

# 57. An algorithm based on the convergent cross mapping method for the detection of causality in uni-directionally connected chaotic systems

**Kazimieras Pukenas**

Lithuanian Sports University, Kaunas, Lithuania

**E-mail:** [kazimieras.pukenas@lsu.lt](mailto:kazimieras.pukenas@lsu.lt)

Received 13 April 2018; received in revised form 17 May 2018; accepted 29 May 2018

DOI <https://doi.org/10.21595/mme.2018.19989>



Copyright © 2018 Kazimieras Pukenas. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Abstract.** In this paper, we present some improvements to the convergent cross mapping (CCM) algorithm for detecting causality in uni-directionally connected chaotic systems. The basic concept of the CCM algorithm is that the causal influence of system  $X$  on system  $Y$  appears as mapping of the neighbouring states in the reconstructed  $d$ -dimensional manifold,  $M_y$ , to the neighbouring states in the reconstructed  $d$ -dimensional manifold,  $M_x$ , and this effect is evaluated using the correlation coefficient between the estimated and observed values of  $M_x$ . We proposed a composite indicator of causality as the ratio between the correlation coefficient and the Shannon entropy of the distribution of the residuals between the estimated and observed values of  $M_x$ . Application of the proposed approach to four master-slave Rössler and Lorenz systems and real-world data showed that the new algorithm allowed a slight increase in capability to reveal the presence and direction of couplings.

**Keywords:** causality, synchronization, unidirectional coupling, cross-mapping, phase space reconstruction, Shannon entropy.

## 1. Introduction

Quantification of the causal effects between simultaneously observed systems from the analysis of time series recordings is essential in many scientific fields, including economics, climatology, ecosystems, electrical activity of the brain, or cardiorespiratory relations. Estimating the interdependence between the observed variables provides valuable knowledge regarding the processes that generate time series [1]. The Granger causality method [2] is the most well-known and principal method for identifying directional interactions between variables from their time series, and many modifications and extensions of the Granger causality test has been developed. The Granger test focuses on determining whether one time series is useful in forecasting another. If useful, the first system causally affects the second system. The conventional Granger's causality test is based on autoregressive models and is particularly useful in stochastic linearly interconnected systems [3]. However, its appropriateness of direct application to nonlinear systems depends on the specific problem to be analyzed [4-6]. Therefore, new approaches have been proposed, including nonlinear extensions of Granger's causality [4], transfer entropy [7], conditional mutual information [8] and measures evaluating distances of conditioned neighbors in reconstructed state spaces [9-13], to name a few. In 2012, Sugihara et al. [5] introduced another method based on state space reconstruction. The method was called convergent cross-mapping (CCM) and involves evaluating distances between conditioned neighbours in reconstructed state spaces and tests for causation between the driver ("master") system  $X$  and the driven ("slave") system  $Y$ , by measuring the extent to which the historical states of reconstructed state space,  $M_y$ , can reliably estimate states of reconstructed state space,  $M_x$ . Similarity between points,  $X_n$ , and estimates,  $\hat{X}_n$ , is evaluated by assessing their correlation. To infer the direction of coupling, the opposite direction must also be assessed. Then, the direction of coupling is inferred based on asymmetries emerging from the calculations of the two possible causal directions.

In this paper, we propose a simple improvement to the CCM method by introducing a composite parameter for inferring the direction of causation composed of two components: (1) the correlation coefficient between the cross-mapped and observed values of the reconstructed state space, and (2) Shannon entropy of the distribution of the residuals between the cross-mapped and observed values of the reconstructed state space.

## 2. Convergent cross mapping

CCM [5, 14] is a method for detecting causality between two systems represented by time series. CCM is based on Takens' embedding theorem, exploiting the geometry of attractors of coupled dynamical systems and constructs a map between mutual neighborhoods in state spaces of the coupled dynamical systems under study.

Let  $X = \{x_t\}_{t=1}^L$ , and  $Y = \{y_t\}_{t=1}^L$  be two time series of finite length  $L \in \mathbb{N}$ . To establish cross mapping from  $X$  to  $Y$ , first we reconstruct the attractor manifold  $\mathbf{M}_X$  as a set of  $E$ -dimensional vectors [15]  $\mathbf{x}_t = [x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-(E-1)\tau}]^T$  for  $t = 1 + (E - 1)\tau$  to  $t = L$ , i.e.,  $\mathbf{M}_X = \{\mathbf{x}_t\}_{t=1+(E-1)\tau}^L$ , where  $E$  is the embedding dimension,  $\tau$  is the time delay, and  $(\cdot)^T$  denotes the transpose of a real matrix. We find the  $E + 1$  nearest neighbors of  $\mathbf{x}_t$  in  $\mathbf{M}_X$  and denote their time indices (from the closest to the farthest) by  $t_1, \dots, t_{E+1}$ . These indices will be used in the construction of the cross mapping of  $\hat{y}_t$  for  $y_t, t = 1 + (E - 1)\tau, \dots, L$  by:

$$\hat{y}_t | \mathbf{M}_X = \sum_{i=1}^{E+1} w_i y_{t_i}, \quad (1)$$

where  $w_i = u_i / \sum_{j=1}^{E+1} u_j$ ,  $u_i = \exp[-d(\mathbf{x}_t, \mathbf{x}_{t_i}) / d(\mathbf{x}_t, \mathbf{x}_{t_1})]$  and  $d(A, B)$  is the Euclidean distance between vectors  $A$  and  $B$ . The cross mapping from  $Y$  to  $X$  is defined analogously. The skill of cross-map estimates is indicated by the correlation coefficient ( $\rho$ ) between  $\{y_t\}_{t=1+(E-1)\tau}^L$  and  $\{\hat{y}_t | \mathbf{M}_X\}_{t=1+(E-1)\tau}^L$  (or between  $\{x_t\}_{t=1+(E-1)\tau}^L$  and  $\{\hat{x}_t | \mathbf{M}_Y\}_{t=1+(E-1)\tau}^L$ , respectively). Sugihara et al. [5] declare in the supplement: "If  $X$  and  $Y$  are dynamically coupled, the nearest neighbors of  $\mathbf{M}_X$  should identify the time indices of corresponding nearest neighbors on  $\mathbf{M}_Y$ . As  $L$  increases, the attractor manifold fills in and the distances among the  $E + 1$  nearest neighbors shrinks. Consequently,  $\hat{y}_t | \mathbf{M}_X$  should converge to  $y_t$  and  $\hat{x}_t | \mathbf{M}_Y$  should converge to  $x_t$ . In this way, we use convergence of the nearest neighbors to test whether there is a correspondence between states on  $\mathbf{M}_X$  and states on  $\mathbf{M}_Y$ ". Cross-mapping that converges in only one direction is the criterion for unidirectional causality. One of the fundamental concepts of CCM is that when causation is unilateral,  $X \rightarrow Y$  ( $X$  drives  $Y$ ), then it is possible to estimate  $X$  from  $Y$ , but not  $Y$  from  $X$ . It follows that a high value of the correlation coefficient between  $\{x_t\}_{t=1+(E-1)\tau}^L$  and  $\{\hat{x}_t | \mathbf{M}_Y\}_{t=1+(E-1)\tau}^L$  comparing to the value of the correlation coefficient between  $\{y_t\}_{t=1+(E-1)\tau}^L$  and  $\{\hat{y}_t | \mathbf{M}_X\}_{t=1+(E-1)\tau}^L$  indicates that system  $X$  drives system  $Y$ . This runs counter to Granger's intuitive scheme. For more details on the CCM method see [5, 14, 16, 17].

## 3. Description of the algorithm

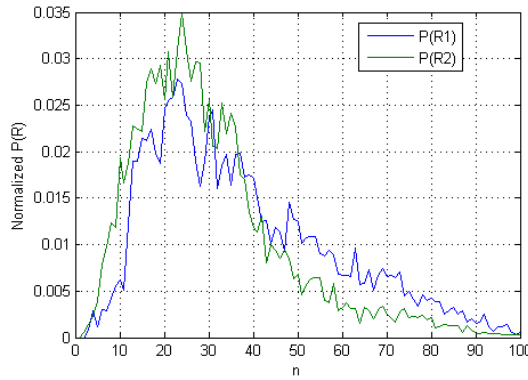
The simulation results obtained from the CCM method applied to uni-directionally coupled non-identical oscillators ( $X$  drives  $Y$ ) shows that probability distribution of the residuals taken as the Euclidean distance  $R(\mathbf{M}_X, \mathbf{M}_{\hat{X}} | \mathbf{M}_Y)$  between manifold  $\mathbf{M}_X = \{\mathbf{x}_t\}_{t=1+(E-1)\tau}^L$  and manifold  $\mathbf{M}_{\hat{X}} | \mathbf{M}_Y = \{\hat{\mathbf{x}}_t\}_{t=1+(E-1)\tau}^L$  obtained by cross mapping using a manifold  $\mathbf{M}_Y$ , and the Euclidean distance  $R(\mathbf{M}_Y, \mathbf{M}_{\hat{Y}} | \mathbf{M}_X)$  between manifold  $\mathbf{M}_Y = \{y_t\}_{t=1+(E-1)\tau}^L$  and manifold  $\mathbf{M}_{\hat{Y}} | \mathbf{M}_X = \{\hat{y}_t\}_{t=1+(E-1)\tau}^L$  obtained by cross mapping using a manifold  $\mathbf{M}_X$ , differs. More

specifically, the probability distribution of  $R(\mathbf{M}_X, \mathbf{M}_{\hat{X}}|\mathbf{M}_Y)$  has a taller, narrower shape (Fig. 1). It is therefore reasonable to combine the Shannon entropy of the probability distributions of  $R(\mathbf{M}_X, \mathbf{M}_{\hat{X}}|\mathbf{M}_Y)$  or  $R(\mathbf{M}_Y, \mathbf{M}_{\hat{Y}}|\mathbf{M}_X)$ , and the corresponding correlation coefficient to enhance the inference of the direction of instantaneous causality. Furthermore, we compute the relationship between vectors of the manifold  $\mathbf{M}_X$  and its cross-mapped estimate  $\mathbf{M}_{\hat{X}}|\mathbf{M}_Y$  and between vectors of the manifold  $\mathbf{M}_Y$  and its cross-mapped estimate  $\mathbf{M}_{\hat{Y}}|\mathbf{M}_X$ , by determining the 2D correlation coefficient, and introduce interdependence measures  $CME_{X|Y}$  and  $CME_{Y|X}$  as:

$$CME_{X|Y} = \frac{\rho_{M_X}}{En_{R_X}}, \quad (2)$$

$$CME_{Y|X} = \frac{\rho_{M_Y}}{En_{R_Y}}, \quad (3)$$

where  $\rho_{M_X}$  is the 2D correlation coefficient between vectors of the manifold  $\mathbf{M}_X$  and its cross-mapped estimate  $\mathbf{M}_{\hat{X}}|\mathbf{M}_Y$ ,  $\rho_{M_Y}$  is the 2D correlation coefficient between vectors of the manifold  $\mathbf{M}_Y$  and its cross-mapped estimate  $\mathbf{M}_{\hat{Y}}|\mathbf{M}_X$ ,  $En_{R_X}$  is the Shannon entropy of the probability distribution of  $R(\mathbf{M}_X, \mathbf{M}_{\hat{X}}|\mathbf{M}_Y)$ , and  $En_{R_Y}$  is the Shannon entropy of the probability distribution of  $R(\mathbf{M}_Y, \mathbf{M}_{\hat{Y}}|\mathbf{M}_X)$ .



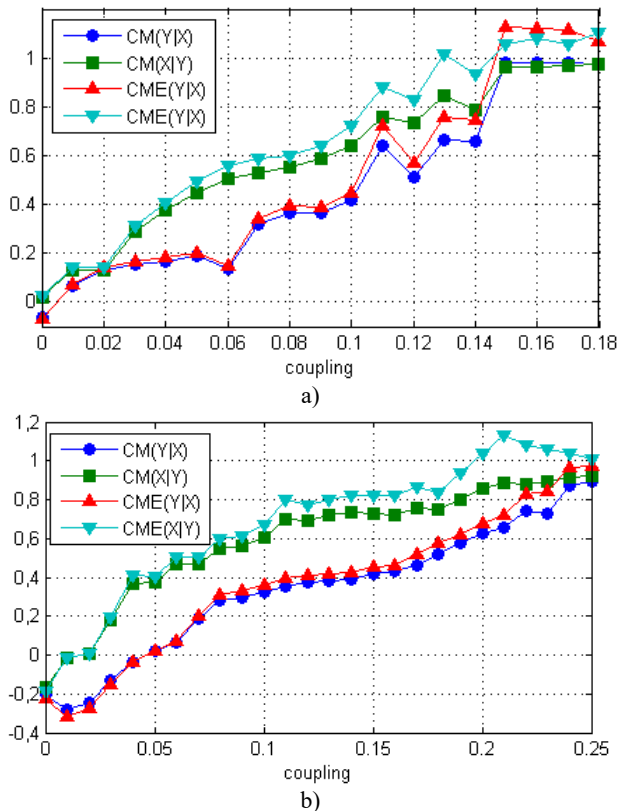
**Fig. 1.** Example of the probability distribution of the Euclidean distances  $R(\mathbf{M}_X, \mathbf{M}_{\hat{X}}|\mathbf{M}_Y)$  and  $R(\mathbf{M}_Y, \mathbf{M}_{\hat{Y}}|\mathbf{M}_X)$

## 4. Results

To demonstrate the effectiveness of the cross-mapping evaluated by the proposed interdependence measures (*CME*), we compared it with the cross-mapping evaluated using the correlation coefficient (*CM*) on four test examples of uni-directionally connected chaotic Rössler and Lorenz-type systems that were coupled with variable coupling strengths. In the examples used in this paper, the possibility of correlation instead of causation was excluded, and therefore the aspect of convergence was not used. The first and second datasets originated from the coupling of two Rössler systems [17], which were coupled via a one-way driving relationship between variables  $x_1$  of the driving system and variable  $y_1$  of the responsive system:

$$\begin{aligned} \dot{x}_1 &= -\omega_1(i)x_2 - x_3, \\ \dot{x}_2 &= \omega_1(i)x_1 + 0.15x_2, \\ \dot{x}_3 &= 0.2 + x_3(x_1 - 10), \\ \dot{y}_1 &= -\omega_2(i)y_2 - y_3 + c(x_1 - y_1), \\ \dot{y}_2 &= \omega_2(i)y_1 + 0.15y_2, \\ \dot{y}_3 &= 0.2 + y_3(y_1 - 10), \end{aligned} \quad (4)$$

where  $\omega_1(1) = 1.015$ ,  $\omega_2(1) = 0.985$  and  $\omega_1(2) = 1.075$ ,  $\omega_2(2) = 1.0$ . The coupling strength  $c$  was chosen from 0 to 0.25 with the step 0.01. The data were generated by Runge-Kutta integration with a step size of 0.1. The first 2000 data points were discarded. The total number of obtained data was 14000 and this resulted in around 60 samples per one average orbit around the attractor. The causal relationship between the two systems was calculated using 7000 time-delayed vectors of  $x_1$  and  $y_1$  with a time delay equal to 3 and embedding dimension of 7. Eight nearest neighbours were used. All data processing and analyses were performed using Matlab software (MathWorks, Natick, MA) by adopting the ideas from reference [18], among others. Consequently, we denoted the direction from  $X$  to  $Y$  as  $X|Y$  and the direction from  $Y$  to  $X$  as  $Y|X$ . Taking, for example, the measure CM: if  $X$  drives  $Y$ , the measure  $CM(X|Y)$  is expected to be higher than  $CM(Y|X)$ . As we can see from Fig. 2, both CME and CM measures show that  $X$  drives  $Y$  until the onset of synchronization. The effect of the entropy contribution in the  $Y|X$  direction is little, but is considerable in the  $X|Y$  direction, especially at moderate and high coupling strengths. As a result, the gap between  $CME(X|Y)$  and  $CME(Y|X)$  becomes larger, compared to the gap between  $CM(X|Y)$  and  $CM(Y|X)$ , i.e. the direction of causation can be determined more reliable. Furthermore, in the case of the coupled Rössler systems with  $\omega_1 = 1.075$ ,  $\omega_2 = 1.0$ , the independent-samples t-test showed that this difference between gaps across all coupling strengths is near to statistical significance ( $p = 0.0866$ ).

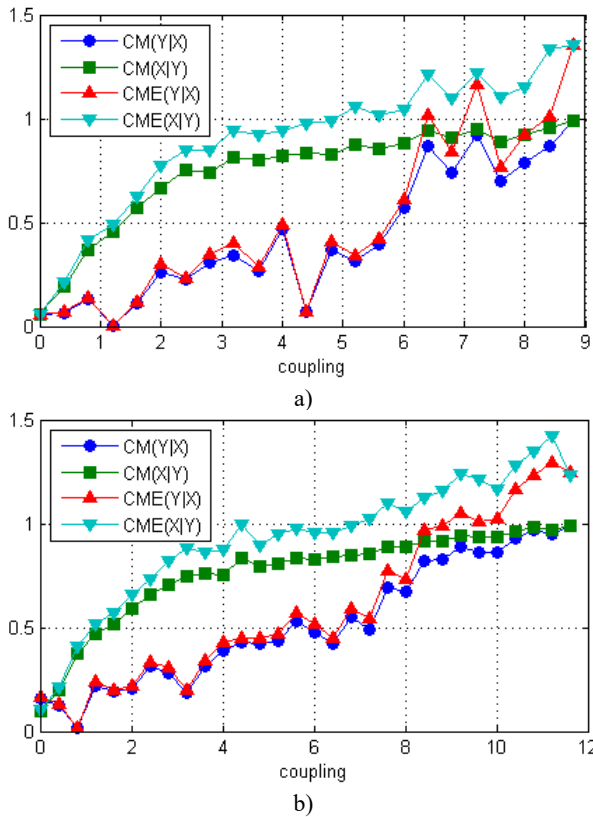


**Fig. 2.** Measures CME and CM computed for two uni-directionally coupled Rössler systems: a)  $\omega_1 = 1.015$ ,  $\omega_2 = 0.985$ , b)  $\omega_1 = 1.075$ ,  $\omega_2 = 1.0$

The third and fourth datasets originated from the coupling of two Lorenz systems [17], which are coupled via one-way driving relationship between variables  $x_1$  of the driving system and variable  $y_1$  of the responsive system:

$$\begin{aligned}
 \dot{x}_1 &= 10(-x_1 + x_2), \\
 \dot{x}_2 &= \omega_1(i)x_1 - x_2 - x_1x_3, \\
 \dot{x}_3 &= x_1x_2 - (8/3)x_3, \\
 \dot{y}_1 &= 10(-y_1 + y_2) + c(x_1 - y_1), \\
 \dot{y}_2 &= \omega_2(i)y_1 - y_2 - y_1y_3, \\
 \dot{y}_3 &= y_1y_2 - (8/3)y_3,
 \end{aligned} \tag{5}$$

where  $\omega_1(1) = 28.5$ ,  $\omega_2(1) = 27.5$  and  $\omega_1(2) = 39$ ,  $\omega_2(2) = 35$ . The coupling strength  $c$  was chosen from 0 to 11.6 with a step size 0.4. The data were generated using the Matlab solver of ordinary differential equations ode45. The first 2000 data points were discarded. The total number of obtained data was 14000 and the causal relationship between the two systems was calculated using 7000 time-delayed vectors of  $x_1$  and  $y_1$ , with a time delay equal to 3 and embedding dimension of 7. As we can see from Fig. 3, the results are similar to the coupled Rössler systems, although the effect of entropy is greater.



**Fig. 3.** Measures CME and CM computed for two uni-directionally coupled Lorenz systems:  
 a)  $\omega_1 = 28.5$ ,  $\omega_2 = 27.5$ , b)  $\omega_1 = 39$ ,  $\omega_2 = 35$

To demonstrate the performance of the proposed algorithm on real-world data, we applied the algorithm for the quantification of coupling between respiration and heart rate variability (HRV) because is well known that HRV is strongly related to respiration. For this purpose, we used data from the Fantasia database (<https://physionet.org/physiobank/database/fantasia/>). The data that are free of artifacts were acquired from 12 young subjects. The R-peaks of the electrocardiogram (ECG) signal were detected by means of a threshold method and the RR time series were obtained. In this study, we removed the mean of every RR time series either as of every respiration signal

so that they oscillate around zero. The independent-samples t-test showed a statistically significant ( $p = 0.0097$ ) difference between  $CM(X|Y)$  and  $CM(Y|X)$  and much more statistically significant ( $p = 0.000454$ ) difference between  $CME(X|Y)$  and  $CME(Y|X)$ .

## 5. Conclusions

A simulation involving examples of uni-directionally connected chaotic of Rössler and Lorenz-type systems and real-world time series shows that the proposed substitutions in the CCM method reinforce the asymmetries between the calculations of the two possible causal directions and thus improve determination of the presence and direction of coupling, and also detection of the onset of full synchronization in uni-directionally connected chaotic systems.

## References

- [1] **Papana A., Kyrtsov C., Kugiuntzis D., Diks C.** Simulation study of direct causality measures in multivariate time series. *Entropy*, Vol. 15, 2013, p. 2635-2661.
- [2] **Granger C. W. J.** Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, Vol. 37, 1969, p. 424-438.
- [3] **Krakovska A., Hanzely F.** Testing for causality in reconstructed state spaces by an optimized mixed prediction method. *Physical Review E*, Vol. 94, 2016, p. 052203.
- [4] **Chen Y., Rangarajan G., Feng J., Ding M.** Analyzing multiple nonlinear time series with extended Granger causality. *Physics Letters A*, Vol. 324, 2004, p. 26-35.
- [5] **Sugihara G., May R., Ye H., Hsieh C.-H., Deyle E., Fogarty M., Munch S.** Detecting causality in complex ecosystems. *Science*, Vol. 338, 2012, p. 496-500.
- [6] **Lusch B., Maia P. D., Kutz J. N.** Inferring connectivity in networked dynamical systems: challenges using Granger causality. *Physical Review E*, Vol. 94, 2016, p. 032220.
- [7] **Schreiber T.** Measuring information transfer. *Physical Review Letters*, Vol. 85, 2000, p. 461-464.
- [8] **Paluš M., Vejmelka M.** Directionality of coupling from bivariate time series: how to avoid false causalities and missed connections. *Physical Review E*, Vol. 75, 2007, p. 056211.
- [9] **Arnhold J., Grassberger P., Lehnertz K., Elger C. E.** A robust method for detecting interdependences: application to intracranially recorded EEG. *Physica D*, Vol. 134, Issue 4, 1999, p. 419-430.
- [10] **Quiroga R. Q., Kraskov A., Kreuz T., Grassberger P.** Performance of different synchronization measures in real data: a case study on electroencephalographic signals. *Physical Review E*, Vol. 65, 2002, p. 041903.
- [11] **Andrzejak Kraskov R. G. A., Stögbauer H., Mormann F., Kreuz T.** Bivariate surrogate techniques: Necessity, strengths, and caveats. *Physical Review E*, Vol. 68, 2003, p. 066202.
- [12] **Smirnov D. A., Andrzejak R. G.** Detection of weak directional coupling: phase-dynamics approach versus state-space approach. *Physical Review E*, Vol. 71, 2005, p. 036207.
- [13] **Chicharro D., Andrzejak R. G.** Reliable detection of directional couplings using rank statistics. *Physical Review E*, Vol. 80, 2009, p. 026217.
- [14] **Coufal D., Jakubík J., Jajcay N., Hlinka J., Krakovská A., Paluš M.** Detection of coupling delay: a problem not yet solved. *Chaos*, Vol. 27, 2017, p. 083109.
- [15] **Kantz H., Schreiber T.** *Nonlinear Time Series Analysis*. Cambridge University Press, Cambridge, 2004.
- [16] **Ye H., Deyle E. R., Gilarranz L. J., Sugihara G.** Distinguishing time-delayed causal interactions using convergent cross mapping. *Scientific Reports*, Vol. 5, 2015, p. 14750.
- [17] **Krakovská A., Jakubík J., Budáčová H., Holeciová M.** Causality Studied in Reconstructed State Space. Examples of uni-directionally connected chaotic systems. 2016, <https://arxiv.org/abs/1511.00505>.
- [18] **Jakubik J.** *Convergent Cross Mapping*. Version 1.3, <http://www.mathworks.com/matlabcentral/fileexchange/52964-convergent-cross-mapping>.