Robust methods to detect coupling among nonlinear time series

Timothy Sauer
George Mason University
Fairfax, VA 22030

George Sugihara
Scripps Institution of Oceanography at UCSD
La Jolla, CA 92037
(Dated: December 17, 2024)

Two numerical methods are proposed for detection of coupling between multiple time series generated by deterministic nonlinear systems. The first detects interdepence or the existence of coupling between time series. The second ascertains directionality of coupling, or alternatively, latent coupling, the case when multiple series are driven by another, unobserved system. In either case, the driver and the recipients of the coupling may be periodic or aperiodic, and in particular may be chaotic. The only inputs to the method are two or more simultaneously recorded time series. The methods rely solely on ranking distances between states in time-delay reconstructions of the data, and for that reason tend to be robust to observational noise.

I. INTRODUCTION

Determining causal relationships is a fundamental challenge across scientific and engineering disciplines. For nonlinear deterministic systems, inferring causality from observational time series data is particularly complex and often counterintuitive. This happens when the assumption of separability of causes breaks down due to interdependent coupling, or where nonlinear dynamics allows causation to occur in the absence of correlation. Nevertheless, significant progress has been made in leveraging non-separability to uncover instances where the dynamics of one system may influence or drive another. In this article, we introduce a robust method to address this problem, utilizing order statistics derived from relative distances within a reconstructed phase space.

Given two nonlinear systems X and Y, and their time series observations $\{x_t\}$ and $\{y_t\}$, we develop criteria to decide which of the following relations hold:

1. **Unidirectional Coupling:** X drives Y $(X \to Y)$ or Y drives X $(Y \to X)$.

2. **Bidirectional Coupling/GS:** Both X and Y drive each other $(X \leftrightarrow Y)$, or the systems are in generalized synchrony (GS) caused by strong unidirectional coupling between X and Y.

3. **Latent Coupling:** X and Y do not influence each other but appear to be coupled because they are driven by a third, unobserved dynamical system.

4. **Independence:** X and Y are dynamically independent of one another.

Our methods rely on *genericity* of the dynamics in that counterexamples to these criteria may be artificially constructed, but are not expected to be common in general circumstances.

Generalized synchrony [1] (GS) is the circumstance when each state of X occurs simultaneously with a unique corresponding state of Y, and vice versa. This is distinct from the special case of *identical synchrony* [2] where $X = Y$ and the corresponding states are identical. Generalized synchrony is often observed as a result of unidirectional driving that is relatively strong, although (perhaps surprisingly) it can also occur for weak unidirectional driving. An example of this, and a more detailed description of GS, is contained in Section IV. Under generalized synchrony, it is likely that no conclusions can be made from observed time series alone concerning directionality of the coupling.

In systems that are dominated by linear dynamics or noise, Granger causality [3] is recognized as an effective means of ascertaining directionality of coupling between systems from observed time series. Granger causality is based on the intuitive idea that causal coupling $X \to Y$ is established when the prediction of system Y improves when observations from system X are included in the model. Implementations using linear cointegration, entropy [4] or mutual information [5, 6] have been developed and used successfully for this purpose.

Therefore, it is somewhat surprising that for finite-dimensional nonlinear systems under generic conditions, Granger causality-based methods are ill-posed. Indeed, the phenomenon of generalized synchrony means that if X drives Y, once the state of Y is identified from observations (for example by time-delay embedding, assuming sufficient data is available), X obviously can provide no further information about the state of Y, meaning that Granger causality has no power to identify the fact that $X \to Y$. Moreover, even without synchrony, Takens' theorem [7, 8] shows that X cannot add more information because the state of X can be determined from histories of Y. Conceptually, the lack of separability of effects that lies at the foundation of Granger causality eliminates its theoretical basis for making accurate directionality as-

sessments in deterministic nonlinear systems in the large data limit (see [9] for a discussion of this issue).

The concepts of time-delay embedding [7, 8] imply that each state of a dynamical system can be represented uniquely by a sufficiently long delay vector constructed from time series observations. Suppose X and Y are deterministic dynamical systems and that X has an input to Y, but that there is no return input from Y incident on X. Then, under technical conditions that are satisfied in general circumstances, we can expect to reconstruct the entire combined dynamical system $X \rightarrow Y$ from observations y_t of Y. The same is not possible from observations on X only, since observations on X will lack information about Y, and therefore fail to reconstruct the entire coupled system. We exploit this asymmetry in our methods.

This asymmetry has been exploited before, for example in [10], and more widely in the Convergent Cross Mapping method [9]. This latter reference included the crucial ingredient of tracking the asymmetry as a function of the length of data, and was successfully applied to experimental data in [11, 12], and in a version based on sorting in [13]. Exploitation of this asymmetry to study latent coupling in deterministic systems was addressed in [14] and more recently in [15].

Inspired by these prior results, here we describe a unifying approach, that attempts to affirmatively detect directional and latent coupling, and is compatible with statistical methods that provide confidence intervals. We exploit the asymmetry in delay-coordinate reconstructions to establish these statistical tests. The approach is robust in that it depends only the relative sizes of the pairwise distances of points in the reconstructions with respect to others.

II. DELAY EMBEDDING

The tests in the next section rely on the reconstruction of dynamics from a history of observations, in particular time series. The theory behind reconstruction was put on a firm theoretical footing by the embedding theorem of Takens [7], following similar ideas in [16]. This foundational work led to applications in a wide array of fields, including geophysics [17, 18], neuroscience [10, 19], ecology [9, 20, 21], fault analysis [22], and data assimilation [23].

Assume that $\{x_t\}$, $1 \le t \le N$, is a time series measured from a finite-dimensional compact attractor X of a dynamical system at discrete times t. For a given time t, define the vector of observations $X_t = [x_t, x_{t-1}, \ldots, x_{t-e+1}]$ for some integer e which we call the embedding dimension. The collection of such X_t comprise the time-delay reconstruction of the attractor X. Here we are assuming that there is a measurement function h_X from the system X to the real numbers that produces $x_t = h_X(t)$, an observation of system X at time t.

The main theorem of [7] states that for generic conditions and sufficiently large e, there is a one-to-one bijection between the states of a smooth manifold X and the reconstructed states $X_t = [x_t, x_{t-1}, \ldots, x_{t-e+1}]$. The underlying genericity assumption for this theorem is that both the dynamics of X and the observation function h_X are not special, meaning that if it fails for X, there is another system infinitesimally close to X for which the conclusion holds. Later work [8] simplified the main idea somewhat, by allowing X to be fractal (a non-manifold), and requiring genericity on the observation function only, assuming mild and specific conditions on the equilibria of the attractor X. The latter article also replaced the concept of genericity itself with a more probabilistic concept called prevalence.

Expanded versions of Takens' theorem show that in place of univariate time series, multivariate time series may be used to reconstruct the dynamical system [8, 24]. In this article, the results remain unchanged whether the reconstruction is made from univariate or multivariate times series, since only the fundamental property of a one-to-one correspondence is relevant to the success of the statistical tests.

The two tests we propose depend on the bijection promised by the theory under general conditions. That is, for each state of X, or X and Y if they are coupled, we assume that both the dynamics and the observations are not special, so that the delay coordinate reconstruction faithfully replicates the upstream dynamical states.

III. STATISTICAL TESTS

Two statistical tests are proposed. The goal of the first (DetC) is to determine whether two systems are coupled at all. If the pair of systems passes this test, we propose a second test (DirC) to determine the direction of coupling, or more precisely, to identify one of the relationships (1) - (3) above.

A. Detection of coupling

Delay coordinate embedding leads to a test of the existence of coupling between X and Y, which we call the Detection of Coupling Test. Assume we have simultaneous time series $\{x_t\}$ and $\{y_t\}$ of length N measured from X and Y, respectively, and that we form delay coordinate vectors $\{X_t\}$ and $\{Y_t\}$ as above. Fix time t and a number of nearest neighbors n. Find the n nearest neighbors $\{Y_{t_1}, \ldots, Y_{t_n}\}$ of Y_t in terms of Euclidean distance. Consider the set $\{X_{t_1}, \ldots, X_{t_n}\}$, in other words the simultaneous appearances of the reconstructed X_{t_j} in the X system, and consider their relationship as neighbors to X_t, the simultaneous correlate to Y_t. If X and Y are independent, the set $\{X_{t_1}, \ldots, X_{t_n}\}$ should have no special relationship to X_t. On the other hand, if X drives Y, then each Y_t determines a unique state in X and Y, and the nearest neighbors Y_{t_j} correspond to simultaneous X_{t_j} that are relatively near X_t (in particular much

nearer than random choices from the X time series). The same is true if instead of $X \to Y$, there is a latent coupling, meaning a third system D such that $D \to X$ and $D \to Y$.

This distinction can be analyzed statistically, due to the theory of order statistics (see, e.g., [25]). Given a uniform random choice of n numbers from the interval $[0, 1]$, the jth smallest of the n numbers follows a beta distribution $\text{Beta}(j, n - j + 1)$, whose mean and variance are $\mu = j/(n+1)$ and $\sigma^2 = j(n+1-j)/((n+1)^2(n+2))$, respectively.

We will apply this fact in the following way. Consider a time step $1 \le t \le N$, the reconstructed state Y_t, and its n nearest neighbors $Y_{s_1}, \ldots, Y_{s_n}$ according to the Euclidean distance $d_i = |Y_t - Y_{s_i}|$ from Y_t. Define the subset $X_{s_1}, \ldots, X_{s_n}$ of contemporaneous delay vectors at the times $s_1, \ldots, s_n$, and their distances $d_i = |X_t - X_{s_i}|$ from X_t. As mentioned above, we would like to know whether the X subset is as close to X_t as the Y subset is close to Y_t.

Each distance $d_s = |X_t - X_s|$, $1 \le s \le N - e$, can be assigned a relative rank among all $N - e$ distances of the reconstructed states, as follows. If d_s is the R_sth smallest distance from X_s, let $r_s = R_s/(N-e)$ denote the relative rank. By construction, the r_s are uniformly distributed between 0 and 1. If the subset $S = \{r_{s_1}, \ldots, r_{s_n}\}$ is randomly chosen from the entire set, order statistics gives the expected relative position of the jth smallest number $r_{(j)}$ of the subset to be $j/(n+1)$. On the other hand, if the subset S is chosen randomly from a reduced proportion p of the entire set, then the expected relative position will be $jp/(n+1)$. Therefore we can use

$$\hat{p} = \frac{(n+1)r_{(j)}}{j} \tag{1}$$

for each $1 \le j \le n$, as an estimator for p. If the dynamical systems X and Y are uncorrelated, we expect to recover $p = 1$ for each j. If there is a direct or indirect coupling between the two systems, the contemporaneous delay vectors in S will be chosen from a portion of the attractor that is limited in extent, with a corresponding proportion of the state space $p < 1$.

Summarizing this discussion, the following test averages the estimate over the time series X_t as a null hypothesis for independence of the two time series.

Detection of Coupling Test: DetC(X,Y). For each t, find the n nearest neighbors of the delay vector Y_t and denote their times $s_1, \ldots s_n$. Sort the entire set of $N - e$ distances $|X_t - X_s|$ for $1 \le s \ne t \le N - e + 1$, and for each $1 \le j \le n$, find the relative rank $r_{(j)}$ of the jth smallest distance in the subset $S = \{|X_t - X_{s_1}|, \ldots, |X_t - X_{s_n}|\}$, among the entire set $\{|X_t - X_s|\}$. For each j, $1 \le j \le n$, an estimate $\hat{p}$ is found from the expected value

$$\hat{p} = \mathbb{E}\left[\frac{(n+1)r_{(j)}}{j}\right]$$

averaged over t. The value of p is equal to 1 if and only if there is no coupling between X and Y.
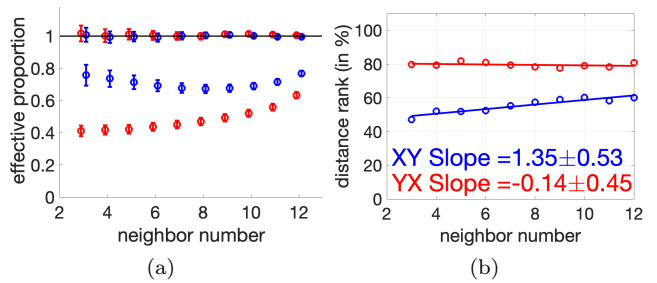

FIG. 1. DetC and DirC tests applied to time series from skew system $X \to Y$. (a) For zero forcing (upper curves), both DetC(X,Y) (blue traces) and DetC(Y,X) (red traces) show that 1 lies inside the confidence intervals for p, and for positive forcing (lower curves), 1 lies outside the confidence intervals. (b) For nonzero forcing, the slope estimate for DirC(X,Y) (blue trace) rejects zero, and DirC(Y,X) (red trace) does not, correctly implying a forcing $X \to Y$.

We apply a one-sample Student's t-test to assign a confidence interval to the estimate $\hat{p}$, A 95% interval around $\hat{p}$ contains the possible values of p with 95% certainty, and we would like to know whether the interval contains $p = 1$. Since the one-sample t-distribution has $N - 2$ degrees of freedom, it is well approximated by the normal distribution for large N, which represents two standard deviations for the 95% level. (Note that we can further average over e to decrease the variance even more, although we have not done so in the examples, where a fixed $e = 8$ was used.)

Figure 1(a) shows the results of the DetC test for observations of X and Y where $X \to Y$ for two different coupling strengths. When the coupling is zero (top red and blue curves) the null hypotheses $p = 1$ is not rejected, and we conclude there is no coupling. When the coupling is positive (lower red and blue curves) the null is rejected, and coupling is concluded.

The length-1000 time series in Fig. 1 were generated by coupled discrete dynamical systems. Both X and Y are two-cell networks of Hénon-like [26] maps. Specifically, each X or Y consists a four-dimensional discrete map of two Hénon-like maps coupled together, with equations

$$\begin{aligned}
x_{i+1} &= b_1 \cos x_i + c_1 y_i + d_1 u_i \\
y_{i+1} &= x_i \\
u_{i+1} &= b_2 \cos u_i + c_2 v_i + d_2 x_i \\
v_{i+1} &= u_i
\end{aligned} \tag{2}$$

Here the parameters b_i, c_i, and d_i were chosen randomly near $2.2, 0.1$, and 0.1, respectively, which resulted in chaotic dynamics in the network. The time series observation from each network is the x variable. The coupling between the two networks is achieved by adding the x-variable from one system to the u-variable of the other, multiplied by a coupling strength (either 0 or 0.2) in Fig. 1.

The map in Hénon's original paper [26] is often studied due to the fact that it has a chaotic attractor near a basin

boundary. Here we use a variant of the original Hénon map. The version used in (2) replaces a quadratic term in the x equation with a cosine, which retains the chaotic dynamics but moves the basin boundary of the attractor far away, to enhance stability when used in a network. In equation (2), two such maps are coupled together to produce a hyperchaotic attractor of correlation dimension slightly greater than 2.0. Each of the attractors X and Y in Fig. 1 and 2 is an attractor of this type.

B.   Direction of coupling

The second test assesses directionality of the coupling. Our general assumption is that there is an unknown finite-dimensional attractor embodying one of the above scenarios (1) - (3), and that it is observed in some way by two time series that have each been used to reconstruct dynamics as in Takens' theorem.

First assume that relation (1) holds. Thus a unidirectional coupling exists, which we will without loss of generality suppose is $X \to Y$. Further assume that time series recorded from both X and Y are used for time delay reconstruction using generic observations functions and sufficiently large embedding dimension e required by Takens' Theorem. Let $R(X)$ (resp. $R(Y)$) denote the reconstructed attractors, and let $T_{X,X}$, $T_{Y,X}$ and $T_{Y,Y}$ denote the projections from the reconstructions to X and Y as shown in the diagram:
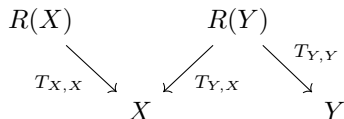

Takens' theorem implies that generically, there is a homeomorphism between $R(X)$ and X, and also between $R(Y)$ and the skew product $X \to Y$. Since there are projection maps from the skew product to X and Y, we see that the three maps in the diagram are well-defined. While the map $T_{X,X}$ is one-to-one, the other two maps are not: Although $R(Y)$ reconstructs the skew product $X \to Y$, unless X and Y are in general synchronization, $T_{Y,X}$ and $T_{Y,Y}$ will fail to be one-to-one, because there is no return coupling $Y \to X$. In particular, the inverse image $T_{Y,X}^{-1}(x)$ will generally be more than one point, since the state of X does not determine the contemporaneous state of Y.

As observers, our access to this diagram is only along the top row. According to the diagram, the reconstructed state Y_t in $R(Y)$ corresponds to a single point $T_{X,X}^{-1} T_{Y,X} Y_t$ in $R(X)$. That means for each Y_t, only a single pair of simultaneous states (X_t, Y_t) will exist. On the other hand, a reconstructed state X_t in $R(X)$ corresponds to a non-singleton set of points $T_{Y,X}^{-1} T_{X,X} X_t$ in $R(Y)$, since $T_{Y,X}$ is not one-to-one. This is the asymmetry referred to earlier. For each X_t, there is a multiple point set of simultaneous pairs (X_t, Y_t). If we can use the

observed time series to distinguish between the two cases, one point versus multiple points, for pairs (X_t, Y_t), we can infer in which direction the coupling exists.

The case (2) of bidirectional coupling is simpler, because under generic conditions, each of the reconstructions $R(X)$ and $R(Y)$ are in one-to-one correspondence with states of the bidirectionally-coupled system $X \leftrightarrow Y$, and therefore $R(X)$ and $R(Y)$ are in one-to-one correspondence. Unlike the asymmetric case (1), in case (2), for each X_t in $R(X)$ there is a unique Y_t in $R(Y)$, and vice versa.

Now consider the case of latent coupling (3), where there is an unobserved driver D that is coupled to both the observed systems X and Y, but no direct connection between X and Y, as in the diagram:
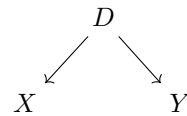
Then observations from X (resp. Y) reconstruct the skew product $D \to X$ (resp., $D \to Y$). This results in the diagram
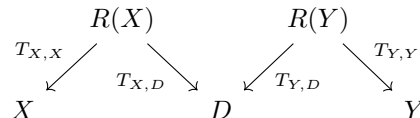
where none of the four maps are one-to-one. Now both $T_{X,D}^{-1} T_{Y,D} Y_t$ and $T_{Y,D}^{-1} T_{X,D} X_t$ comprise sets of multiple points, i.e. more than one. This observation suggests a criterion for X and Y being driven by an unobserved system, which is that for each X_t, there are multiple simultaneous (X_t, Y_t), and the same for each Y_t.

Although not essential for the goals of the present article, this structure can be exploited to theoretically reconstruct a latent driver D from time series observations as follows. Consider a state Y_t in $R(Y)$. In this case, $T_{X,D}^{-1} T_{Y,D} Y_t$ consists of multiple points $\{X_i\}$ by assumption, usually dispersed along a proper subset of $R(X)$. For each of the X_i, we can repeat the process in the opposite direction, which represent different indices in the time series. By assumption, the $T_{Y,D}^{-1} T_{X,D} X_i$ are multiple points sets in Y, only one of which one is the original Y_t. Recursing this process identifies an equivalence class of point sets in both $R(X)$ and $R(Y)$, which corresponds to a theoretical state d linked to the multiple points in $R(X)$ and $R(Y)$. If we repeat this construction for all Y_t in $R(Y)$ and X_t in $R(X)$, these equivalence classes represent the states of an unobserved attractor D that implies scenario (3), i.e. latent coupling $D \to X$ and Y. This process of reconstructing the driver was the subject of [14], which was recently put on a more stable numerical footing [15].

The preceding discussion clarifies the critical distinction between the cases (1) - (3), which depends on whether the sets (X_t, Y_t), for fixed X_t or Y_t, are singleton

sets or more than one element. The possibilities can be summarized in a diagram:

	$(X_t$, many $Y_s)$	$(X_t$, single $Y_s)$
(many X_s, Y_t)	D $\swarrow \quad \searrow$ $X \qquad Y$	$Y \to X$
(single X_s, Y_t)	$X \to Y$	$X \leftrightarrow Y$ or Generalized Synchrony

The following test is designed to determine which of these cases hold, using only the observed time series. The logic is as follows: If for a given Y_t, there is a unique corresponding time delay vector X_t, then if we look at the X_s that are time-contemporaneous with Y_s that are close to Y_t, these X_s should be close to X_t. The test checks to see whether the X_s progressively move farther from X_t as the Y_s are moving away from Y_t.

Direction of Coupling Test: DirC(X,Y). For each t and $1 \leq j \leq n$, find the jth nearest neighbor of the delay vector Y_t and denote its time s_j. Sort the entire set of distances $|X_t - X_s|$ for $1 \leq s \neq t \leq N$, and find the percentile rank $r_{(j)}$ of the distance $|X_t - X_{s_j}|$ among the entire set. Let $\hat{r}_{(j)}$ denote the average percentile rank over all t. Plot the best fit line through the points $(j, \hat{r}_{(j)})$.

A slope greater than zero is evidence that for each Y_t, the set of possible simultaneous pairs (X_t, Y_t) is a single pair. A slope indistinguishable from zero is evidence that it is a set of multiple pairs (i.e., not a singleton).

Confidence intervals for the DirC test can be developed similarly to the DetC test. For the line fit $y = \alpha + \beta t$ where we use a one-sample Student-t test to compare the slope β with the null hypothesis $\beta = 0$, the radius of the 95% confidence interval is $s = (2 \sum_{i=1}^{n}(y_i - \hat{y}_i)^2)/((n-2) \sum_{i=1}^{n}(x_i - \overline{x})^2)$.

Figure 1(b) demonstrates the use of DirC. Here the two 2-cell Henon maps X and Y are coupled as $X \to Y$. The slope in DirC(X,Y) is positive, which provides evidence that for each Y_t, there is a unique X_t that can exist simultaneously with Y_t. On the other hand, the slope of DirC(Y,X) cannot be distinguished from zero, meaning that there are multiple Y_t corresponding to a given X_t. These results are consistent with case (1), unidirectional coupling from X to Y.

The two tests proposed in this article can be used together in the following way. If the Detection of Coupling Test is passed in either of the X or Y directions, the remaining possible relations are (1) - (3). Then, the Direction of Coupling Test applied to both X and Y will distinguish between the remaining four possibilities. In the table, DirC(X,Y) + means the slope is determined to be greater than zero, and DirC(X,Y) 0 means zero slope cannot be rejected.
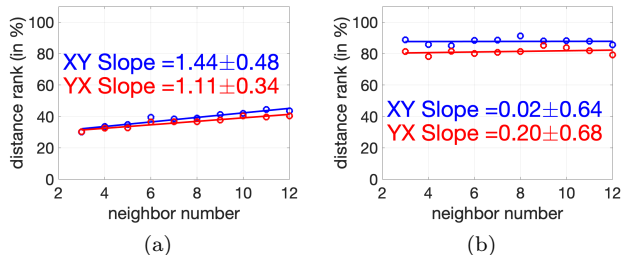

FIG. 2. DirC test applied to (a) bidirectional coupling $X \leftrightarrow Y$ and (b) latent coupling $D \to X, D \to Y$. DirC(X,Y) in blue and DirC(Y,X) in red. (a) Both slope estimates reject zero slope, correctly implying bidirectional driving. (b) Neither of the slope estimates reject the null hypothesis of zero slope, correctly implying latent coupling from an unobserved common driver.

	DirC(Y,X) 0	DirC(Y,X) +
DirC(X,Y) 0	latent $D \to X, Y$	$Y \to X$
DirC(X,Y) +	$X \to Y$	$X \leftrightarrow Y$ or GS

Fig. 2 demonstrates the use of the DirC test for time series collected from two 2-cell Henon maps in cases (2) and (3). In Fig. 2(a), the times series are generated from a system $X \leftrightarrow Y$ with two-way driving. The DirC(X,Y) and DirC(Y,X) both identify positive slope outside the confidence interval, concluding bidirectional driving as in case (2). In Fig. 2(b), zero slope cannot be rejected in either DirC(X,Y) or DirC(Y,X), indicating latent driving from an unobserved system as in case (3).

The tests can also be applied to continuous systems such as coupled Lorenz [27] systems. For example, let $X, Y,$ and Z each be systems of the form

$$\dot{x} = \sigma(y - x)$$
$$\dot{y} = -xz + \rho x - y \qquad (3)$$
$$\dot{z} = xy - \beta z$$

with slightly different parameter values near $\sigma = 10, \rho = 28, \beta = 8/3$. Further, we couple them in a small network as in (A) or (B) below,

(A) $\qquad X$
 $\swarrow \quad \nwarrow$
 $Y \longrightarrow Z$

(B) $\qquad X$
 $\swarrow \quad \searrow$
 $Y \longrightarrow Z$
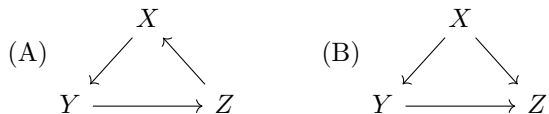
where each arrow represents a term cx proportional to the x-variable of the driving Lorenz system added to the y-variable of the target system. The following table shows the results of applying DirC to both scenarios, using $s = 0.3$:
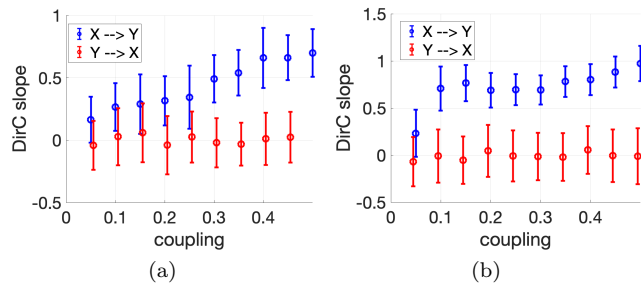
FIG. 3. Comparison of DirC versus coupling strength for unidirectional driving $X \to Y$. The mean of the DirC slope over 50 realizations of length 1000 times series of $X \to Y$ is plotted along with its standard error. (a) Two-cell Hénon example (b) Lorenz equations.
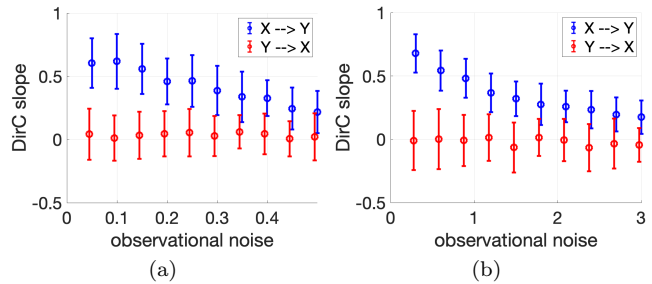

FIG. 4. Comparison of DirC versus observational noise for unidirectional driving $X \to Y$. The mean of the DirC slope over 50 realizations of length 1000 times series of $X \to Y$ is plotted along with its standard error. (a) Two-cell Hénon example (b) Lorenz equations.

(A) slope		
$X \to Y$	0.68 ± 0.13	+
$Y \to X$	0.50 ± 0.26	+
$Y \to Z$	0.61 ± 0.17	+
$Z \to Y$	0.39 ± 0.22	+
$Z \to X$	0.62 ± 0.24	+
$X \to Z$	0.42 ± 0.14	+

(B) slope		
$X \to Y$	0.76 ± 0.20	+
$Y \to X$	0.15 ± 0.36	0
$Y \to Z$	0.25 ± 0.17	+
$Z \to Y$	0.13 ± 0.16	0
$Z \to X$	0.07 ± 0.23	0
$X \to Z$	0.47 ± 0.22	+

The slopes are shown along with the 95% confidence intervals. A plus sign in the right column means that the corresponding coupling was determined to exist by the test, and 0 means the coupling was not detected to the confidence level. There are six possible couplings. We note that in scenario (A), all three systems are driven by each of the others, directly or indirectly, and accordingly, the DirC test detects that coupling in each case. In scenario (B), the DirC test correctly detects the three drivings $X \to Y, Y \to Z$, and $X \to Z$, and finds no evidence of the other three possibilities.

To investigate the sensitivity of the DirC test with respect to parameters, we compared the results for different coupling strengths and observational noise levels. Fig. 3(a) shows the dependence on the DirC test on coupling strength. The data series of length 1000 was generated from a coupled system $X \to Y$ with dynamics from the two-cell Hénon map (2). The coupling between the two networks is achieved by adding a constant c times the x-variable from X to the u-variable of Y. Fig. 3(a) plots the DirC statistic versus the coupling strength c. For small c, the 95% confidence intervals of statistics DirC(X,Y) and Dir(Y,X) overlap. For increased c, they no longer overlap and definitively pick up the correct direction of driving. Fig. 3(b) shows a similar result from coupled Lorenz systems $X \to Y$.

Fig. 4(a) shows that the DirC test degrades gracefully with increased observational noise. For small noise the test easily identifies the driving $X \to Y$, and as noise is increased, the test begins to fail. Fig. 4(b) shows similar outcomes but for Lorenz systems. In Fig. 4, time series of length 1000 were sampled from the Lorenz systems,

using sampling interval $\Delta t = 0.05$.

IV. GENERALIZED SYNCHRONY

As mentioned above, the phenomenon of generalized synchrony (GS) plays a clarifying role in the study of causation between dynamical systems. A typical example is given by a coupled skew system $X \to Y$ where there is no feedback from Y to X, but where a one-to-one correspondence develops between states of X and states of Y. That is, after transient behavior, each state of X coexists with a unique state of Y, and vice versa. This can occur even when X and Y are different dynamical systems.

On the one hand, by its existence, GS shows that Granger causality cannot detect the fact that $X \to Y$, since knowledge of the X state cannot add to the ability to predict future states of Y. (In fact, the logic of such an ability would be circular, since one could say the same in the opposite direction.) Furthermore, GS is also a wild card for the DirC calculation, because the sets of possible pairs (X_t, Y_t) discussed in the derivation of the DirC method are singletons, both in the non-GS case of bidirectional coupling $X \leftrightarrow Y$ and the case of GS caused by unidirectional coupling. In other words, for deterministic dynamical systems, it is challenging for any method to distinguish GS under $X \to Y$ from the relationship $X \leftrightarrow Y$ on the basis of observed time series alone.

We include here two illuminating examples of generalized synchronization for chaotic flows. The first is constructed from two Lorenz systems with different parameters. Let X and Y be systems of form (3) with parameters $\rho = 27$ and $\rho = 30$, respectively. With coupling of $1.0x$ from X added to the y-variable of Y, the system is not in GS. However, changing the coupling to $1.5x$ induces GS between X and Y. On the other hand, a coupling of cz from X to the y-variable of Y for virtually any $c > 0$ causes GS.

To positively verify GS in this example, we first will assume knowledge of all six phase variables of X and
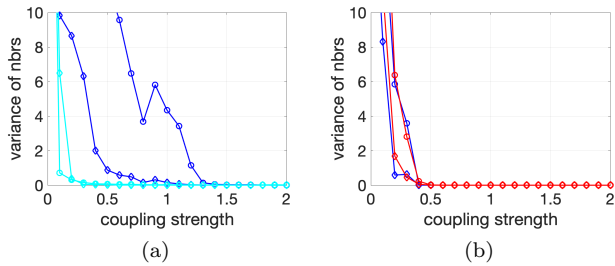
(a)

(b)

FIG. 5. Test for generalized synchrony. The plot shows the variance of neighbors of Y points chosen contemporaneously from neighbors of an X point (upper curves, plotted as circles), and vice versa (lower curves, plotted with diamonds). GS is indicated when both variances drop to near zero. (a) Dark blue curves: Coupling cx from Lorenz X to y-variable of Lorenz Y. Light blue curves: Coupling $c(x + z)$ from Lorenz X to y-variable of Lorenz Y. (b) Dark blue curves: coupling cx from Rössler X to y-variable of Lorenz Y. Red curves: coupling cy from Rössler X to y-variable of Lorenz Y. A set of $k = 10$ nearest neighbors was used in these plots.

Y. A long trajectory of the coupled system $X \to Y$ is generated, and a strategy similar to DirC, but that does not require delay coordinates, is used. Specifically, let p_t denote the state of the three Lorenz variables from the X system at time t, and denote by q_t the three Y system variables at the same time. For a given p_t on the X attractor at time t, we find the nearest k neighbors p_s for some number k. Then we calculate the variance of the set of distances $|q_s - q_t|$ for the same t and group of s. The signature of global synchrony is when this variance drops to near zero.

The median of this variance as we sweep over p_t on X is plotted in Figure 5(a), along with the reverse: Starting with k neighbors of q_t on Y and gathering contemporaneous points on X. For small coupling $X \to Y$ achieved by adding a coupling strength times the x variable of X to the y variable of Y, there is no synchronization. When coupling strength reaches 1.5, the systems are in general synchronization. Figure 5(a) also shows that using a coupling proportional to $x + z$ instead of to x leads to GS for almost all positive coupling strengths.

When we can only measure time series of X and Y, and have a smaller data set, we can apply the DirC test to distinguish between unidirectional coupling $X \to Y$ that does and does not result in GS. The results of analyzing a 1000 point time series of x-coordinates observed from X and Y is shown in Fig. 6. In Figure 6(a), X drives Y by adding $0.4x$ to the y-variable of Y. Fig. 5(a) indicates that generalized synchrony does not occur in this case. Accordingly, DirC correctly concludes unidirectional coupling $X \to Y$. When the driving is increased to $1.4x$, generalized synchrony occurs, and we find that DirC rejects zero slope in both directions, as shown in Fig. 6(b). The explanation for this result is generalized synchronization caused by unidirectional coupling $X \to Y$, not two-way driving $X \leftrightarrow Y$.

(a)

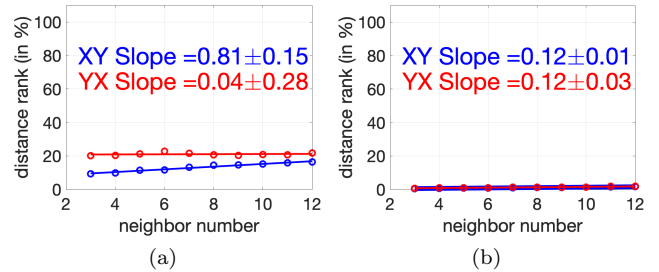(b)

FIG. 6. Outcome of DirC for data from Lorenz systems $X \to Y$, DirC(X,Y) in blue and DirC(Y,X) in red. (a) Coupling $c = 0.4x$, (not in general synchronization), only XY slope statistically greater than zero, implying case (1). (b) Coupling $c = 1.4x$ (forces generalized synchronization), XY and YX slopes greater than zero implies case (2). Data consists of the x-coordinate time series of 1000 points observed from X and Y with sampling interval $\Delta t = 0.05$.

For a Lorenz attractor driven by the output of another Lorenz attractor, GS is relatively easy to achieve, even for non-identical systems as above. If the driving system sends a signal cx to the response system for c sufficiently large (as in the above paragraph), the system enters GS. Also, the signal $c(x + z)$ used in the same way causes GS for any $c > 0$.

The second example of generalized synchrony occurs for the coupling $X \to Y$ where X is a Rössler attractor driving a Lorenz attractor Y, with the standard parameters $\sigma = 10, \rho = 28, \beta = 8/3$. The Rössler equations [28] are

$$
\begin{aligned}
\dot{x} &= -y - z \\
\dot{y} &= x + ay \\
\dot{z} &= b + (x - c)z
\end{aligned}
\tag{4}
$$

and we will use the parameter settings $a = 0.1, b = 0.1$, and $c = 14$. As before, the Lorenz attractor Y will be driven by adding a signal from the X system to the y variable of Y. Fig. 5(b) shows that relatively small couplings from the x or y variable of X to the y variable of Y causes generalized synchrony.

V. DISCUSSION

The tests proposed here are designed to learn the coupling characteristics of a pair of time series, produced simultaneously by two different deterministic processes. Our goal is to exploit the fact of asymmetry between the dual reconstructions as simply and robustly as possible, in order to tease conclusions from time series that may be short and perturbed by observational noise. These tests may be applied to univariate time series (or multivariate time series) measured from two or more processes, as long as the state space reconstructions afford a one-to-one correspondence to the original systems.

Among the innovations of our methods are the reduction of the information of distances between points by treating them ordinally, in order to make the results robust to measurement noise. This also allows us to apply the theory of order statistics in the DetC test to determine dependence of time series, and to retrieve statistical conclusions by assigning confidence intervals. The DirC test exploits the asymmetry of paired Takens delay coordinate reconstructions in order to distinguish unidirectional, bidirections, and latent coupling. Our goal in this article is to deploy these innovations in the simplest possible way so that they may succeed with minimal data requirements.

These tests will gradually fail in general for stochastic systems, due to the degradation of the delay coordinate embedding that is foundational for this approach. Other methods more suited toward working under stochastic assumptions will work better in general, although where the trade-off occurs will be dependent on the details of the situation.

A key requirement for faithfulness of the delay coordinate embedding is genericity of the dynamics and observations. This includes both the internal dynamics of X and Y, and in addition, the connections (if any) between the two systems. We can expect this genericity to exist normally in natural systems, in the absence of a particular structural reason that defeats the hypotheses of Takens' theorem.

The existence of generalized synchrony presents an interesting complication to the problem of distinguishing coupling direction from time series. On the one hand, when couplings exist in both directions, i.e. $X \leftrightarrow Y$, under generic conditions, the delay coordinate embedding from either X or Y will uniquely reconstruct a corresponding state from the other that will occur simultaneously, so that a one-to-one correspondence exists between reconstructed states. This is also the definition of generalized synchrony, but as demonstrated here, GS can occur solely from unidirectional coupling $X \rightarrow Y$. Therefore evidence such as the DirC results in Fig. 6 cannot distinguish between these two cases, and it is unlikely that any method based on time series alone can do so, in the deterministic nonlinear case.

[1] N. F. Rulkov, M. M. Sushchik, L. S. Tsimring, and H. D. Abarbanel, Generalized synchronization of chaos in directionally coupled chaotic systems, Physical Review E **51**, 980 (1995).

[2] L. M. Pecora and T. L. Carroll, Synchronization in chaotic systems, Physical Review Letters **64**, 821 (1990).

[3] C. W. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica: Journal of the Econometric Society , 424 (1969).

[4] T. Schreiber, Measuring information transfer, Physical review letters **85**, 461 (2000).

[5] J. Sun and E. M. Bollt, Causation entropy identifies indirect influences, dominance of neighbors and anticipatory couplings, Physica D: Nonlinear Phenomena **267**, 49 (2014).

[6] J. Runge, Causal network reconstruction from time series: From theoretical assumptions to practical estimation, Chaos: An Interdisciplinary Journal of Nonlinear Science **28** (2018).

[7] F. Takens, Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence, Warwick*, Lecture Notes in Mathematics, Vol. 898, edited by D. Rand and L.-S. Young (Springer Berlin / Heidelberg, 1981) pp. 366–381.

[8] T. Sauer, J. Yorke, and M. Casdagli, Embedology, Journal of Statistical Physics **65**, 579 (1991).

[9] G. Sugihara, R. May, H. Ye, C.-h. Hsieh, E. Deyle, M. Fogarty, and S. Munch, Detecting causality in complex ecosystems, Science **338**, 496 (2012).

[10] S. J. Schiff, P. So, T. Chang, R. E. Burke, and T. Sauer, Detecting dynamical interdependence and generalized synchrony through mutual prediction in a neural ensemble, Physical Review E **54**, 6708 (1996).

[11] H. Ye, E. R. Deyle, L. J. Gilarranz, and G. Sugihara, Distinguishing time-delayed causal interactions using convergent cross mapping, Scientific Reports **5**, 14750 (2015).

[12] B. Gao, J. Yang, Z. Chen, G. Sugihara, M. Li, A. Stein, M.-P. Kwan, and J. Wang, Causal inference from cross-sectional earth system data with geographical convergent cross mapping, Nature Communications **14**, 5875 (2023).

[13] L. Breston, E. J. Leonardis, L. K. Quinn, M. Tolston, J. Wiles, and A. A. Chiba, Convergent cross sorting for estimating dynamic coupling, Scientific Reports **11**, 20374 (2021).

[14] T. Sauer, Reconstruction of shared nonlinear dynamics in a network, Physical Review Letters **93**, 198701 (2004).

[15] W. Gilpin, Recurrences reveal shared causal drivers of complex time series, arXiv preprint arXiv:2301.13516 (2023).

[16] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, Geometry from a time series, Physical Review Letters **45**, 712 (1980).

[17] E. Hossain, M. O. Gani, D. Dunmire, A. Subramanian, and H. Younas, Time series classification of supraglacial lakes evolution over Greenland ice sheet, ArXiv **abs/2410.05638**, 10.48550/arXiv.2410.05638 (2024).

[18] H. Astudillo, F. A. Borotto, and R. A. del Río, Embedding reconstruction methodology for short time series – application to large El Niño events, Nonlinear Processes in Geophysics **17**, 753 (2010).

[19] E. Kwessi and L. J. Edwards, Analysis of EEG data using complex geometric structurization, Neural Computation **33**, 1942 (2021).

[20] E. R. Deyle, R. M. May, S. B. Munch, and G. Sugihara, Tracking and forecasting ecosystem interactions in real time, Proceedings of the Royal Society B: Biological Sciences **283**, 20152258 (2016).

[21] S. B. Munch, T. L. Rogers, C. C. Symons, D. Anderson, and F. Pennekamp, Constraining nonlinear time

series modeling with the metabolic theory of ecology, Proceedings of the National Academy of Sciences **120**, e2211758120 (2023).

[22] D. Lee, J. Rorie, and A. Sabater, An investigation of time series embeddings and topological data analysis for fault analysis, 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE) , 1483 (2023).

[23] F. Hamilton, T. Berry, and T. Sauer, Ensemble Kalman filtering without a model, Physical Review X **6**, 011021 (2016).

[24] E. R. Deyle and G. Sugihara, Generalized theorems for nonlinear state space reconstruction, Plos one **6**, e18295 (2011).

[25] J. E. Gentle, *Computational statistics*, Vol. 308 (Springer, 2009).

[26] M. Hénon, A two-dimensional mapping with a strange attractor, in *The Theory of Chaotic Attractors* (Springer, 1976) pp. 94–102.

[27] E. N. Lorenz, Deterministic nonperiodic flow, Journal of the Atmospheric Sciences **20**, 130 (1963).

[28] O. E. Rössler, An equation for continuous chaos, Physics Letters A **57**, 397 (1976).