Convergent Cross-Mapping and Pairwise Asymmetric Inference

James M. McCracken[*] and Robert S. Weigel[†]

School of Physics, Astronomy, and Computational Sciences
George Mason University
4400 University Drive MS 3F3, Fairfax,VA. 22030-4444
(Dated: July 8, 2014)

Convergent Cross-Mapping (CCM) is a technique for computing specific kinds of correlations between sets of times series. It was introduced by Sugihara *et al.* [1] and is reported to be "a necessary condition for causation" capable of distinguishing causality from standard correlation. We show that the relationships between CCM correlations proposed in [1] do not, in general, agree with intuitive concepts of "driving", and as such, should not be considered indicative of causality. It is shown that CCM causality analysis implies causality is a function of system parameters for simple linear and nonlinear systems. For example, in a RL circuit, both voltage and current can be identified as the driver depending on the frequency of the source voltage. It is shown that CCM causality analysis can, however, be modified to identify asymmetric relationships between pairs of time series that are consistent with intuition for the considered example systems for which CCM causality analysis provided non-intuitive driver identifications. This modification of the CCM causality analysis is introduced as "pairwise asymmetric inference" (PAI) and examples of its use are presented.

I. INTRODUCTION

Modern time series analysis includes techniques meant to discern "driving" relationships between different data sets. These techniques have found application in a wide range of fields including neuroscience (e.g., [2]), economics (e.g., [3, 4]), and climatology (e.g., [5]). General casual relationships in time series data are also being studied in an effort to understand causality itself (e.g., [6]).

To date, most techniques for "causal inference" in time series fall into two broad categories: those related to transfer entropy and those related to Granger causality. Transfer entropy (introduced in [7]) and Granger causality (introduced in [8]) are known to be equivalent under certain conditions [9]. In this article, we investigate a casual inference technique, called Convergent Cross-Mapping (CCM), that was recently introduced by Sugihara *et al.* [1]; [10].

CCM is described as a technique that can be used to identify a necessary condition for causality between time series and is intended to be useful in situations where Granger causality is known to be invalid (i.e., in dynamic systems that are "nonseperable" [1]). Granger causality is not causality as it is typically understood in physics [11–13]. We show that a similar conclusion can be made regarding CCM causality.

CCM has been used to draw conclusions regarding the sardine-anchovy-temperature problem [1]; confirm predictions of climate effects on sardines [14]; compare the driving effects of precipitation, temperature, and solar radiation on the atmospheric CO_2 growth rate [15]; study the driving relationship between pressure and displacement of abdominal parts in insects [16]; and to quantify cognitive control in developmental psychology [17]. The wide range of applications already appearing for the relatively new CCM technique is a testament to the importance of time series causality studies. This work presents examples in which CCM does not provide consistent qualification of an intuitive notion of causality. (However, the domain of applicability of CCM remains an open question; i.e., the method may have worked as expected in the above-cited papers despite its apparent failure in the examples presented in this article.)

We begin with a review of the work of Sugihara *et al.* [1], including an extended evaluation of the coupled logistic map example. After showing examples where CCM analysis gives results that are inconsistent with intuitive notions of driving, we introduce "pairwise asymmetric inference" (PAI) and show that it can be used to identify asymmetric relationships that are consistent with intuitive notions of driving.

II. CONVERGENT CROSS-MAPPING

CCM is closely related to simplex projection [18, 19], which predicts a point in the times series X at a time $t + 1$, labeled X_{t+1}, by using the points with the most similar histories to X_t. Similarly, CCM uses points with the most similar histories to X_t to estimate Y_t. The CCM correlation is the squared Pearson correlation coefficient [20] between the original time series Y and an estimate of Y made using its convergent cross-mapping with X, which is labeled as $Y|\tilde{\mathbf{X}}$:

$$C_{YX} = \left[\rho(Y, Y|\tilde{\mathbf{X}}) \right]^2 \ .$$

Any pair of times series, X and Y, will have two CCM correlations, C_{YX} and C_{XY}, which are compared to de-

---
[*] jmccrac1@masonlive.gmu.edu
[†] rweigel@gmu.edu

termine the CCM causality. For example, Sugihara *et al.* [1] define a difference of CCM correlations

$$\Delta = C_{YX} - C_{XY} \tag{1}$$

and use the sign of Δ to determine the CCM causality between X and Y.

If X can be estimated using Y better than Y can be estimated using X (e.g., if $\Delta < 0$), then X is said to "CCM cause" Y.

A. CCM Algorithm

The CCM algorithm may be written in terms of five steps:

1. Create the *shadow manifold* for X, called $\tilde{\mathbf{X}}$;

2. Find the nearest neighbors to a point in the shadow manifold at time t, which is labeled $\tilde{\mathbf{X}}_t$;

3. Create weights using the nearest neighbors;

4. Estimate Y using the weights; (this estimate is called $Y|\tilde{\mathbf{X}}$); and

5. Compute the correlation between Y and $Y|\tilde{\mathbf{X}}$.

The steps are described in more detail below.

1. Create Shadow Manifold $\tilde{\mathbf{X}}$

Given an embedding dimension E, the *shadow manifold* of X, called $\tilde{\mathbf{X}}$, is created by associating an E-dimensional vector (also called a *delay vector*) to each point X_t in X, i.e., $\tilde{\mathbf{X}}_t = \left(X_t, X_{t-\tau}, X_{t-2\tau}, \ldots, X_{t-(E-1)\tau}\right)$. The first such vector is created at $t = 1 + (E-1)\tau$ and the last is at $t = L$ where L is the number of points in the time series (also called the *library length*).

2. Find Nearest Neighbors

The minimum number of points required for a bounding simplex in an E-dimensional space is $E + 1$ [18, 19]. Thus, the set of $E + 1$ nearest neighbors must be found for each shadow manifold $\tilde{\mathbf{X}}_t$. For each $\tilde{\mathbf{X}}_t$, the nearest neighbor search results in a set of distances that are ordered by closeness $\{d_1, d_2, \ldots, d_{E+1}\}$ and an associated set of times $\{\hat{t}_1, \hat{t}_2, \ldots, \hat{t}_{E+1}\}$. The distances from $\tilde{\mathbf{X}}_t$ are

$$d_i = D\left(\tilde{\mathbf{X}}_t, \tilde{\mathbf{X}}_{\hat{t}_i}\right) \ ,$$

where $D(\mathbf{a}, \mathbf{b})$ is the Euclidean distance between vectors $\mathbf{a}$ and $\mathbf{b}$.

3. Create Weights

Each of the $E + 1$ nearest neighbors are be used to compute an associated weight. The weights are defined as

$$w_i = \frac{u_i}{N} \ ,$$

where $u_i = e^{-d_i/d_1}$ and the normalization factor is $N = \sum_{j=1}^{E+1} u_j$.

4. Find $Y|\tilde{\mathbf{X}}$

A point Y_t in Y is estimated using the weights calculated above. This estimate is

$$Y_t|\tilde{\mathbf{X}} = \sum_{i=1}^{E+1} w_i Y_{\hat{t}_i} \ .$$

5. Compute the Correlation

The CCM correlation is defined as

$$C_{YX} = \left[\rho\left(Y, Y|\tilde{\mathbf{X}}\right)\right]^2 \ ,$$

where $\rho(A, B)$ is the standard Pearson's correlation coefficient between A and B.

The CCM algorithm depends on the embedding dimension E and the lag time step τ. A dependence on E and τ is a feature of most state space reconstruction (SSR) methods [21–23], so an E and τ dependence is not unexpected. Sugihara *et al.* mention that "optimal embedding dimensions" are found using univariate SSR [1] (supplementary material), and other methods for determining E and τ for SSR algorithms can be found in the literature (e.g., [21, 23, 24]).

B. CCM Example

Consider the example system used by Sugihara *et al.* [1]:

$$X_t = X_{t-1}\left(r_x - r_x X_{t-1} - \beta_{xy} Y_{t-1}\right) \tag{2}$$
$$Y_t = Y_{t-1}\left(r_y - r_y Y_{t-1} - \beta_{yx} X_{t-1}\right) \tag{3}$$

where the parameters $r_x, r_y, \beta_{xy}, \beta_{yx} \in \mathbb{R} \geq 0$. This pair of equations is a specific form of the two-dimensional coupled logistic map system [25].

In this example, the CCM causality relationship between X and Y is determined using a sampling of both the initial conditions and the system parameters, calculating Δ, and demonstrating the necessary convergence. The term *convergence* is used here in the same sense as

it was used in [1]; i.e., "convergence means that cross-mapped estimates improve in estimation skill with time-series length L" [26]. The dynamic parameters r_x and r_y are sampled from normal distributions $\mathcal{N}(\mu_{rx}, \sigma_{rx})$ and $\mathcal{N}(\mu_{ry}, \sigma_{ry})$, respectively. The initial conditions X_0 and Y_0 are also sampled from normal distributions, specifically $\mathcal{N}(\mu_{x0}, \sigma_{x0})$ and $\mathcal{N}(\mu_{y0}, \sigma_{y0})$. The coupling parameters β_{xy} and β_{yx} are then varied over the interval $[10^{-6}, 1]$ in steps of 0.02 to produce the plots seen in Figure 1.

Sugihara *et al.* consider convergence to be critically important for determining CCM causality, and note that it is "a key property that distinguishes causation from simple correlation" [1]. Figure 1 shows plots created with several different library lengths to illustrate the convergence of Δ for this example. Typically, for convenience, the (approximately) converged CCM correlation values will be reported and proof of convergence will be implied, rather than shown.
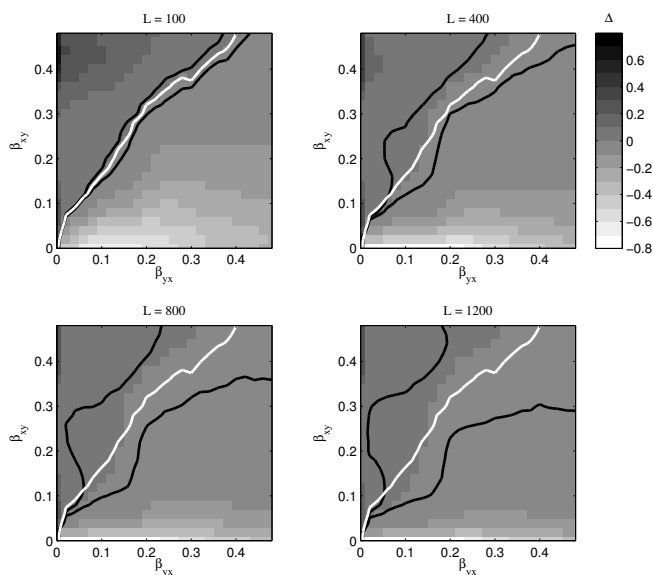
FIG. 1: The dependence of Eqn. 1 on β_{xy} and β_{yx}. The white line is a contour for $\Delta = 0$; Black lines are contours for $\Delta = \pm 0.01$.

The idea is that $\beta_{xy} > \beta_{yx}$ intuitively implies Y "drives" X more than X "drives" Y. Stated more formally, $\beta_{xy} > \beta_{yx} \Rightarrow \Delta > 0$, which is reported as "$Y$ CCM causes X". Likewise, $\beta_{xy} < \beta_{yx}$ implies X CCM causes Y and $\beta_{xy} = \beta_{yx}$ implies no CCM causality in the system. It will be shown below that CCM causality is not necessarily related to causality as it is typically understood in physics.

III. SIMPLE EXAMPLE SYSTEMS

The usefulness of the CCM algorithm in identifying drivers among sets of time series can be explored by using simple example systems. Each of the following examples intuitively supports the conclusion that X drives Y, and CCM analysis (with $E = 3$ and $\tau = 1$) yields values of Δ that support conclusions that do not agree with intuition for all parameter choices.

A. Linear Example

Consider the linear example dynamical system of

$$X_t = \sin(t) \tag{4}$$
$$Y_t = AX_{t-1} + B\eta_t, \tag{5}$$

with $A, B \in \mathbb{R} \geq 0$ and $\eta_t \sim \mathcal{N}(0, 1)$. Specifically, consider $A, B \in [0, 10]$ in increments of 0.1. Figure 2 shows Δ for this example given a library length of $L = 2000$.
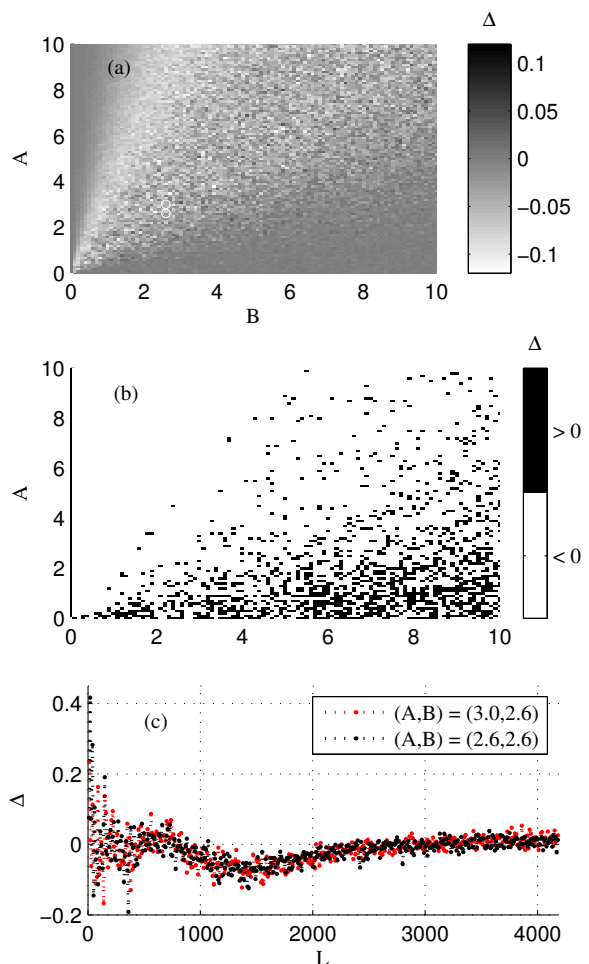

FIG. 2: Dependence of the sign of Δ from Eqn. 1 on A and B. (a) Δ is calculated as described in the text; (b) Two-color version of (a) where black indicates $\Delta > 0$, i.e., Y CCM causes X, and white indicates $\Delta < 0$, i.e., X CCM causes Y; (c) The dependence of Δ on library length L is shown for points $(A, B) = (2.6, 2.6)$ and $(A, B) = (3.0, 2.6)$ marked in (a) with white circles.

Figure 2 (c) shows (for $(A, B) = (2.6, 2.6)$ and $(A, B) = (3.0, 2.6)$) that Δ is more negative at shorter library lengths but converges to a point near zero as the library length is increased. The convergence of CCM correlations is emphasized [1], so the seemingly counter-intuitive behavior of Δ (and C_{XY} and C_{YX}) in Figure 2 implies that the CCM correlations may not be a reliable measure of "driving" (at least not the intuitive definition) for this simple linear example system.

The expected conclusion that X drives Y, corresponding to X CCM causes Y requires $\Delta < 0$. But, it can be seen from Figure 2 (b), the sign of Δ depends on A and B. Given that the intuitive conclusion of X drives Y in Eqn. 4 does not depend on A and B, it would seem that Δ does not reliably reflect the intuitive conclusion in this linear example system.

B. Non-Linear Example

Consider the non-linear dynamical system of

$$X_t = \sin(t) \tag{6}$$
$$Y_t = AX_{t-1}(1 - BX_{t-1}) + C\eta_t, \tag{7}$$

with $A, B, C \in \mathbb{R} \geq 0$ and $\eta_t \sim \mathcal{N}(0,1)$. Specifically, consider $A, B, C \in [0, 5]$ in increments of 0.5. Figure 3 shows Δ for specific values of C for a library length of $L = 2000$. As in the previous example, the expectation
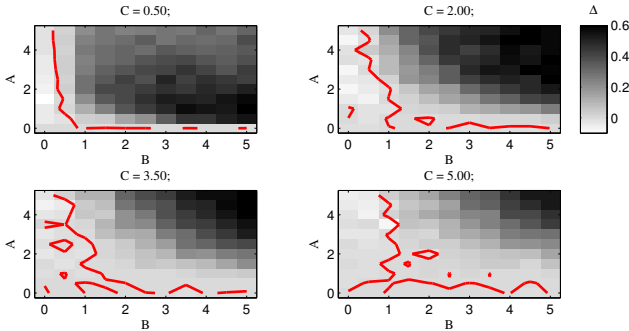
FIG. 3: The sign of Δ, and thus the CCM causality, depends on A, B, and C. Contour lines indicate where $\Delta = 0$.

for this system is that Δ should be negative, independent of the parameters A, B, and C. However, it can be seen from the plots that the sign of Δ can depend on all three parameters. Thus, this simple non-linear example leads to a conclusion similar to the previous linear example; i.e., Δ does not reliably reflect intuitive notions of driving.

C. RL Circuit Example

Both of the previous examples included a noise term, η_t. The failure of CCM analysis to give expected con-

clusions about the drivers in the previous examples may be due to a limitation of the algorithm with respect to noise. This can be investigated by considering a system without noise. Consider a series circuit containing a resistor, inductor, and time varying voltage source related by

$$\frac{dI}{dt} = \frac{V(t)}{L} - \frac{R}{L}I, \tag{8}$$

where I is the current at time t, $V(t) = \sin(\Omega t)$ is the voltage at time t, R is the resistance, and L is the inductance. Eqn. 8 was solved using the *ode45* integration function in MATLAB. The time series $V(t)$ is created by defining values at fixed points and using linear interpolation to find the time steps required by the ODE solver.

Consider the situation where $L = 10$ Henries and $R = 5$ Ohms are constant. Physical intuition is that V drives I, and so we expect to find that V CCM causes I (i.e., $C_{VI} > C_{IV}$ or $\Delta = C_{VI} - C_{IV} > 0$).

Consider evaluating the CCM correlations C_{VI} and C_{IV} for each $\Omega \in [0.01, 2.0]$ in steps of 0.01. The CCM correlations are found using $E = 2$ and $\tau = 1$ and are used to calculate $\Delta = C_{VI} - C_{IV}$, which is plotted in Figure 4. Δ does not consistently agree with intuition in this
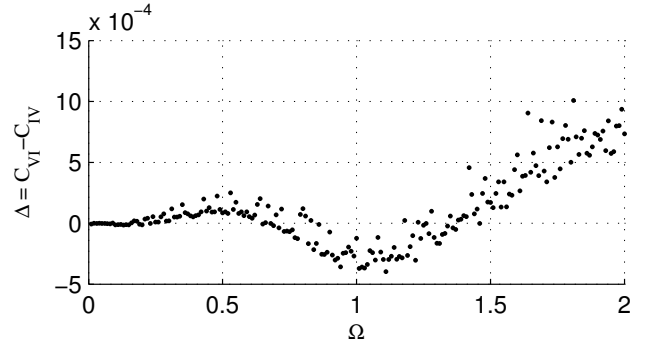

FIG. 4: Δ dependence on Ω for a series RL circuit with a sinusoidal voltage source. The sign of Δ implies V CCM causes I and I CCM causes V depending on the voltage frequency.

example either. Changing the embedding dimension, E, used to calculate Δ leads to plots that are different than Figure 4, but in all of the cases tested (i.e., $E = 2, 3, 4$), the sign of Δ changes over the domain $\Omega \in [0.1, 1.5]$.

The resistance and inductance of the circuit are fixed and the voltage is varied from 1×10^{-2} to 2.0 volts in discrete steps of 0.01 volts as described by Eqn. 8. Physically changing the voltage and witnessing a resulting change in the current is enough to convince most people that the voltage "drives" the current. Rigorous statistical hypothesis testing can be performed to strengthen the confidence in such a conclusion. Yet, from Figure 4, the voltage does not consistently "CCM cause" the current as Ω is changed.

It may be argued that the relatively small values (as compared to the previous examples) of Δ plotted in Fig-

ure 4 indicate that the correct conclusion should be either (1) there is no CCM causality in the system or (2) CCM causality is not applicable to this system. However, conclusion (1) conflicts with the intuitive notion of an RL circuit as a strongly driven system and conclusion (2) conflicts with identifying CCM causality as a general qualifier of "driving" in dynamical systems.

IV. PAIRWISE ASYMMETRIC INFERENCE

Consider the example system of Eqn. 2 with $r_y = r_y = 3.7$, $X_0 = 0.2$, $Y_0 = 0.4$, $\beta_{xy} = 0$, and $\beta_{yx} = 0.32$. These parameters correspond to Figure 3C and D of [1] (with $E = 2$, $\tau = 1$, and $L = 1000$). Plots of the correlation between X and $X|\tilde{\mathbf{Y}}$ (i.e., X estimated using the weights found from the shadow manifold of Y), as well as, Y and $Y|\tilde{\mathbf{X}}$ are shown in Figure 5. This leads to $\Delta =$
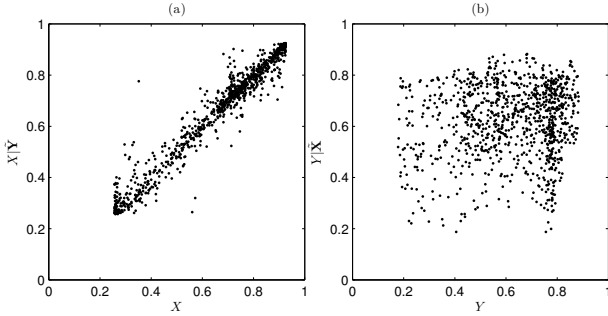
FIG. 5: Correlation plots between a given time series and its convergent cross-mapped estimate. (a) Reproduction of Figures 3C from [1]. (b) Reproduction of Figures 3D from [1].

$C_{YX} - C_{XY} \approx 0.11 - 0.97 = -0.86 < 0$, which implies X CCM causes Y. This result agrees with intuition because $\beta_{xy} = 0 < \beta_{yx} = 0.32$.

The correlations shown in Figure 5 are not the only correlations that can be tested. Consider, for example, the correlation between X and the corresponding $X|\mathbf{X}$, which is estimated using weights found from the shadow manifold of X itself. The time series X may also be estimated using a multivariate shadow manifold consisting of points from both X and Y [14]. For example, an $E+1$ dimensional point in the a multivariate shadow manifold constructed using both X and Y may be defined as $\tilde{\mathbf{X}}_t = (X_t, X_{t-\tau}, X_{t-2\tau}, \ldots, X_{t-(E-1)\tau}, Y_t)$. An estimate of X using weights from a shadow manifold using this specific construction will be referred to as $X|(XY)$ and the correlation between this estimate and the original time series will be labeled $C_{X(XY)}$. See Figure 6.

A difference in CCM correlations similar to Δ can be defined using the multivariate embedding. Consider $\Delta' = C_{Y(YX)} - C_{X(XY)}$. It might be argued, in close parallel to the arguments given in [1] for Δ, that an intuitive definition of "driving" might be captured by the sign of
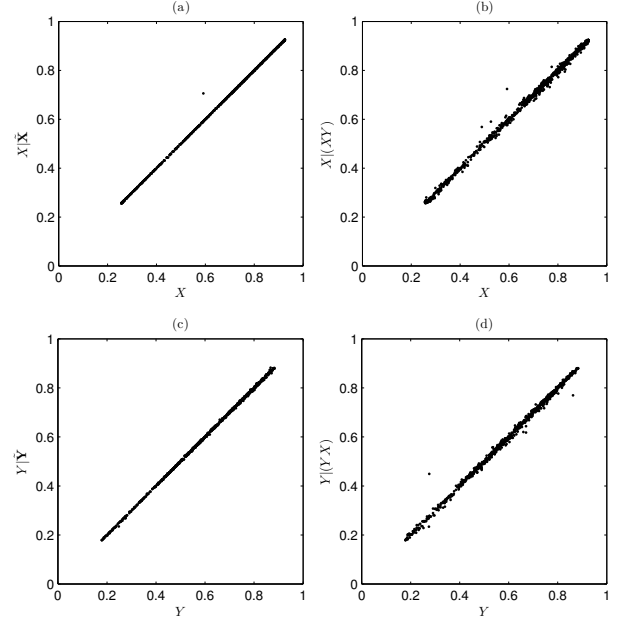

FIG. 6: Stronger correlations, as compared to Figure 5, can be seen between a time series and its estimate when the shadow manifold includes points from the time series it is estimating.

Δ'. For example, if $\Delta' < 0$, then the single time step of Y added to the delay vectors constructed from X create stronger estimators of X than the single time step of X added to the delay vectors constructed from Y do for Y. Thus, it might be argued, that Y contains more "information" about X, which leads to the conclusion X drives Y. The example system and parameters (including E, τ, and L) described at the beginning of this section leads to $\Delta' \approx -3 \times 10^{-4}$ which agrees with the previously discussed conclusions of "X CCM causes Y" and "X drives Y". Using the multivariate embedding described above to explore "driving" relationships between pairs of time series will be referred to as *pairwise asymmetric inference* or *PAI*.

Consider a comparison of PAI and CCM given the linear example system from above, i.e., Eqn. 4. Figure 7 shows Δ' as a function of A and B using of the same E, τ, L, and step sizes as was used to produce Figure 2. $\Delta' < 0 \; \forall A, B$ in the domains shown in the figure. Thus, the sign of Δ' is in agreement with an intuitive notion of driving more consistently than Δ. Δ' is significantly smaller than Δ, which is expected since the correlation of X and Y with their "self estimation" counterparts of $X|\tilde{\mathbf{X}}$ and $Y|\tilde{\mathbf{Y}}$ are initially very high, even without the multivariate additions. But, if the concept of driving is determined solely on the sign of Δ', then, at least for the simple linear example presented here, PAI is a consistent qualifier of "driving".

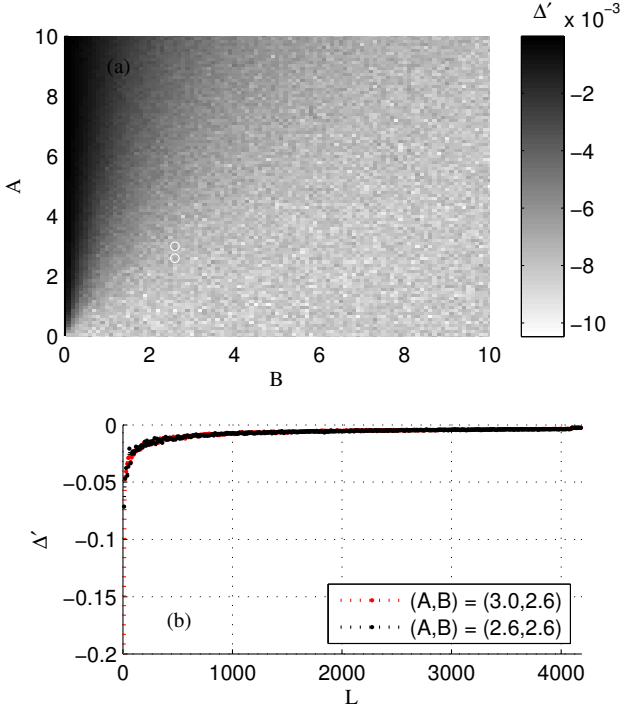Reproducing Figure 2 (c) using PAI shows an apparent reduction in some of the erratic behavior seen in CCM.

X PAI drives Y? Such questions are still open. It may be argued that a different measure may be more suitable, such as $\Delta'' = |C_{Y(YX)} - C_{YY}| - |C_{X(XY)} - C_{XX}|$. For this example, $\Delta'' \approx 3.9 \times 10^{-4}$, which does not agree with intuition, despite the agreement of both Δ and Δ'. There are still many open questions in the study of driving relationships among time series sets using state space methods.

Finally, care should be taken in any discussion of causality and especially in discussions of time series causality. We have made many statements about failure to agree with "intuition" in this paper. While some authors argue that any discussion of causality will necessarily involve appeals to intuition [27], the possibility of intuition failing cannot be ignored completely.

Consider the RL circuit example of Section III C. The intuitive definition of causality was motivated by an example of the experimenter physically manipulating a voltage source to create the V and I times series. Suppose instead that two such experiments where conducted in isolation: one with an experimenter, Alice, physically manipulating a voltage source and measuring the current to create the V and I time series (call this set **VI**), and another, different experiment with an experimenter, Bob, physically manipulating a current source and measuring the voltage to create the V and I time series (call this set **IV**). Both **VI** and **IV** are handed to a third party, Charlie, who has no *a priori* knowledge of how the time series are created.

Intuition for Alice is V causes I and she believes **VI** supports that conclusion. Likewise, Bob believes **IV** supports his intuition that I causes V. Charlie, however, must rely on time series analysis alone to determine the causality in the system. The argument we present here is not that CCM causality is insufficient because it does not provide Charlie with a definitive answer (which it does not). Such a task is difficult and may not even be possible with time series analysis alone [27]. The main problem is that the CCM method, as it has been explored in this work, is inconsistent. Any method Charlie uses must be consistent if it is to be useful. Neither Alice nor Bob would change their causality conclusions if they changed their respective input frequencies (i.e., Ω in Eqn. 8). However, if Charlie used the CCM method, his causality conclusions would depend on the frequency of the signal controlled by Alice (as seen in Fig. 4). Thus, CCM causality would not be a *consistent* tool for Charlie. PAI was shown to give consistent results for the considered examples but does not address the ambiguity identified in this example.

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

[2] M. Kaminski, M. Ding, W. A. Truccolo, and S. L. Bressler, Biological Cybernetics **85**, 145 (2001).

[3] J.-M. Dufour and E. Renault, Econometrica , 1099 (1998).

[4] J.-M. Dufour, D. Pelletier, and É. Renault, Journal of Econometrics **132**, 337 (2006).

[5] T. J. Mosedale, D. B. Stephenson, M. Collins, and T. C. Mills, Journal of climate **19** (2006).

[6] M. Eichler and V. Didelez, arXiv preprint arXiv:1206.5246 (2012).

[7] T. Schreiber, Phys. Rev. Lett. **85**, 461 (2000).

[8] C. W. Granger, Econometrica: Journal of the Econometric Society , 424 (1969).

[9] L. Barnett, A. B. Barrett, and A. K. Seth, Phys. Rev. Lett. **103**, 238701 (2009).

[10] Currently, there are no published equivalence conditions for CCM to either transfer entropy or Granger causality.

[11] C. Granger, Journal of Economic Dynamics and Control **2**, 329 (1980).

[12] Y. Liu and M. T. Bahadori, University of Southern California, http://www-bcf. usc. edu/~ liu32/granger. pdf (12. 03.2013) (2012).

[13] D. L. Roberts and S. Nord, Applied Economics **17**, 135 (1985).

[14] E. R. Deyle, M. Fogarty, C.-h. Hsieh, L. Kaufman, A. D. MacCall, S. B. Munch, C. T. Perretti, H. Ye, and G. Sugihara, Proceedings of the National Academy of Sciences **110**, 6430 (2013).

[15] X. Wang, S. Piao, P. Ciais, P. Friedlingstein, R. B. My-neni, P. Cox, M. Heimann, J. Miller, S. Peng, T. Wang, *et al.*, Nature (2014).

[16] A. Bozorgmagham and S. Ross, in *SIAM (Society for Industrial and Applied Mathematics), Student Chapter at Virginia Tech, Oct 2013* (2013).

[17] J. R. Anastas, *Individual Differences on Multi-Scale Measures of Executive Function*, Ph.D. thesis, University of Connecticut (2013).

[18] G. Sugihara, B. Grenfell, R. M. May, P. Chesson, H. M. Platt, and M. Williamson, Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **330**, 235 (1990).

[19] G. Sugihara and R. M. May, Nature **344**, 734 (1990).

[20] This definition differs slightly from the definition in [1], which uses the un-squared Pearsons correlation coefficient. We use the square of this value to avoid dealing with negative correlation values. This subtle change in the definition does not affect the conclusions drawn in [1], as can be seen in our reproduction of key plots from that work: Figure 1 and Figure 5.

[21] H.-g. Ma and C.-z. Han, Frontiers of Electrical and Electronic Engineering in China **1**, 111 (2006).

[22] I. Vlachos and D. Kugiumtzis, in *Topics on Chaotic Systems: Selected Papers from Chaos 2008 International Conference* (World Scientific, 2009) p. 378.

[23] M. Small and C. Tse, Physica D: Nonlinear Phenomena **194**, 283 (2004).

[24] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, Phys. Rev. A **45**, 3403 (1992).

[25] A. L. Lloyd, Journal of Theoretical Biology **173**, 217 (1995).

[26] In terms of Δ, convergence to a value κ means $\lim_{L\to\infty}|\Delta - \kappa| = 0$. As the "estimation skill" of the CCM algorithm increases, each of the CCM correlations, C_{XY} and C_{YX}, converge to some fixed value. Thus, Δ converges to some fixed value κ. Δ depends on the library length L because the CCM algorithm depends on L. However, the complicated algorithmic dependence makes formally solving this limit difficult. As such, convergence is determined from plots, following the method used in [1].

[27] J. Pearl, *Causality: Models, Reasoning, and Inference* (Cambridge University Press, 2000).