gedeon, K. Spendlove, On the efficacy of state space reconstruction methods in determining causality, *SIAM Journal on Applied Dynamical Systems* 14 (1) (2015) 335–381.
- [21] H. Ye, E. R. Deyle, L. J. Gilarranz, G. Sugihara, Distinguishing time-delayed causal interactions using convergent cross mapping, *Scientific reports* 5 (1) (2015) 14750.
- [22] E. De Brouwer, A. Arany, J. Simm, Y. Moreau, Latent convergent cross mapping, in: *International Conference on Learning Representations*, 2020.
- [23] A. T. Clark, H. Ye, F. Isbell, E. R. Deyle, J. Cowles, G. D. Tilman, G. Sugihara, Spatial convergent cross mapping to detect causal relationships from short time series, *Ecology* 96 (5) (2015) 1174–1181.
- [24] G. Feng, J. G. Quirk, P. M. Djurić, Detecting causality using deep gaussian processes, in: *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, IEEE, 2019, pp. 472–476.
- [25] H. Ma, K. Aihara, L. Chen, Detecting causality from nonlinear dynamics with short-term time series, *Scientific reports* 4 (1) (2014) 7464.

- [26] S. Cobey, E. B. Baskerville, Limits to causal inference with state-space reconstruction for infectious disease, *PloS one* 11 (12) (2016) e0169050.
- [27] A. Krakovská, J. Jakubík, Implementation of two causal methods based on predictions in reconstructed state spaces, *Physical Review E* 102 (2) (2020) 022203.
- [28] S. Leng, H. Ma, J. Kurths, Y.-C. Lai, W. Lin, K. Aihara, L. Chen, Partial cross mapping eliminates indirect causal influences, *Nature communications* 11 (1) (2020) 2632.
- [29] A. Ghouse, L. Faes, G. Valenza, Inferring directionality of coupled dynamical systems using gaussian process priors: Application on neurovascular systems, *Physical Review E* 104 (6) (2021) 064208.
- [30] G. Chen, T. Ueta, Yet another chaotic attractor, *International Journal of Bifurcation and chaos* 9 (07) (1999) 1465–1466.
- [31] R. Shaw, Strange attractors, chaotic behavior, and information flow, *Zeitschrift für Naturforschung A* 36 (1) (1981) 80–112.
- [32] D. Li, A three-scroll chaotic attractor, *Physics Letters A* 372 (4) (2008) 387–393.
- [33] C. Letellier, E. M. Mendes, J.-M. Malasoma, Lorenz-like systems and lorenz-like attractors: Definition, examples, and equivalences, *Physical Review E* 108 (4) (2023) 044209.
- [34] C. Letellier, P. Dutertre, J. Reizner, G. Gouesbet, Evolution of a multimodal map induced by an equivariant vector field, *Journal of Physics A: Mathematical and General* 29 (17) (1996) 5359.
- [35] P. Ashwin, I. Melbourne, Symmetry groups of attractors, *Archive for rational mechanics and analysis* 126 (1994) 59–78.
- [36] M. Field, I. Melbourne, M. Nicol, Symmetric attractors for diffeomorphisms and flows, *Proceedings of the London Mathematical Society* 3 (3) (1996) 657–696.
- [37] J. C. Sprott, Simplest chaotic flows with involutorial symmetries, *International Journal of Bifurcation and Chaos* 24 (01) (2014) 1450009.
- [38] D. J. Cross, R. Gilmore, Equivariant differential embeddings, *Journal of mathematical physics* 51 (9) (2010).
- [39] G. Marghoti, T. d. L. Prado, S. R. Lopes, Y. Hirata, Involution symmetry quantification using recurrences, *Physical Review E* 110 (2) (2024) 024203.
- [40] I. Ali, A. U. Rehman, D. M. Khan, Z. Khan, M. Shafiq, J.-G. Choi, Model selection using k-means clustering algorithm for the symmetrical segmentation of remote sensing datasets, *Symmetry* 14 (6) (2022) 1149.

- [41] C. E. Gonzalez, C. Lainscsek, T. J. Sejnowski, C. Letellier, Assessing observability of chaotic systems using delay differential analysis, *Chaos: An Interdisciplinary Journal of Nonlinear Science* 30 (10) (2020).
- [42] Y. Wang, J. Singer, H. H. Bau, Controlling chaos in a thermal convection loop, *Journal of Fluid Mechanics* 237 (1992) 479–498.
- [43] T. Shimizu, N. Morioka, On the bifurcation of a symmetric limit cycle to an asymmetric one in a simple model, *Physics Letters A* 76 (3-4) (1980) 201–204.
- [44] A. M. Rucklidge, Chaos in models of double convection, *Journal of Fluid Mechanics* 237 (1992) 209–229.
- [45] J. C. Sprott, Some simple chaotic flows, *Physical review E* 50 (2) (1994) R647.
- [46] T. Rikitake, Oscillations of a system of disk dynamos, in: *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 54, Cambridge University Press, 1958, pp. 89–105.
- [47] J. Lü, G. Chen, D. Cheng, A new chaotic system and beyond: the generalized lorenz-like system, *International Journal of Bifurcation and Chaos* 14 (05) (2004) 1507–1537.
- [48] L. Chongxin, L. Ling, L. Peng, et al., A new butterfly-shaped attractor of lorenz-like system, *Chaos, solitons & fractals* 28 (5) (2006) 1196–1203.
- [49] J. Lü, G. Chen, A new chaotic attractor coined, *International Journal of Bifurcation and chaos* 12 (03) (2002) 659–661.
- [50] L. Huang, Z. Zhang, J. Xiang, S. Wang, A new 4d chaotic system with two-wing, four-wing, and coexisting attractors and its circuit simulation, *Complexity* 2019 (2019) 1–13.
- [51] Q. Yang, K. Zhang, G. Chen, Hyperchaotic attractors from a linearly controlled lorenz system, *Nonlinear Analysis: Real World Applications* 10 (3) (2009) 1601–1617.
- [52] G. Hu, Generating hyperchaotic attractors with three positive lyapunov exponents via state feedback control, *International Journal of Bifurcation and Chaos* 19 (02) (2009) 651–660.
- [53] J. Wu, C. Li, X. Ma, T. Lei, G. Chen, Simplification of chaotic circuits with quadratic nonlinearity, *IEEE Transactions on Circuits and Systems II: Express Briefs* 69 (3) (2021) 1837–1841.
- [54] H. Zeghlache, P. Mandel, Influence of detuning on the properties of laser equations, *JOSA B* 2 (1) (1985) 18–22.
- [55] P. Reiterer, C. Lainscsek, F. Schürer, C. Letellier, J. Maquet, A nine-dimensional lorenz system to study high-dimensional chaos, *Journal of Physics A: Mathematical and General* 31 (34) (1998) 7121.

- [56] R. Hermann, A. Krener, Nonlinear controllability and observability, IEEE Transactions on automatic control 22 (5) (1977) 728–740.
- [57] C. Letellier, I. Sendiña-Nadal, E. Bianco-Martinez, M. S. Baptista, A symbolic network-based nonlinear theory for dynamical systems observability, Scientific reports 8 (1) (2018) 3785.
- [58] A. Katok, A. Katok, B. Hasselblatt, Introduction to the modern theory of dynamical systems, no. 54, Cambridge university press, 1995.
- [59] N. Marwan, M. C. Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems, Physics reports 438 (5-6) (2007) 237–329.
- [60] S. Bartsev, M. Saltykov, P. Belolipetsky, A. Pianykh, Imperfection of the convergent cross-mapping method, in: IOP Conference Series: Materials Science and Engineering, Vol. 1047, IOP Publishing, 2021, p. 012081.

Appendix A. Observability of nonlinear dynamic system

According to Taken’s embedding theorem, using a delay-coordinate mapping, an n -dimensional dynamical system can be reconstructed from a one-dimensional observation. This reconstruction process is akin to splitting n -dimensional linearly independent information. However, the effectiveness of this approach diminishes as the dimension of the dynamical system increases. To illustrate these limitations, we consider the nine-dimensional Lorenz system [55] as an example, demonstrating the inefficient performance of the delay-coordinate mapping in high-dimensional settings.

The mathematical formulation of the nine-dimensional Lorenz system is given by:

$$\begin{aligned}
\dot{x}_1 &= -\sigma(b_1x_1 + b_2x_7) + x_4(b_4x_4 - x_2) + b_3x_3x_5, \\
\dot{x}_2 &= -\sigma x_2 + x_1x_4 - x_2x_5 + x_4x_5 - \frac{2}{\sigma}x_9, \\
\dot{x}_3 &= \sigma(b_2x_8 - b_1x_3) + x_2x_4 - b_4x_2^2 - \frac{b_3x_1x_5}{\sigma}, \\
\dot{x}_4 &= -\sigma x_4 - x_2x_3 - x_2x_5 + x_4x_5 + \frac{x_9}{2}, \\
\dot{x}_5 &= -\sigma b_5x_5 + \frac{x_2^2}{2} - \frac{x_4^2}{2}, \\
\dot{x}_6 &= -b_6x_6 + x_2x_9 - x_4x_9, \\
\dot{x}_7 &= -b_1x_7 - Rx_1 + 2x_5x_8 - x_4x_9, \\
\dot{x}_8 &= -b_1x_8 + Rx_3 - 2x_5x_7 + x_2x_9, \\
\dot{x}_9 &= -x_9 + (R + 2x_6)(x_4 - x_2) + x_4x_7 - x_2x_8,
\end{aligned} \tag{A.1}$$

where $b_1 = \frac{5}{1.5}$, $b_2 = 0.6$, $b_3 = 1.2$, $b_4 = 0.2$, $b_5 = \frac{2}{1.5}$, $b_6 = \frac{4}{1.5}$, and $R = 14.3$. The initial state is $\mathbf{x}_0 = [0.01, 0, 0.01, 0, 0, 0, 0, 0, 0.01]$, with a time step of $\Delta t = 0.02$ and the simulation time domain of

$[0, 800]$. The measurement function is $h = x_9$, with the lag value $\tau = 12$ and an embedding dimension $n = 9$. The original attractor and the reconstructed shadow manifold are shown in Fig. 14. It seems that

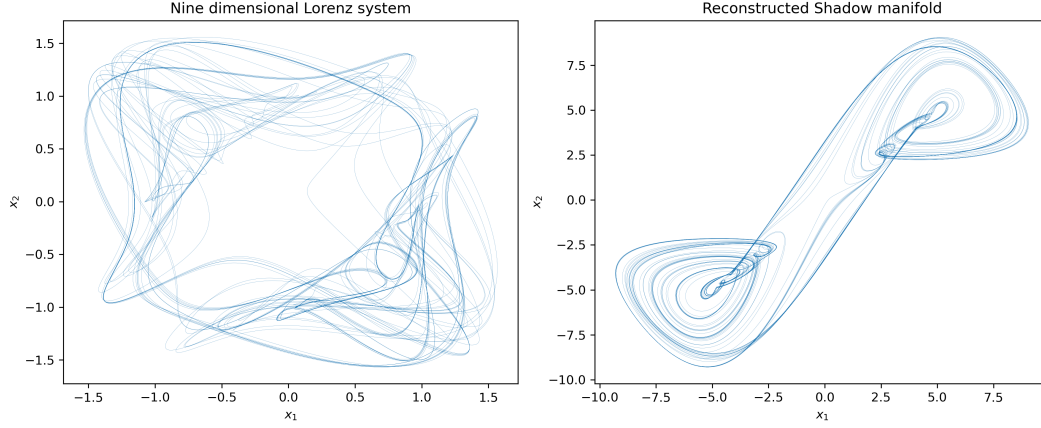


Figure 14: **Left: Chaotic attractor produced by nine-dimensional Lorenz system. Right: Reconstructed shadow manifold by delay-coordinate mapping.**

the reconstructed shadow manifold exhibits a butterfly attractor similar to the Lorenz63 system rather than reflecting the more complex dynamic behavior in the nine-dimensional attractor. This discrepancy can be attributed to the observability of the system. Observability refers to the ability to reconstruct the entire state of the system using a subset of measurable variables over a finite period. Consider the following nonlinear system:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\ s(t) &= h(\mathbf{x})\end{aligned}\tag{A.2}$$

where $\mathbf{x} \in \mathbb{R}^m$ represents the m -dimensional state vector, $h : \mathbb{R}^m \rightarrow \mathbb{R}$ is the measurement function, $s(t) \in \mathbb{R}$ is the measurement, and $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ denotes the vector field. Differentiating $s(t)$ yields

$$\dot{s}(t) = \frac{d}{dt}h(\mathbf{x}) = \frac{\partial h}{\partial \mathbf{x}} \dot{\mathbf{x}} = \frac{\partial h}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}) = \mathcal{L}_f h(\mathbf{x})\tag{A.3}$$

where $\mathcal{L}_f h(\mathbf{x})$ is the Lie derivative of h along the vector field $\mathbf{f}$ and the time derivative of s in Lie derivatives form can be written as $s^j = \mathcal{L}_f^j h(\mathbf{x})$. The j -th order Lie derivative is given by

$$\mathcal{L}_f^j h(\mathbf{x}) = \frac{\partial \mathcal{L}_f^{j-1} h(\mathbf{x})}{\partial \mathbf{x}} \cdot \mathbf{f}(\mathbf{x})\tag{A.4}$$

where $\mathcal{L}_f^0 h(\mathbf{x}) = h(\mathbf{x})$.

A dynamic system is said to be *state observable* at time t_f if its initial state $\mathbf{x}_0$ can be uniquely determined from the measurement vector $s(t)$ over the interval $0 \leq t \leq t_f$. The observability matrix, the Jacobian matrix of the Lie derivatives of $h(\mathbf{x})$, determines whether a dynamic system is observable. The

observability matrix $\mathcal{O} \in \mathbb{R}^{m \times m}$ is expressed as:

$$\mathcal{O}(\mathbf{x}) = \begin{bmatrix} \frac{\partial \mathcal{L}_f^0 h(\mathbf{x})}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial \mathcal{L}_f^{m-1} h(\mathbf{x})}{\partial \mathbf{x}} \end{bmatrix}. \quad (\text{A.5})$$

For the linear system,

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} + B\mathbf{u} \\ \mathbf{s} &= C\mathbf{x}, \end{aligned} \quad (\text{A.6})$$

where A, B, C are constant dynamics matrices, the observability matrix simplifies to:

$$\mathcal{O} = \begin{bmatrix} C & CA & CA^2 & \dots & CA^{n-1} \end{bmatrix}^T. \quad (\text{A.7})$$

Theorem 1. The m -dimensional dynamic system is said to be *state observable* if and only if the observability matrix has full rank, that is, $\text{rank}(\mathcal{O}) = m$.

The above theorem is a generalization of the observability property of both linear and nonlinear systems and has been proved in [56]. It states that if a system is observable for both linear and nonlinear dynamic systems, it is possible to recover every initial condition from the measured series $s(t)$ for $t \leq 0$. In the case of the nine-dimensional Lorenz system, the system is nearly unobservable from a single observation, as the observability matrix is deficient. For this system, at least six variables should be measured to achieve sufficient observability of the high dimensional dynamics [57].

Thus, for high-dimensional dynamical systems, a single variable is generally insufficient to construct a diffeomorphic shadow manifold, implying that CCM results may be unreliable.

Moreover, a single variable may still provide poor observability even for low-dimensional dynamic systems ($n \leq 3$). For instance, consider the Rössler system, the mathematical formulation is:

$$\begin{aligned} \dot{x}_1 &= -y - z, \\ \dot{x}_2 &= x + ay, \\ \dot{x}_3 &= b + xz - cz, \end{aligned} \quad (\text{A.8})$$

where $a = 0.2$, $b = 0.2$, $c = 5.7$. The time step is set to $\Delta t = 0.01$, with a total of 400000 iterations, starting from the initial condition $\mathbf{x}_0 = [1, 1, 0]$. The original attractor and the reconstructed shadow manifold are shown in Fig. 15. Here, we modify the measurement function to $h' = x + z$ and obtain the shadow manifold $\mathcal{M}_{z'}$. The lag value is $\tau = 40$ for both shadow manifolds, and the embedding dimension is $n = 3$. It can be observed that the single observation $h = z$ yields poor observability, such that the observability matrix $\mathcal{O}(\mathbf{x})$ is deficient near the original point, leading to significant distortion in the shadow manifold $\mathcal{M}_z$. In contrast, the shadow manifold reconstructed from the multivariate measurement function h' provides a better result.

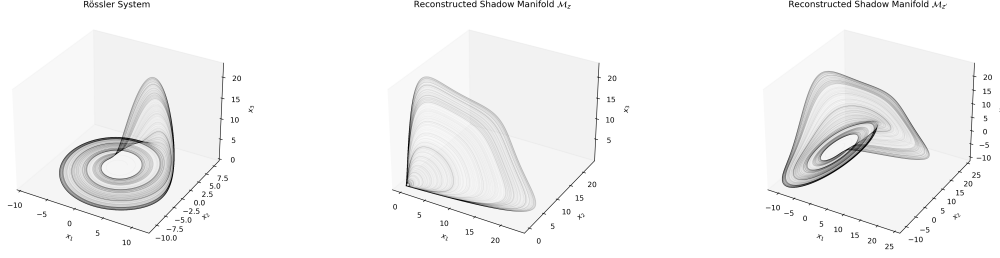


Figure 15: **Left:** Chaotic attractor produced by the Rössler system. **Middle:** Reconstructed shadow manifold $\mathcal{M}_z$. **Right:** Reconstructed shadow manifold $\mathcal{M}_{z'}$.

Appendix B. The Importance of Recurrence Property: Explanations for a case

Here, we emphasize that the recurrence property is necessary for CCM. The recurrence property is one of the fundamental properties of the dynamic systems. As stated in the Poincaré recurrence theorem [58], in any measure-preserving transformation on a dynamic system's attractor, the trajectories will eventually reappear at the neighborhood of the former points in the state space. The recurrences can be captured through distance plots [59]. For a series of trajectories $\{\mathbf{x}_i\}_{i=1}^N$ of a dynamic system in its phase space, the corresponding distance plot is based on the matrix:

$$R_{i,j} = \|\mathbf{x}_i - \mathbf{x}_j\|, \quad i, j = 1, 2, \dots, N \quad (\text{B.1})$$

where N is the number of considered states $\mathbf{x}_i$ and $\|\cdot\|$ represents a norm (e.g., the Euclidean norm). An example of Lorenz63 system is presented in Fig. 16, it delineates the distances between every pair of

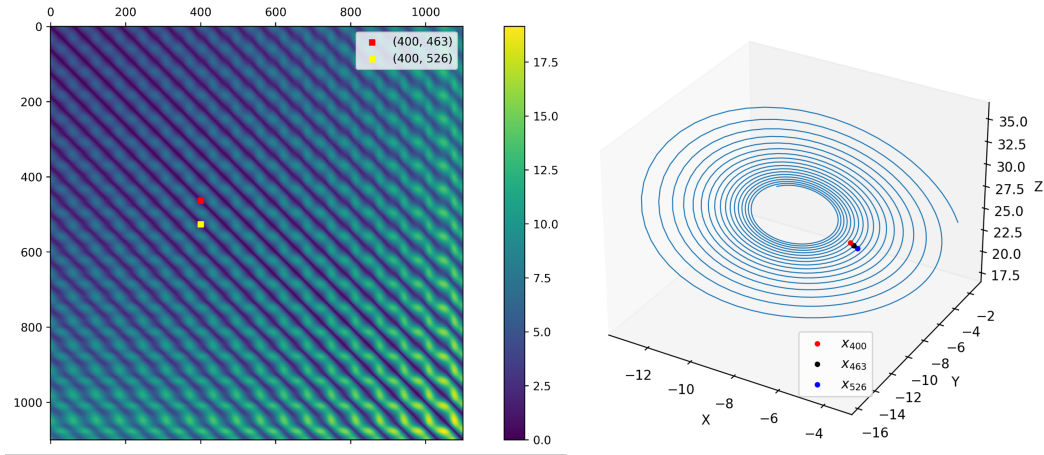


Figure 16: Distance plot using the Euclidean distance and corresponding trajectories of Lorenz63 attractor. This system is simulated as same as Fig. 3. **Left:** Distance plot for trajectories lies in $t \in [1, 12]$. **Right:** Trajectories for the Lorenz63 at $t \in [1, 12]$.

points $\mathbf{x}_i$ and $\mathbf{x}_j$ in the state space. The points $\mathbf{x}_{400}$ and $\mathbf{x}_{526}$ would revisit the position within $\mathbf{x}_{400}$.

Consider the counterexample mentioned in [60]. This paper discusses that when the time series have certain forms – one of the series is a periodic function of time, and the second series has a time trend

- it may mislead the results of CCM because points from the manifold with a periodic shape are close to points from the manifold with a linear shape. However, we point out that the second monotonic increasing time series does not satisfy the recurrence requirement for CCM, which means the shadow manifold for this monotone time series is not an attractor.

One typical example is:

$$\begin{aligned} x(t) &= \frac{1}{2000}t \\ y(t) &= \sin\left(\frac{\pi}{50}t\right) \end{aligned} \tag{B.2}$$

Here t refers to the time. We generate 2000 examples containing 20 complete periods about $y(t)$, and the results of causal links are shown in the following Figure 17. The monotone variable causes the periodic

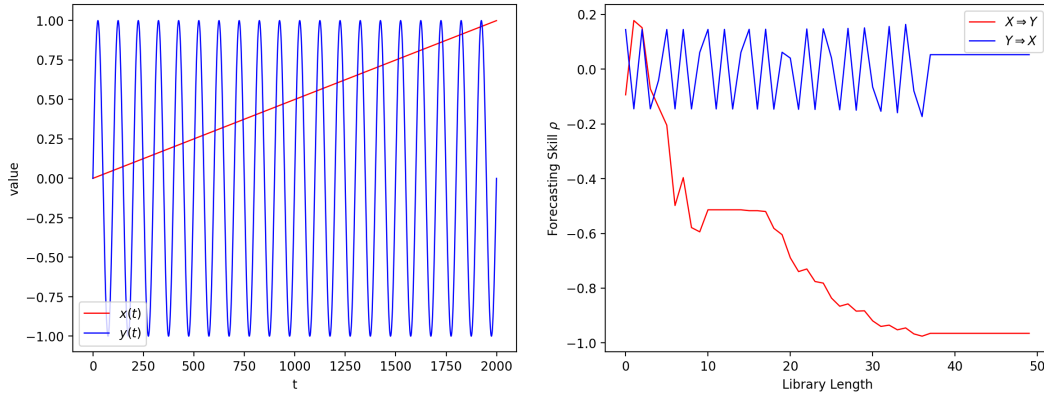


Figure 17: Left: Trajectories of Eq. (B.2). Right: Causal links obtained by CCM.

one as the Pearson correlation coefficient ρ_{xy} converges to 0.965. In [60], it is stated that a monotone time series has no causal effect on the periodic time series. However, the mathematical expression of $y(t)$ can be simply solvable by $x(t)$ as we wrote it into the discrete form:

$$\begin{aligned} x(t+1) &= x(t) + \frac{1}{2000} \\ y(t+1) &= \sin(40\pi[x(t+1) - \frac{1}{2000}] + \frac{1}{50}) \end{aligned} \tag{B.3}$$

which means that the monotone variable causes the other one and corresponds to what we see. However, the inverse relationship also holds:

$$\begin{aligned} y(t+1) &= y(t) \cos\left(\frac{\pi}{50}\right) + \cos\left(\frac{\pi}{50}t\right) \sin\left(\frac{\pi}{50}\right) \\ x(t+1) &= \frac{\arcsin(y(t+1)) - \frac{1}{50}}{40\pi} + \frac{1}{2000} \end{aligned} \tag{B.4}$$

From the above analysis, if the nonlinear periodic term can be explicitly solved, as in the case above, it can be expressed in terms of the linear term and vice versa. However, it is important to emphasize that the CCM test becomes nonsensical under these conditions. It does not conform to the requirement of CCM; precisely, with the library length increase, the monotone shadow manifold reconstructed from the linear term would not become denser with the nonlinear one. In this scenario, the prediction ability does not improve as the library length increases, making the CCM result unreliable.