AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the accepted manuscript version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

D. Mønster, R. Fusaroli, K. Tylen, A. Roepstorff, J.F. Sherson, Causal inference from noisy time-series data — testing the convergent cross-mapping algorithm in the presence of noise and external influence, Future Generation Computer Systems (2016), <http://dx.doi.org/10.1016/j.future.2016.12.009>

Publication metadata

Title:	Causal inference from noisy time-series data — testing the convergent cross-mapping algorithm in the presence of noise and external influence
Author(s):	Dan Mønster, Riccardo Fusaroli, Kristian Tylen, Andreas Roepstorff, Jacob F. Sherson
Journal:	Future Generation Computer Systems
DOI/Link:	https://doi.org/10.1016/j.future.2016.12.009
Document version:	Accepted manuscript (post-print)
Document license:	

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

If the document is published under a Creative Commons license, this applies instead of the general rights.

Accepted Manuscript

Causal inference from noisy time-series data — testing the convergent cross-mapping algorithm in the presence of noise and external influence

Dan Mønster, Riccardo Fusaroli, Kristian Tylén, Andreas Roepstorff, Jacob F. Sherson

PII: S0167-739X(16)30742-7

DOI: <http://dx.doi.org/10.1016/j.future.2016.12.009>

Reference: FUTURE 3250

To appear in: *Future Generation Computer Systems*

Received date: 31 August 2016

Revised date: 20 November 2016

Accepted date: 6 December 2016

Please cite this article as: D. Mønster, R. Fusaroli, K. Tylén, A. Roepstorff, J.F. Sherson, Causal inference from noisy time-series data — testing the convergent cross-mapping algorithm in the presence of noise and external influence, *Future Generation Computer Systems* (2016), <http://dx.doi.org/10.1016/j.future.2016.12.009>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

- The CCM algorithm is tested on a model system of two coupled logistic maps
- Noise and an external driving signal are added to test CCM robustness
- CCM can fail even for low and intermediate coupling
- We propose low R^2 of fit to exponential convergence as indicator of CCM failure
- Controlled injection of noise can be used to improve accurate causal inference

Causal inference from noisy time-series data — testing the Convergent Cross-Mapping algorithm in the presence of noise and external influence

Dan Mønster^{a,b,*}, Riccardo Fusaroli^{c,b}, Kristian Tylén^{b,c}, Andreas Roepstorff^b,
Jacob F. Sherson^{d,e}

^a*Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark*

^b*Interacting Minds Centre, Aarhus University, Jens Chr. Skous Vej 4, 8000 Aarhus C, Denmark*

^c*Center for Semiotics, Aarhus University, Jens Chr. Skous Vej 2, 8000 Aarhus C, Denmark*

^d*Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, 8000 Aarhus C, Denmark*

^e*AU Ideas Center for Community Driven Research, Aarhus University, Ny Munkegade 120, 8000 Aarhus C, Denmark*

Abstract

Convergent Cross-Mapping (CCM) has shown high potential to perform causal inference in the absence of detailed models. This has implications for the understanding of complex information systems, as well as complex systems more generally. This article assesses the strengths and weaknesses of the CCM algorithm by varying coupling strength and noise levels in a model system consisting of two coupled logistic maps. As expected, it is found that CCM fails to accurately infer coupling strength and even causality direction in strongly coupled synchronized time-series, but surprisingly also in the presence of intermediate coupling. It is further found that the presence of noise reduces the level of cross-mapping fidelity, where the converged value of the CCM correlation decreases roughly linearly

*Corresponding author

Email address: danm@econ.au.dk (Dan Mønster)

as a function of the noise, while the convergence rate of the CCM correlation shows little sensitivity to noise. The article proposes controlled noise injections in intermediate-to-strongly coupled systems could enable more accurate causal inferences. Initial investigation of an external driving signal indicates robustness of CCM toward this potentially confounding influence. Given the inherent noisy nature of real-world systems, the findings enable a more accurate evaluation of CCM applicability and the article advances suggestions on how to overcome the method's weaknesses.

Keywords: Convergent Cross-Mapping, Causality, Logistic map, Noise, Time-series analysis, Nonlinear dynamics, Complex systems

1. Introduction

Complex information systems play a crucial role in infrastructures vital for modern societies. Prime examples are the global Internet, cellular and PSTN networks, power grids, supply networks and the global financial system. The optimal design and operation of complex information systems is an important research topic, and new types of critical infrastructures appear, such as cloud computing and the internet of things. The interaction between many coupled heterogeneous components in such systems makes it difficult to predict the actual behavior at a system-wide level. This is true in particular for predicting events that may be detrimental or even catastrophic to system performance, e.g., network failure [1], rolling blackouts [2], financial crises [3] or even multi-infrastructure failures [4]. Methods that allow us to understand or predict complex systems based on observations are therefore an important alternative approach to trying to predict their behavior simply based on a model or knowledge of their design. Complex information systems usually contain a large number (size) of components that are

not identical (heterogeneity); that interact (coupling); and that change over time (dynamics). The combination of size, heterogeneity, coupling, and dynamics means that to characterize complex information systems, time-series of many variables are needed. One of the simplest and most difficult questions one may ask in order to characterize such a data set is: what is the causal relationship between a pair of variables? In other words: does X cause Y , does Y cause X , or does causality go in both directions?

The ability to infer causality from the relation between two variables is an intensely researched area. Common approaches involve having both a model of the system being studied and a series of measurements of that same system [5]. In many cases, however, an adequate model of the system is not available, or there are several conflicting models. Complex natural, social, and technical systems are prime examples, e.g., ecosystems, brains, the climate and the complex information systems mentioned above. Data from such systems are often non-stationary and with nonlinear relationships between variables. This presents a challenge not only to traditional linear statistics based methods, but also to machine learning and big data analytics methods, although for specific applications progress has been made in detecting relevant patterns in the data (e.g., [6]). In most such cases, however, inferring whether one part has a causal influence on another part has to rely on model-free methods. Convergent Cross-Mapping (CCM) [7] is a relatively new method that promises to ‘distinguish causality from correlation’ in time series data (ibid., p. 496).

While nonlinear methods promise better results they often require more input from the user, e.g. in terms of setting parameters, making unsupervised use more difficult. Further, nonlinear methods often require more computational

resources [8]. CCM, however, only requires the specification of two parameters:
the embedding dimension and time delay, which can both be estimated with existing
algorithms [9, 10, 11, 12].

For a system with several variables, for which time series data are available, the
CCM method produces a causal network structure describing which variables are
causally connected, including the direction of causality. Like mutual information
[13], transfer entropy [14], and cross-recurrence quantification [15, 16] CCM is a
state space method relying on time-delayed embedding of the time series data in
a higher dimensional space. While CCM is not a prediction method, using CCM
to infer the causal connections between variables can be used to constrain viable
predictive models.

CCM has already been used in a wide range of different fields for different kinds
of data [17, 18, 19], and given the interest and relevance of CCM, it is important
to better understand its strengths and limitations. Accordingly, in this article the
authors will present results of an in-depth study of a simple model system, the
coupled logistic map, with particular emphasis on how the CCM results reflect
the model input, varying strength of coupling, levels of noise, and the influence
of an external variable. This article will briefly describe how CCM is applied to
the model. Then it is examined how coupling strength affects the CCM results, the
result of adding noise to the system, and finally the effect of an external driving
variable will be investigated. Although noise in real-world data is ubiquitous, the
inclusion of noise in model investigations has been largely neglected. It is found
that noise can dramatically change the strength of causal inferences. Crucially, it is
also observed that the appropriate injection of noise into the dynamics can be used
as means of inferring the relative strengths of the coupling (to the extent that the

66 system can be controlled).

67 2. Related work

68 CCM was introduced as an alternative to other methods that detect causality
69 between two time series, principally Granger causality [20].

70 Granger causality [20] was introduced to assess causality between variables
71 in both linear and non-linear systems. The method has gained wide application
72 from contexts of neuroscience to economics, and today several graphical Granger
73 algorithms are available. In a recent study Arnold, Liu and Abe [21] compared a
74 range of the algorithms on both synthetic and real-world data sets, and characterized
75 their performance in terms of ability to recover the correct causal graph, and
76 computational complexity. While Granger causality performs well for certain types
77 of coupled systems, it rests on the assumption of easily separable variables with
78 little or no feedback. As a consequence, it fails to correctly detect the direction
79 of causality in a range of naturally occurring biological, ecological, and social
80 systems that are rather characterized by weak to moderately coupled dynamics [7].

81 In contrast, CCM is primarily suited for weakly coupled components of non-
82 linear dynamic systems, while it works less well for cases of strong coupling.
83 Accordingly, the two methods have slightly diverging notions of causality, and the
84 statement ‘X causes Y’ would more accurately be phrased as ‘X Granger causes Y’
85 or ‘X CCM causes Y’. Sugihara et al. [7] refer to the type of causality captured
86 by CCM as *dynamic causation*, reminiscent of what Lakoff [22] terms systemic
87 causation.

88 Since its first appearance in 2012, a number of modifications and improvements
89 of the CCM algorithm has been suggested. Cummins, Gedeon, and Spendlove

[23] have elaborated the mathematical foundation for CCM to explain in more detail why CCM fails in the limit of strong coupling. They also show that an alternative approach building on the results by Pecora, Carroll, and Heagy [24] can be used, and is more consistent with the mathematical framework developed. Along similar lines, McCracken and Weigel [19] noted that CCM results are not always consistent with theoretical intuitions, and they extended CCM to pairwise asymmetric inference (PAI), which they demonstrated to give results that are in better accordance with the intuitively expected outcomes for several physical systems.

While the original CCM model needs quite extensive time series, Ma, Aihara, and Chen [25] developed cross-map smoothness (CMS)—a method related to CCM—which has the advantage of requiring fewer points in the time series. They also compared CMS to CCM using data from mathematical models, and from a gene regulatory networks, and found that CMS can be used for time-series of length $n \approx O(10)$, whereas CCM requires time-series of length $n \approx O(10^3)$ to yield reliable results.

More recently, Clark et al. [26] developed the CCM technique into multispatial CCM (MCCM) to allow for the analysis of even very short time series ($n < 10$) in the presence of multiple time series, sampled from different spatial locations.

The fact that CCM and related techniques lend themselves to the investigation of also short time-series is important for many applications, however the focus in this article will mainly be on applications with large data sets, so the low n limit, is not a concern here.

Another concern when analyzing complex, naturalistic time series data is the methods' tolerance to noise. Jiang et al. [27] recently investigated the influence

of noise on the CCM model, and found that if asymmetric noise is injected in the observed time-series it can lead to increased confidence in the detected direction of causality. While their study is related to the present investigation, these authors consider noise injected after data gathering as part of the analysis process. In contrast, the focus in the present study will be on considering noise that comes from other parts of the system or noise that is injected to experimentally perturb the system.

While most previous studies have investigated CCM under the ideal, but unrealistic, conditions without noise, the study presented in this article focuses on how the CCM algorithm performs in the presence of noisy data. Another understudied challenge to CCM is the influence of an external variable, which can potentially obscure or reverse the causal results obtained with the CCM algorithm. The present article will therefore address this point by including such an external influence. These issues will be investigated using the original CCM algorithm, while leaving the analysis of developments such as PAI, CMS and MCCM for future work.

3. Model system

The logistic map has long been a model system of nonlinear dynamics displaying regular periodic behavior as well as deterministic chaos [28]. A system of two coupled logistic maps has been used as a simple model of chemical reaction dynamics [29] and population dynamics [30]. This makes it an ideal model system for testing CCM. Sugihara et al. [7] is followed to study two logistic maps coupled through linear terms

$$\begin{aligned}x_{t+1} &= x_t(r_x(1 - x_t) - \beta_{xy}y_t) \\ y_{t+1} &= y_t(r_y(1 - y_t) - \beta_{yx}x_t)\end{aligned}\tag{1}$$

138 The two variables x and y have a nonlinear dependence on their own past values
 139 parameterized by the growth rates r_x, r_y , and are coupled to each other through
 140 linear terms with coupling constants that parameterize the strength of the coupling
 141 from x to y (β_{yx}) and from y to x (β_{xy}). Times series for x and y are shown in
 142 Figure 1 for a particular choice of the parameters in the model.

Figure 1: Plot of the time series of x (shown in blue) and y (shown in red) after the initial 300 time steps. Note that $\beta_{xy} = 0$, so there is no coupling from y to x and the causality is therefore unidirectional from x to y ($\beta_{yx} = 0.05$). Both $r_x = 3.65$ and $r_y = 3.77$ are chosen to be in a chaotic regime of the logistic map.

143 A single logistic map has well-known domains of periodic behavior controlled

by the growth rate r . As r becomes larger, increasingly frequent period-doublings occur, which ultimately give way to chaotic behavior at $r \approx 3.57$. This well-known phase diagram or bifurcation diagram is shown in the bottom panel of Figure 2, representing the x variable. The top panel shows the phase diagram for another logistic map representing the y variable. In order to illustrate the effect of coupling between two logistic maps the growth rate for the y variable r_y is related to r_x by the equation $r_x + r_y = 6.5$. This is the equation for a line in the (r_x, r_y) -plane with a slope of -1 . The middle panel in Figure 2 illustrates the phase diagram of y for a family of two coupled logistic maps along this line in the (r_x, r_y) -plane as described by Eq. 1 with $\beta_{xy} = 0$ and $\beta_{yx} = 0.2$.

The middle and lower panels show the effect of unidirectional coupling between two logistic maps. It is evident that the dynamics of x leaves an imprint on y as indicated by the fixed points. For low r_y (high r) where the isolated system (top panel) is stationary, the effect of coupling from the chaotic dynamics of x induces chaotic dynamics in y , albeit with a much smaller amplitude. For high r_y (low r) where the isolated system is chaotic, the coupling from x which is in the stationary domain has the effect of stabilizing the periodic phases of y which persist to higher values of r_y than in the isolated system.

3.1. Noise

Real-world systems present varying degrees of environmental and measurement noise and the application of the method to such systems thus critically depends on the method's ability to cope with noise. The model with added noise terms is therefore also considered:

$$\begin{aligned} x_{t+1} &= x_t(r_x(1 - x_t) - \beta_{xy}y_t) + \epsilon_{x,t} \\ y_{t+1} &= y_t(r_y(1 - y_t) - \beta_{yx}x_t) + \epsilon_{y,t} \end{aligned} \quad (2)$$

Figure 2: Phase diagram for the coupled logistic map in Eq. 1 for $\beta_{xy} = 0$. The growth rates r_x, r_y are parameterized by the variable r on the horizontal axis as $r_x = r$ and $r_y = 6.5 - r$. For each value of r the plots represent a different pair of parameters r_x, r_y . The bottom panel shows the fixed points x^* as a function of the growth rate, and the top panel ($\beta_{yx} = 0$) shows the fixed points y^* . The center panel shows the fixed points for $\beta_{yx} = 0.2$.

Here $\epsilon_{x,t}$ and $\epsilon_{y,t}$ are the noise terms on x and y which are modeled as stochastic variables sampled from normal distributions $\mathcal{N}(0, \sigma^2)$ with zero mean and standard

deviation σ . The standard deviation σ will be referred to as the noise level. Note that noise in either variable in Eq. 2 propagates to later values of the variable and to the other variable through the coupling terms. Hence, the noise terms in this model represent perturbations from the environment rather than noise from the measurement process as in [27].

3.2. External driving signal

A fundamental problem in inferring causality between two observed variables is whether an observed connection between the variables is due to a third unobserved variable that has a causal influence on both variables. In ecology this is called the Moran effect, since Moran [31] was the first to state that synchrony between two populations could be due to external environmental influence, and that the correlation between population sizes would equal the correlation between the respective environments. This equality between the correlation coefficients was based on assumptions of linear dynamics, and it was later shown that the Moran effect is weaker in nonlinear systems [32]. The Moran effect has been modeled in nonlinear population dynamics using both stochastic [33] and periodic [34] signals to model the effect of the environment. An external driving signal is introduced as a modification of the growth rates:

$$\begin{aligned} x_{t+1} &= x_t [(r_x + \eta_x H_t)(1 - x_t) - \beta_{xy} y_t] + \epsilon_{x,t} \\ y_{t+1} &= y_t [(r_y + \eta_y H_t)(1 - y_t) - \beta_{yx} x_t] + \epsilon_{y,t} \end{aligned} \quad (3)$$

The external driver H_t modulates the growth rate parameters r_x and r_y with strengths η_x and η_y , respectively. This has been used in ecology to model seasonal changes in environmental conditions, and is a specific example of the Moran effect.

192 Summers, Cranford and Healy [34] are followed in choosing

$$193 \quad H_t = \cos(2\pi t/p), \quad (4)$$

194 where $p = 2, 3, \dots$ is the period. The full model with an external driving signal and
 195 noise given by Eq. 3 is depicted graphically in Figure 3, where arrows indicate
 196 direction of causality and the parameters controlling the strength of each causal
 197 influence is shown. In the special case where $\eta_x = \eta_y = 0$ the model reduces to
 198 Eq. 2, and if, further, $\sigma_x = \sigma_y = 0$ the original model, without noise and external
 199 driving signal is obtained, given by Eq. 1.

Figure 3: Graphical depiction of the model in Eq. 3. The direction of the arrows indicate the direction of causal influence between the variables in the model, and the annotation next to the arrows indicate the parameter that controls the strength of the interaction.

200 Example time series for x_t , y_t , and H_t are shown in Figure 4. It is clear that both

201 x and y are modulated by H , with the effect that x and y show larger fluctuations
 202 when H is close to its maximum value, and smaller fluctuations when H is near its
 203 minimum.

204 Note that the addition of periodic external forcing has been shown to drastically
 205 alter the dynamics of a single logistic map system [34], so with the added influence
 206 from H_t , the phase diagram in Figure 2 is no longer a valid representation of the
 207 system.

208 4. Method

209 Convergent cross-mapping is a state space method that relies on Takens' theorem
 210 [35] to reconstruct the underlying dynamics of a system in a model-free fashion, by
 211 using time-delayed embedding to reconstruct its attractor landscape (see, e.g., [9]).
 212 The model system in Eq. 1 can be described by its attractor, that is, the trajectory
 213 consisting of consecutive points in two-dimensional Euclidean space given by
 214 the Cartesian coordinates $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ resulting from the dynamics
 215 described by Eq. 1. The first 300 points are discarded to avoid transient behavior of
 216 the model from affecting the results. According to Takens' theorem the attractor
 217 can approximately be reconstructed from one of the variables x or y alone using
 218 time-delayed embedding of the points in one of the time series, say x , where each
 219 point in E -dimensional space is given by $\mathbf{x}_t = (x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-(E-1)\tau})$. The
 220 embedding thus depends on two parameters: the time delay τ and the dimension E
 221 of the space in which the reconstructed attractor is embedded. To reconstruct an
 222 attractor in a phase space of dimension E using time delayed embedding with a
 223 delay of τ from a time series of length n a total number of points $N = n - (E - 1)\tau$
 224 is obtained.

225 Since there are only two independent variables in the model system it is evident
226 that $E = 2$, but estimating the embedding dimension from the times series obtained
227 from Eq. 1 using the false nearest neighbor method [9] gives the same result.
228 The time delay for the embedding was set to $\tau = 1$ based on the average mutual
229 information criterion [10]. The reconstructed attractors, referred to as *shadow*
230 *manifolds* by Sugihara et al. [7], are shown in Figure 5.

231 The two key ingredients of the CCM method are the concept of *cross-mapping*
232 and the *convergence* property that are explained in the following sections.

233 4.1. Cross-mapping

234 The time series data from each variable—in this case x and y —can be used
235 to construct shadow manifolds— M_x and M_y —that are approximations to the true
236 attractor. Given the non-separability of the system, Taken’s theorem demonstrates
237 that when two different variables represent different parts of the same dynamical
238 system their shadow manifolds are diffeomorphic to the true attractor and therefore
239 to each other. Intuitively the two variables are connected because they are part
240 of the same dynamical system as evidenced by the fact that they both represent
241 a dimension in the state space. So, if x has a causal influence on the dynamics
242 of y then x will influence the dynamics of y . This ‘imprint’ of x on y means that
243 knowledge of the shadow manifold M_y obtained from the time series of y can be
244 used to estimate values of x . This estimate is called the cross-map and is denoted
245 $\hat{x}|M_y$.

246 To find the cross-mapped estimate $\hat{x}_t|M_y$ of x_t start by identifying the corre-
247 sponding point $\mathbf{y}_t$ in M_y . Since M_y is diffeomorphic to M_x , a small region around
248 $\mathbf{y}_t$ will map to a small region around $\mathbf{x}_t$ and this can be used to estimate x_t . To form
249 a bounding simplex around $\mathbf{y}_t$ at least $E + 1$ points are needed [36], so the $E + 1$

250 nearest neighbors of $\mathbf{y}_t$ in M_y are found. Sorted from the closest to the farthest
 251 point from $\mathbf{y}_t$ these are called $\mathbf{y}_{t_1}, \mathbf{y}_{t_2}, \dots, \mathbf{y}_{t_{E+1}}$. Next the points in the time series
 252 for x at the corresponding times, i.e., $x_{t_1}, x_{t_2}, \dots, x_{t_{E+1}}$ are used to estimate x_t as

$$253 \quad \hat{x}_t | M_y = \sum_{i=1}^{E+1} w_i x_{t_i} \quad (5)$$

254 The weights w_i are exponentially weighted with the Euclidean distance between $\mathbf{y}_t$
 255 and the nearest neighbor points:

$$256 \quad w_i = u_i / \sum_{j=1}^{E+1} u_j, \quad u_i = \exp \left(- \frac{\|\mathbf{y}_t - \mathbf{y}_{t_i}\|}{\|\mathbf{y}_t - \mathbf{y}_{t_1}\|} \right) \quad (6)$$

257 where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^E$.

258 Estimating one point alone is not sufficient to show how well $\hat{x}_t | M_y$ estimates
 259 the true value x_t . A library consisting of L points from M_y is therefore used to
 260 provide estimates of L points in the time series for x . The Pearson correlation
 261 coefficient $\rho_{x\hat{x}}$ between the L true values from x and the L cross-mapped estimates
 262 is an indicator of how much the dynamics of x influences the dynamics of y . Scatter
 263 plots of observed values and cross-mapped estimates of x and y are shown in
 264 Figure 6.

265 The results indicate that the cross-mapped estimates of x obtained from M_y
 266 ($\rho_{x\hat{x}} \approx 0.6$) are better than the cross-mapped estimates of y obtained from M_x
 267 ($\rho_{y\hat{y}} \approx 0.2$). It is therefore tempting to conclude that x causes y , but, although
 268 correct in this case, such a conclusion is potentially misleading. Indeed, A slight
 269 variation in the parameters, can produce results that would lead to the opposite
 270 conclusion. The additional criterion of convergence is crucial to correctly infer the
 271 direction and relative strength of causality.

272 4.1.1. Implementation

273 Cross-mappings were performed with XMAP [37] which was developed in
 274 MATLAB and validated by reproducing the results by Sugihara et al. [7] and by
 275 comparing with an independently developed algorithm [38]. A small regularizing
 276 term δ is added to the denominator in the argument to the exponential function in
 277 Eq. 6 in order to avoid floating point overflow in cases where $\mathbf{y}_{t_1} = \mathbf{y}_t$.

278 The MATLAB function `xmap( X, Y, MX, MY, E, tau, L, sampling)` is
 279 called with vectors X and Y of length n holding the time series variables x and y ,
 280 matrices MX and MY of dimension $N = n - (E - 1)\tau \times E$ holding the reconstructed
 281 attractors M_x and M_y and scalar variables E and τ to specify the parameters E
 282 and τ used for phase-space embedding. Additional optional parameters specify the
 283 library size L and the sampling method (see section 4.3). Phase-space embedding
 284 can be performed by the user, or with the function `psembed` supplied as part of
 285 XMAP. All variables are held in memory and hence performance is constrained
 286 by CPU speed, memory, and cache size. For very large data sizes CPU cache
 287 optimization and out-of-memory methods will be needed, but with the low number
 288 of variables and relatively short time-series used here, this has not been necessary.
 289 Parallelization or use of GPU computations could also be employed to speed up
 290 computations for large data sets.

291 4.2. Convergence

292 The convergence properties of the correlation between observed values and
 293 cross-mapped estimates as a function of the library length L is the second key
 294 ingredient in CCM. If x causes y then the estimate of x obtained from M_y should
 295 improve as the number of points L sampled from M_y becomes larger, since the
 296 library of samples will become a better and better representation of the attractor,

297 and the nearest neighbor points will be closer and closer to $\mathbf{y}_t$.

298 This convergence phenomenon is illustrated in Figure 7 where correlations
299 derived from the data shown in Figure 6 have been computed for increasing values
300 of the library length. Figure 7 shows that $\rho_{x\hat{x}}$ converges toward a much higher
301 value than $\rho_{y\hat{y}}$ does. This is evidence that M_y enables much better estimates for x
302 than M_x does for y . Note that cross-mapped estimation skill is from y to x when x
303 causes y .

304 4.2.1. Fitting the convergence

305 Rather than relying only on the value of the correlation coefficient for the
306 largest obtained value of L , the computed values of ρ are fitted as a function of L
307 to the following function

$$308 \quad \rho(L) = \alpha e^{-\gamma L} + \rho_{\infty}. \quad (7)$$

309 This function, containing three constants α , γ and ρ_{∞} , is in accordance with the
310 computed values of the correlation coefficient. It can be interpreted as convergence
311 toward the value ρ_{∞} as $L \rightarrow \infty$ where the speed of convergence is given by the
312 constant γ . The third constant α is necessary to obtain a good fit, but does not
313 provide additional interpretive value.

314 The solid lines in Figure 7 represent the fitted correlation coefficients to the
315 function in Eq. 7 using the same parameters as in Figure 1. The fitted value for $\rho_{x\hat{x}}$
316 is $\rho_{\infty} = 0.71 \pm 0.046$ (95% CI), and the fitted value for $\rho_{y\hat{y}}$ is $\rho_{\infty} = 0.096 \pm 0.024$.

317 4.3. A note on phase-space sampling

318 As just described, to investigate the convergence of the correlation coefficients,
319 a library of L points is sampled from the reconstructed attractors. How this sample
320 is drawn from an attractor containing N points is an important implementation

321 detail. The XMAP [37] implementation of the CCM algorithm has two different
 322 sampling options, ‘linear’ that samples the first L points of the reconstructed
 323 attractor, and ‘random’ that draws points from the attractor from a list of uniform
 324 random numbers between 1 and N (without replacement).

325 The convergence of the correlation coefficients for ‘linear’ and ‘random’ sam-
 326 pling of the attractors is shown in figure 8 a and b, respectively. It is evident from
 327 the figure that the consecutive (‘linear’) sampling method can produce misleading
 328 results at small library sizes, since only a small part of the reconstructed attractors
 329 are sampled. The ‘random’ sampling method (Figure 8b) is a distinct improvement,
 330 that gives a much clearer picture of the convergence of $\rho_{x\hat{x}}$ and $\rho_{y\hat{y}}$. The main
 331 problem is that the curves are very noisy at small library sizes, which may be a
 332 problem if there is fast convergence, since this can be masked by the noise (see
 333 section and Figure 9 for an example).

334 To reduce the noise at low library length, the correlation coefficients are av-
 335 eraged over an ensemble of independent random samples at each library length.
 336 At each library size an ensemble of M independent samples is constructed, where
 337 $M = \lfloor N/L \rfloor$, and N is the number of points in the attractor. Hence, a big ensemble
 338 is used at low values of L and $M \rightarrow 1$ as $L \rightarrow N$. The smallest library size is set
 339 to $L = 6$. The result of using an ensemble average is shown in Figure 8c. This
 340 reduces the noise for low values of L and allows better estimation when fitting the
 341 convergence of $\rho_{x\hat{x}}$ and $\rho_{y\hat{y}}$ as a function of L . Note that the example in Figure 8 is
 342 a system without noise ($\sigma = 0$), which shows that the effect of ensemble averaging
 343 is not related to noisy data.

344 In Figure 7 and all subsequent figures an ensemble average is used, as described
 345 here.

5. Results

The previous section introduced CCM and demonstrated how to apply the method on the model system using a particular choice of parameter values as an example. Sugihara et al. [7] summarize their results for different values of the coupling constants β_{xy} and β_{yx} with r_x and r_y chosen uniformly from the interval $[3.6, 4]$ in Figure 3B of their manuscript. This figure indicates that CCM gets the direction of the coupling right, but also indicates that for some choices of the coupling constants the method may give results that are hard to distinguish from symmetric coupling, even when $\beta_{xy} \gg \beta_{yx}$ or vice versa.

To investigate this further, a particular choice of r_x and r_y will be examined as the coupling between the two variables is manipulated.

5.1. Effect of weak and strong coupling

To study how CCM causality estimates depend on the strength of the coupling between the two variables the intrinsic growth rates are fixed: $r_x = 3.625$ and $r_y = 3.77$. To keep things simple, unidirectional coupling is chosen, so $\beta_{xy} = 0$ as in the previous example, and only β_{yx} is varied. The correlation coefficients $\rho_{x\hat{x}}$ and $\rho_{y\hat{y}}$ are shown in Figure 9 for $\beta_{yx} = 0.1, 0.2, 0.5, 2.0$.

For $\beta_{yx} = 0.05$, a plot very similar to that displayed in Figure 7 is produced (not shown in Figure 9), which is consistent with a coupling from x to y , but no coupling from y to x .

For $\beta_{yx} = 0.1$, $\rho_{x,\hat{x}}$ is seen to converge toward $\rho_\infty = 0.89$, whereas $\rho_{y,\hat{y}}$ fluctuates around 0.36. This lack of convergence as a function of the library size is an indication that there is no coupling from y to x .

For $\beta_{yx} = 0.5$, the convergence of $\rho_{x\hat{x}}$ is toward the higher value $\rho_\infty = 0.96$,

and convergence is faster, indicating a stronger coupling. Here convergence of $\rho_{y\hat{y}}$ toward $\rho_{\infty} = 0.43$ is observed, and although the value is much smaller than for x , this could indicate a coupling from y to x . However, this cannot be the case, since $\beta_{xy} = 0$, so instead this is most likely due to the fact that x is driving y , and that the two time series are becoming more synchronized.

For an intermediate β_{yx} value of 0.2 something unexpected happens: both $\rho_{x\hat{x}}$ and $\rho_{y\hat{y}}$ quickly increase to values close to 1, already for very low values of L . Furthermore for $L \lesssim 100$, $\rho_{y\hat{y}} > \rho_{x\hat{x}}$. Naively this could be seen as an indication that there is a strong bidirectional coupling between x and y . Alternatively both signals could be driven by a common external variable—the Moran effect [31]—or one of the variables is driving the other so strongly that they synchronize [39, 40]. It is clear, however, that neither the Moran effect, nor synchronization can be the explanation here, so a different explanation must be found. This will be considered in section 5.1.1, but note here that (i) for both x and y the correlation is uniformly high, i.e., there is not much evidence of convergence, and (ii) the function in Eq. 7 is not a good fit to the data.

For $\beta_{yx} = 2$ very high correlation coefficients for both x and y are again observed, but clearly $\rho_{x\hat{x}} > \rho_{y\hat{y}}$. There is not much evidence of convergence for $\rho_{x\hat{x}}$ and $\rho_{y\hat{y}}$ fluctuates around a value close to 0.8. This is consistent with synchronization due to the strong coupling from x to y and inspection of the time series confirms this. Again, note that $\rho(L)$ from Eq. 7 is not a good fit to the data.

5.1.1. Near-periodicity and bunching

The case in the upper right panel in Figure 9 is now considered again. For $r_x = 3.625$ and $r_y = 3.77$ both logistic maps—considered in isolation—are in the chaotic phase (cf. Figure 2). But for $\beta_{yx} = 0.2$ the dynamics of y become nearly

periodic as illustrated in Figure 10.

Nonlinear dynamic systems are known to display intermittency [41], where periods of chaos are interspersed with periods of nearly periodic behavior. Intermittency, however, does not seem to be the cause here, since the phenomenon persists no matter how many points are discarded as transient before starting to sample. Instead, it seems, that y becomes partially synchronized with x because the range of x -values lies within several relatively narrow intervals. An example of cross-mapped estimates for x and y are shown in figure 11.

Near-periodicity results in $\rho_{y\hat{y}}$ becoming higher than $\rho_{x\hat{x}}$ at low L , and ‘bunching’ of the points leads to ‘fast non-convergence’, i.e., high correlations, that are not reminiscent of the convergence behavior expected in CCM, evidenced by the fact that the fits to $\rho(L)$ in Eq. 7 fail.

To see whether this is an isolated point in parameter space, resulting in pathological behavior, cross-mapped estimates $\hat{y}|M_x$ of y are calculated for a range of parameter values r_x , r_y and β_{yx} . For each value the first 1000 points in the time series are discarded and the next 400 points after that are sampled to construct the shadow manifolds. The result is shown in Figure 12.

There is a clear plateau in Figure 12 that corresponds to the ‘bunching’ observed in Figure 10, indicating that the issue is not confined to a few isolated parameter combinations. Whether similar issues affect the use of CCM for other models or when applied to empirical data is a topic for future research.

5.2. Effects of noise

Since all empirical data contain a certain level of noise it is important to study what effect noise, as included in Eq. 2, has on CCM results. Consider the simple case of unidirectional coupling from x to y , i.e., $\beta_{xy} = 0$. In this case the noise

term $\epsilon_{x,t}$ propagates directly to the variable y , and therefore does not affect the cross-mapped estimates of x based on M_y (a fact that has been verified by numerical simulations). On the other hand, noise in y is not coupled back into x and could potentially be detrimental to the reconstruction. Hence, the simplifying assumption of setting $\epsilon_{x,t} = 0$ is made. This leaves r_x, r_y, β_{yx} and the noise level σ_y as free parameters of the model in Eq. 2. The intrinsic growth rates are fixed at $r_x = 3.8$ and $r_y = 3.5$, so that x is in the chaotic regime and the dynamics of y is governed by a period-4 attractor for $\beta_{yx} = 0$, but as illustrated in Figure 2 the dynamics of y will become increasingly chaotic as β_{yx} increases. Results for three different values of β_{yx} ranging from weak to strong coupling will be presented. At low noise levels CCM gives the correct result for all three values of the coupling strength, namely that x causes y .

When the noise level σ_y is increased the cross-mapped estimates of x from M_y deteriorate, and as a result $\rho_{x\hat{x}}$ decreases. Figure 13 shows the fitted values $\rho_{x,\infty}$ that $\rho_{x\hat{x}}$ converges to as a function of the noise level (top panel).

It is seen that $\rho_{x,\infty}$ decreases in a roughly linear fashion as σ_y is increased until the correlation becomes very low ($0 \lesssim \rho_{x,\infty} \lesssim 0.2$) where the curve becomes flat. For systems corresponding to the flat part of the curves $\rho_{x,\infty}$ is very low and at about the same level as $\rho_{y,\infty}$, so, effectively, the system is too noisy for CCM to extract the direction of causality from the converged value of $\rho_{x\hat{x}}$.

The bottom panel shows that the rate of convergence γ_x from Eq. 7 is relatively unaffected by the noise level, except when $\rho_{x\hat{x}} \approx 0$, where CCM is not applicable under any circumstances. But for combinations of coupling strength and noise level that lie on the linear decline of the curves in the top panel of Figure 13 the fitted rate of convergence γ_x seems to be a better indicator that $\rho_{x\hat{x}}$ converges than

445 $\rho_{x,\infty}$; and hence also a better predictor of causality.

446 5.2.1. *Estimating coupling from noise*

447 As noted above, it is observed from Figure 13 that $\rho_{x,\infty}$ decreases linearly as a
 448 function of noise. Further note, that the inverse of the slope depends linearly on
 449 the strength of the coupling, as shown in Figure 14. For the particular example
 450 studied here this has the implication that it is possible to determine the strength of
 451 the coupling from x to y by controlled injection of noise into the system. Whether
 452 this results generalizes to other systems is a question for future research that has
 453 practical implications for measuring the interaction between different parts of
 454 complex dynamic systems.

455 5.2.2. *The influence of an external driving signal*

456 As discussed in section 3.2, it is an important and long-standing problem to
 457 be able to distinguish a true causal connection between variables from an external
 458 driving force that affects both variables. Hence, the study of the external influence
 459 given by the function H_t (cf. Eqs. 3, 4) is now considered, and the question of
 460 whether CCM can extract the correct causal connection between x and y under this
 461 external influence.

462 To answer this question, consider a case with bi-directional coupling, where
 463 the coupling from x to y is twice as strong as in the opposite direction ($\beta_{xy} = 0.01$,
 464 $\beta_{yx} = 0.02$). Slightly different values of the intrinsic growth rates, $r_x = 3.63$,
 465 $r_y = 3.77$ are chosen, and furthermore the noise term in Eq. 3 is set to zero.
 466 The direction of causality is given by the relative magnitudes of the correlation
 467 coefficients $\rho_{x\hat{x}}$ and $\rho_{y\hat{y}}$, so the effect of non-zero values of η_x and η_y on these
 468 measures is examined. A plot of these correlation coefficients is shown in Figure 15,

where the influence of the external signal on the growth rates of x and y is the same, i.e., $\eta_x = \eta_y = \eta$.

As seen in Figure 15 the general trend is—as expected—that the correlation coefficients become larger for larger values of η . But, more importantly, the direction of causality does not change, i.e., the blue curve (ρ_x) stays above the red curve (ρ_y), although for some of the lower values of η the fitted values of $\rho_{y,\infty}$ are above the fitted values of $\rho_{x,\infty}$. Upon closer inspection of the plots of $\rho_{x\hat{x}}$ and $\rho_{y\hat{y}}$ (not shown, but similar to Figure 7) it turns out that the convergence is slower than the exponential form (Eq. 7), resulting in poorer fits than in the case without the external driving signal. As an extra check, the values of $\rho_{x\hat{x}}$ and $\rho_{y\hat{y}}$ are plotted for the largest library size, $L = 800$, shown in Figure 15 as open circles connected by dashed lines. Note that there are systematic differences between these and the fitted values, and that the difference has opposite signs for x and y . It is evident that in the few cases where $\rho_{y,\infty} > \rho_{x,\infty}$ the fits for y are poor (as evidenced by the large error bars) and $\rho_{x\hat{x}}$ and $\rho_{y\hat{y}}$ are better indicators of the direction of causality.

While the complete parameter space has not been systematically sampled, the example above indicates that CCM is relatively robust to a common external influence of the form given by Eqs. 1 and 3. However, it is clear from Figure 15 that in the example $\rho_{y\hat{y}}$ is more sensitive to variations in η than $\rho_{x\hat{x}}$ is, so it is possible that there exist combinations of the period p and the coupling η of the forcing signal that will disrupt CCM, e.g., by synchronizing the time series of x and y . A systematic search is deferred to future work.

6. Discussion

Methods for reliable inference of causality between two or more variables are of great interest to many fields of research, and is particularly relevant for complex systems consisting of many interacting and heterogeneous components. While the example system studied here has its origins in ecology, the underlying model has become an often used example of a simple system displaying complex behavior. The conclusions are therefore relevant for the broader set of complex systems, including complex information systems. It could be imagined, for example, that CCM can be used to understand the causal link between variables from cyber-physical systems such as sensor networks or IoT.

6.1. Limitations of CCM

Convergent Cross-Mapping shows great potential in this regard. However, based on the systematic assessments presented in the previous sections, it was seen that the method can give erroneous results. It was noted already by Sugihara et al. [7] that CCM does not correctly predict the direction of causality when the coupling is so strong that it results in synchronization of variables, as also seen above. Further, The present study shows that the method seems very sensitive to the particular dynamics of the model system, and CCM can also fail to correctly predict the direction of causality in some cases where the coupling is weak to moderate. Interestingly, it was shown that both types of cases where CCM fails is associated with a failure to fit the observed values of cross-mapped correlations to the function $\rho(L)$ in Eq. 7. It would therefore seem that the failure to produce a good fit to $\rho(L)$ is an indicator that CCM is not applicable to the data.

514 6.2. *Dealing with noise in CCM*

515 Another aspect of CCM performance concerns conditions of noise. Generally,
 516 CCM is observed to be fairly robust to random noise and makes reliable inferences
 517 at varying degrees of coupling under conditions of lower noise levels. However,
 518 for higher levels of noise CCM correlations drop linearly as a function of added
 519 noise. For systems subject to noise the results suggest that the rate of convergence
 520 is a more robust indicator of cross-mapping convergence and hence of causality.
 521 In applications where noise can be controlled, injecting noise into the system at
 522 different noise levels presents an opportunity to gauge the strength of coupling
 523 between variables.

524 6.3. *Dealing with external influence*

525 The initial investigations into the effect of an external variable influencing
 526 the coupled sub-systems indicate that stronger external forcing results in higher
 527 correlations between observed data in one variable and estimates obtained from
 528 cross-mapped estimates based on the other variable. No evidence was found that
 529 this external forcing disrupts CCM, but since it has been shown to drastically
 530 change the dynamics of a single variable [34], such a detrimental effect cannot be
 531 ruled out.

532 7. **Conclusion and future work**

533 The reliability of the CCM algorithm has been investigated in a model system
 534 including noise and the influence of an external variable. The authors conclude that
 535 the CCM algorithm can give wrong results, even in the absence of strong coupling.
 536 All observed cases where CCM fails were also associated with a failure to fit to an

537 exponential convergence of the correlation coefficients as a function of library size.
 538 It is therefore proposed that a poor fit (low R^2) is an indicator that CCM results are
 539 not valid. It has been shown that noise decreases cross-map reliability with a linear
 540 response. In systems where noise can be injected, this linear response may be used
 541 as a way of measuring the coupling between subsystems.

542 Together the results presented in this article warrant caution in the application
 543 of CCM to real-world data for purposes of causal inference, and care must be taken
 544 to look in detail at the convergence properties of the correlations between observed
 545 data and cross-mapped estimates. When applied under the right circumstances,
 546 the method has the potential not only to inform the researcher about the causal
 547 direction of dynamics between coupled variables, but controlled injection of noise,
 548 makes it possible infer the coupling strength between the variables.

549 The work presented here, and related work (cf. section 2), point to several
 550 relevant future advances. Since only one particular system has been studied in
 551 depth, future work is needed to investigate to what extent the results generalize to
 552 other systems. In particular bunching of points in phase space has been identified
 553 to be the reason for a failure of CCM in the system studied, but whether this is also
 554 the case for other systems is a topic of future work.

555 Injecting noise to gauge the coupling between subsystems might have practical
 556 implications, especially in complex information systems, that are more amenable
 557 to controlled manipulation than natural or social systems. In addition, naturally
 558 occurring noise spikes in technological or biological systems could potentially be
 559 detected and utilized for coupling strength inference. It will therefore be interesting
 560 to investigate whether the effects of noise injection generalize to other systems.

561 In the model studied here heterogeneity between the two subsystems was

562 included by choosing different parameters in the equations governing the dynamics
 563 of the two variables. It has been suggested by Osorio et al. [42] that heterogeneity
 564 and coupling strength are the relevant characteristics determining if the dynamics
 565 of systems—such as the brain—composed of interacting threshold oscillators are
 566 governed by synchronization, self-organized criticality, or is in an incoherent state.
 567 The investigation of the extent to which large-scale complex systems switch, for
 568 instance, between a synchronization phase and a self-organized criticality phase
 569 could have important implications for control theory and the design of complex
 570 information systems. Future work investigating the effects of systematic variation
 571 of heterogeneity and coupling strength for CCM and overall system state is needed
 572 to shed light on this.

573 The influence of an external driving signal was explored, mainly by looking
 574 at the effect of varying the strength of the external driver. There is reason to
 575 assume that the period of the driving signal plays an important role, and that the
 576 effect on the driven variables is bigger for particular values of the period, that is,
 577 so-called resonance effects. If resonances occur and are sufficiently strong, they
 578 may render CCM unable to detect the direction of causality. To determine this, a
 579 more systematic investigation is needed.

580 Finally, only the original CCM algorithm introduced by Sugihara et al. [7] was
 581 studied here. However, more recently other algorithms have been presented and it
 582 is equally important to investigate the influence of noise and external signals on for
 583 instance the PAI [19], CMS [25] and MCCM [26] algorithms.

Acknowledgements

The authors would like to acknowledge the Interacting Minds Centre, Aarhus University, for providing the ideal environment for the authors' collaboration. The authors wish to thank the two anonymous reviewers for FGCS, the eight anonymous reviewers who provided feedback on an earlier version of the manuscript presented at COMPLEXIS 2016, and the attendees at the conference for valuable feedback.

References

- [1] D. R. Kuhn, Sources of failure in the public switched telephone network, *Computer* 30 (4) (1997) 31–36. doi:10.1109/2.585151.
- [2] P. Pourbeik, P. S. Kundur, C. W. Taylor, The anatomy of a power grid blackout - Root causes and dynamics of recent major blackouts, *IEEE Power and Energy Magazine* 4 (5) (2006) 22–29. doi:10.1109/MPAE.2006.1687814.
- [3] R. Quax, D. Kandhai, P. M. A. Sloot, Information dissipation as an early-warning signal for the Lehman Brothers collapse in financial time series, *Scientific Reports* 3. doi:10.1038/srep01898.
URL <http://www.nature.com/articles/srep01898>
- [4] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, S. Havlin, Catastrophic cascade of failures in interdependent networks, *Nature* 464 (7291) (2010) 1025–1028. doi:10.1038/nature08932.
URL <http://www.nature.com/nature/journal/v464/n7291/abs/nature08932.html>

- 605 [5] J. Pearl, Causal inference in statistics: An overview, *Statistics Surveys* 3
 606 (2009) 96–146. doi:10.1214/09-SS057.
 607 URL <http://projecteuclid.org/euclid.ssu/1255440554>
- 608 [6] C. Cai, K. Zeng, L. Tang, D. Chen, W. Peng, J. Yan, X. Li, Towards
 609 adaptive synchronization measurement of large-scale non-stationary
 610 non-linear data, *Future Generation Computer Systems* 43–44 (2015) 110–119.
 611 doi:10.1016/j.future.2014.06.003.
 612 URL [http://www.sciencedirect.com/science/article/pii/](http://www.sciencedirect.com/science/article/pii/S0167739X14001241)
 613 [S0167739X14001241](http://www.sciencedirect.com/science/article/pii/S0167739X14001241)
- 614 [7] G. Sugihara, R. May, H. Ye, C.-h. Hsieh, E. Deyle, M. Fogarty, S. Munch,
 615 Detecting Causality in Complex Ecosystems, *Science* 338 (6106) (2012)
 616 496–500. doi:10.1126/science.1227079.
 617 URL <http://science.sciencemag.org/content/338/6106/496>
- 618 [8] A. Lendasse, D. Francois, V. Wertz, M. Verleysen, Vector quantization: a
 619 weighted version for time-series forecasting, *Future Generation Computer*
 620 *Systems* 21 (7) (2005) 1056–1067. doi:10.1016/j.future.2004.03.006.
 621 URL [http://www.sciencedirect.com/science/article/pii/](http://www.sciencedirect.com/science/article/pii/S0167739X04000238)
 622 [S0167739X04000238](http://www.sciencedirect.com/science/article/pii/S0167739X04000238)
- 623 [9] H. Abarbanel, *Analysis of Observed Chaotic Data*, Springer New York, 1996.
- 624 [10] H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge Univer-
 625 sity Press, New York, NY, USA, 1997.
- 626 [11] L. Cao, Practical method for determining the minimum embedding dimension
 627 of a scalar time series, *Physica D: Nonlinear Phenomena* 110 (1) (1997)

- 628 43–50. doi:10.1016/S0167-2789(97)00118-8.
 629 URL [http://www.sciencedirect.com/science/article/pii/](http://www.sciencedirect.com/science/article/pii/S0167278997001188)
 630 [S0167278997001188](http://www.sciencedirect.com/science/article/pii/S0167278997001188)
- 631 [12] M. B. Kennel, R. Brown, H. D. I. Abarbanel, Determining embedding dimen-
 632 sion for phase-space reconstruction using a geometrical construction, *Physical*
 633 *Review A* 45 (6) (1992) 3403–3411. doi:10.1103/PhysRevA.45.3403.
 634 URL <http://link.aps.org/doi/10.1103/PhysRevA.45.3403>
- 635 [13] A. M. Fraser, H. L. Swinney, Independent coordinates for strange attractors
 636 from mutual information, *Physical Review A* 33 (2) (1986) 1134–1140.
 637 doi:10.1103/PhysRevA.33.1134.
 638 URL <http://link.aps.org/doi/10.1103/PhysRevA.33.1134>
- 639 [14] T. Schreiber, Measuring Information Transfer, *Physical Review Letters* 85 (2)
 640 (2000) 461–464. doi:10.1103/PhysRevLett.85.461.
 641 URL <http://link.aps.org/doi/10.1103/PhysRevLett.85.461>
- 642 [15] C. L. Webber, J. P. Zbilut, Dynamical assessment of physiological systems
 643 and states using recurrence plot strategies, *Journal of Applied Physiology*
 644 76 (2) (1994) 965–973.
 645 URL <http://jap.physiology.org/content/76/2/965>
- 646 [16] N. Marwan, M. Carmen Romano, M. Thiel, J. Kurths, Recurrence plots for
 647 the analysis of complex systems, *Physics Reports* 438 (5–6) (2007) 237–329.
 648 doi:10.1016/j.physrep.2006.11.001.
 649 URL [http://www.sciencedirect.com/science/article/pii/](http://www.sciencedirect.com/science/article/pii/S0370157306004066)
 650 [S0370157306004066](http://www.sciencedirect.com/science/article/pii/S0370157306004066)

- [17] J. C. McBride, X. Zhao, N. B. Munro, G. A. Jicha, F. A. Schmitt, R. J. Kryscio, C. D. Smith, Y. Jiang, Sugihara causality analysis of scalp EEG for detection of early Alzheimer's disease, *NeuroImage: Clinical* 7 (2015) 258–265. doi:10.1016/j.nicl.2014.12.005.
URL <http://www.sciencedirect.com/science/article/pii/S2213158214001909>
- [18] A. E. BozorgMagham, S. Motesharrei, S. G. Penny, E. Kalnay, Causality Analysis: Identifying the Leading Element in a Coupled Dynamical System, *PLOS ONE* 10 (6) (2015) e0131226. doi:10.1371/journal.pone.0131226.
URL <http://dx.plos.org/10.1371/journal.pone.0131226>
- [19] J. M. McCracken, R. S. Weigel, Convergent cross-mapping and pairwise asymmetric inference, *Physical Review E* 90 (6) (2014) 062903. doi:10.1103/PhysRevE.90.062903.
URL <http://link.aps.org/doi/10.1103/PhysRevE.90.062903>
- [20] C. W. J. Granger, Investigating Causal Relations by Econometric Models and Cross-spectral Methods, *Econometrica* 37 (3) (1969) 424–438. doi:10.2307/1912791.
URL <http://www.jstor.org/stable/1912791>
- [21] A. Arnold, Y. Liu, N. Abe, Temporal Causal Modeling with Graphical Granger Methods, in: *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '07*, ACM, New York, NY, USA, 2007, pp. 66–75. doi:10.1145/1281192.1281203.
URL <http://doi.acm.org/10.1145/1281192.1281203>

- [22] G. Lakoff, Why it Matters How We Frame the Environment, *Environmental Communication* 4 (1) (2010) 70–81. doi:10.1080/17524030903529749.
URL <http://dx.doi.org/10.1080/17524030903529749>
- [23] 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. doi:10.1137/130946344.
URL <http://epubs.siam.org/doi/abs/10.1137/130946344>
- [24] L. M. Pecora, T. L. Carroll, J. F. Heagy, Statistics for mathematical properties of maps between time series embeddings, *Physical Review E* 52 (4) (1995) 3420–3439. doi:10.1103/PhysRevE.52.3420.
URL <http://link.aps.org/doi/10.1103/PhysRevE.52.3420>
- [25] H. Ma, K. Aihara, L. Chen, Detecting Causality from Nonlinear Dynamics with Short-term Time Series, *Scientific Reports* 4 (2014) 7464. doi:10.1038/srep07464.
URL <http://www.nature.com/articles/srep07464>
- [26] 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. doi:10.1890/14-1479.1.
URL <http://onlinelibrary.wiley.com/doi/10.1890/14-1479.1/abstract>
- [27] J.-J. Jiang, Z.-G. Huang, L. Huang, H. Liu, Y.-C. Lai, Directed dynamical influence is more detectable with noise, *Scientific Reports* 6.

- doi:10.1038/srep24088.
- URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4828632/>
- [28] R. M. May, Simple mathematical models with very complicated dynamics, *Nature* 261 (5560) (1976) 459–467. doi:10.1038/261459a0.
- URL <http://www.nature.com/nature/journal/v261/n5560/abs/261459a0.html>
- [29] A. Ferretti, N. K. Rahman, A study of coupled logistic map and its applications in chemical physics, *Chemical Physics* 119 (2–3) (1988) 275–288. doi:10.1016/0301-0104(88)87190-8.
- URL <http://www.sciencedirect.com/science/article/pii/0301010488871908>
- [30] A. L. Lloyd, The coupled logistic map: a simple model for the effects of spatial heterogeneity on population dynamics, *Journal of Theoretical Biology* 173 (3) (1995) 217–230. doi:10.1006/jtbi.1995.0058.
- URL <http://www.sciencedirect.com/science/article/pii/S0022519385700586>
- [31] P. Moran, The statistical analysis of the Canadian Lynx cycle., *Australian Journal of Zoology* 1 (3) (1953) 291–298.
- URL <http://www.publish.csiro.au/paper/Z09530291>
- [32] E. Ranta, V. Kaitala, J. Lindström, E. Helle, The Moran Effect and Synchrony in Population Dynamics, *Oikos* 78 (1) (1997) 136–142. doi:10.2307/3545809.
- URL <http://www.jstor.org/stable/3545809>

- [33] T. Royama, Moran Effect on Nonlinear Population Processes, *Ecological Monographs* 75 (2) (2005) 277–293. doi:10.1890/04-0770.
 URL <http://onlinelibrary.wiley.com/doi/10.1890/04-0770/abstract>
- [34] D. Summers, J. G. Cranford, B. P. Healey, Chaos in periodically forced discrete-time ecosystem models, *Chaos, Solitons & Fractals* 11 (14) (2000) 2331–2342. doi:10.1016/S0960-0779(99)00154-X.
 URL <http://www.sciencedirect.com/science/article/pii/S096007799900154X>
- [35] F. Takens, Detecting strange attractors in turbulence, *Lecture Notes in Mathematics* 898 (1981) 366–381. doi:10.1007/BFb0091924.
- [36] G. Sugihara, R. M. May, Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series, *Nature* 344 (6268) (1990) 734–741. doi:10.1038/344734a0.
 URL <http://www.nature.com/nature/journal/v344/n6268/abs/344734a0.html>
- [37] D. Mønster, XMAP (2013).
 URL <https://github.com/danm0nster/xmap>
- [38] S. N. Jespersen, Personal communication (Jan. 2013).
- [39] S. Boccaletti, J. Kurths, G. Osipov, D. L. Valladares, C. S. Zhou, The synchronization of chaotic systems, *Physics Reports* 366 (1–2) (2002) 1–101. doi:10.1016/S0370-1573(02)00137-0.

- 740 URL [http://www.sciencedirect.com/science/article/pii/](http://www.sciencedirect.com/science/article/pii/S0370157302001370)
741 [S0370157302001370](http://www.sciencedirect.com/science/article/pii/S0370157302001370)
- 742 [40] A. Pikovsky, M. Rosenblum, J. Kurths, Synchronization: A Universal Con-
743 cept in Nonlinear Sciences, Cambridge University Press, 2003.
- 744 [41] E. Ott, Chaos in Dynamical Systems, 2nd Edition, Cambridge University
745 Press, Cambridge, U.K. ; New York, 2002.
- 746 [42] I. Osorio, M. G. Frei, D. Sornette, J. Milton, Y.-C. Lai, Epileptic
747 seizures: Quakes of the brain?, Physical Review E 82 (2) (2010) 021919.
748 doi:10.1103/PhysRevE.82.021919.
749 URL <http://link.aps.org/doi/10.1103/PhysRevE.82.021919>

Dan Mønster holds an M.Sc. and a Ph.D in physics, both from Aarhus University. He is former CTO of the Danish Research Network, and has a background in high-performance computing at the Danish Computing Centre for Research and Education. Currently, he is Assistant Professor in Business and Behavioral Science at Aarhus University, where his research focuses on experiments, simulation and analysis of social interactions in complex systems.

751

Riccardo Fusaroli investigates interpersonal coordination mixing experimental, clinical, and computational perspectives. His focus is on coordinative efficacy (how do we effectively coordinate and think together?) and social impairment (what can go wrong?), as well as on automated continuous behavioral measures and statistical methods. Riccardo Fusaroli is currently Associate Professor in Cognitive Science at the Interacting Minds Center, Aarhus University.

752

Kristian Tylén, PhD, is Associate Professor in Cognitive Science at the School of Communication and Culture and The Interacting Minds Center, Aarhus University. He has a background in cognitive semiotics from Aarhus University and linguistics from University of Southern Denmark. His research includes experimental investigations of social coordination, collective problem solving, and language evolution, relying on methods from cognitive science, dynamical systems and neuroscience.

753

Andreas Roepstorff is an anthropologist in and of cognitive science, with a primary focus on human interaction. He is professor in cognition, communication and culture and director of the Interacting Minds Centre at Aarhus University.

754

Jacob F. Sherson, Ph.D, is currently Associate Professor in experimental and theoretical quantum physics at Aarhus University. He is also the director of the interdisciplinary AU Ideas Center for Community Driven Research, in which he and his team have created the citizen science platform, www.scienceathome.org, used by 150,000 citizens worldwide. One main focus area of the center is to study processes of human problem solving to inspire novel machine learning algorithms.

Figure 4: Example of time series generated by the model in Eq. 3 with bidirectional, asymmetric coupling between x and y , and equal coupling from the external driving signal H to both x and y . The noise terms are set to zero. Parameters are: $r_x = 3.5$, $r_y = 3.7$, $\beta_{yx} = 0.02$, $\beta_{xy} = 0.01$, $p = 10$, $\eta_x = \eta_y = 0.2$, $\sigma = 0$

Figure 5: The shadow manifolds based on the time series displayed in Figure 1, embedded using $E = 2$ and $\tau = 1$. A total of 60 points $\mathbf{x}_t = (x_t, x_{t-1})$ in M_x are shown in blue and the points $\mathbf{y}_t = (y_t, y_{t-1})$ in M_y are shown in red.

Figure 6: Scatter plot of pairs of observed values of x and estimated values $\hat{x}|M_y$ (left panel) and the equivalent for y (right panel). The cross-mapped estimates were computed according to Eq. 5 and were based on the shadow manifolds in Figure 5.

Figure 7: The correlation coefficient $\rho_{x, \hat{x}}$ (blue circles) and $\rho_{y, \hat{y}}$ (red circles) as a function of library length L . Fits to the function in Eq. 7 are shown as solid lines. For x the converged value is $\rho_{\infty} = 0.71$, and for y it is $\rho_{\infty} = 0.096$. Parameter values: $r_x = 3.65$, $r_y = 3.77$, $\beta_{xy} = 0$, $\beta_{yx} = 0.05$.

Figure 8: Correlation coefficients $\rho_{\hat{x}\hat{x}}$ (blue) and $\rho_{\hat{y}\hat{y}}$ (red) for ‘linear’ sampling (a), ‘random’ sampling (b), and ensemble average with random sampling (c). The parameters used (cf. Eq. 3) were: $r_x = 3.63$, $r_y = 3.77$, $\beta_{xy} = 0.01$, $\beta_{yx} = 0.05$, $\sigma = 0$, $\eta_x = \eta_y = 0$.

Figure 9: Correlations between observed and cross-mapped estimates of x (blue) and y (red) for four different values of β_{yx} . Thin lines are the computed correlations and thick lines are fits to Eq. 7. For $\beta_{yx} = 0.2$ and $\beta_{yx} = 2$ a good fit is not possible, and no fits are shown.

Figure 10: Plot of the time series of x (shown in blue) and y (shown in red) displaying nearly periodic behavior. Parameter values: $r_x = 3.625$, $r_y = 3.77$, $\beta_{xy} = 0$, $\beta_{yx} = 0.2$.

Figure 11: Cross-mapped estimates vs. observed values of x and y . Parameter values: $r_x = 3.625$, $r_y = 3.77$, $\beta_{xy} = 0$, $\beta_{yx} = 0.2$.

Figure 12: Surface plot of p_{yy} for $L = 400$ as a function of r_y and β_{yx} . Note that r_x is also varied, since $r_x + r_y = 7.3984$. $\beta_{xy} = 0$ throughout.

Figure 13: The effect of the noise level σ_y on the converged value $\rho_{x,\infty}$ of $\rho_{x\hat{x}}$ (top panel) and on the rate of convergence γ_x of $\rho_{x\hat{x}}$ (bottom panel). In all cases $r_x = 3.8$, $r_y = 3.5$ and $\beta_{xy} = 0$. The coupling β_{yx} was varied as shown in the figure. The error bars represent 95% confidence intervals on the fitted values.

Figure 14: The inverse negative slope $-1/q$ of the linear portion of the curves in the top panel of Figure 13 as a function of the coupling. The straight line is the result of a linear fit, which has coefficient of determination $R^2 = 0.953$.

Figure 15: Correlation coefficients for x (blue) and y (red) as a function of $\eta = \eta_x = \eta_y$. Solid lines with error bars (95% CI) are the fitted values, $\rho_{x,\infty}$, $\rho_{y,\infty}$. Dashed lines with open circles are the values of $\rho_{x\hat{x}}$ and $\rho_{y\hat{y}}$ for the largest library size, $L = 800$. The parameters in the model (Eq. 3) are: $r_x = 3.63$, $r_y = 3.77$, $\sigma = 0$, $\beta_{xy} = 0.01$, $\beta_{yx} = 0.02$, and $p = 4$.