Detecting dynamical interdependence and generalised synchronisation using the Lorenz method of analogues.

Erik Casagrande
Neural Computing Research Group
School of Engineering and Applied Science
Aston University
Email: casagrae@aston.ac.uk

David Lowe
Neural Computing Research Group
School of Engineering and Applied Science
Aston University
Email: d.lowe@aston.ac.uk

Avgis Hadjipapas
The Wellcome Trust Laboratory
School of Life and Health Sciences
Aston University
Email: hadjipaa@aston.ac.uk

***Abstract*— We provide results comparing two new methods based on different embedding space approaches to infer directional interdependence between bivariate time series, such as the transition to synchronisation. Both methods share a common identification stage which models the underlying generator of each single observable data. This consists of two parts: state space reconstruction, and local model fitting using a class of kernel density estimation methods which is an extension of Lorenz Analogues techniques. The mutual information derived from the density estimation is used to compare the two approaches and is tested on coupled dynamical systems for which the interdependence is controlled.**

I. Introduction

Identifying dynamical directional interdependence among multivariate time series is an important problem in many scientific fields such as engineering, econometrics, biology and biomedical analysis. Generally, the mathematical formulation of the underlying generator of the observed signals is not known a priori and their coupling needs to be inferred using observational data. This latent generator can be represented by a composite interplay of many subsystems which can result in a global complex collective behaviour. Instead of the complete statistical description of hidden systems, it can be sufficient to estimate information about the relationship between observed signals. An accurate analysis of these relationships can bring about useful inference regarding the hidden connections inside the underlying generator.

The nonlinearity of the system is a further complication which is common in nature and can make the generator dynamics more difficult to analyse. Linear tools such as correlation or Fourier analysis have limitations for uncovering nonlinear couplings. Instead, in the literature, nonlinear interdependence concepts such as causality or synchronisation are being used to construct a framework which, in this work, we use as the motivational background for our study.

Causality has been the subject of debate due to different views around its nature and interpretation. While an empirical idea of cause and effect is relatively easy to comprehend, its practical translation can lead to an algorithmic representation providing many issues (see [1] for a discussion). Granger [2] based on ideas from Wiener, studied a particular class of linear stochastic systems where a plausible definition of causality proposed a statistical framework of analysis. Given two observable time series x and y from systems $\mathcal{X}$ and $\mathcal{Y}$, if a k-step ahead predictor of x_i using the past observable of x_i and y_i together increases the predictability likelihood over a predictor using only x_i, than we can say that $\mathcal{Y}$ and $\mathcal{X}$ are causally related. Depending on the flow of information in predicting $\mathcal{Y}$ from $\mathcal{X}$ or $\mathcal{X}$ from $\mathcal{Y}$ allows us to infer directionality of coupling. From a different point of view, we can express the same concept in a probabilistic way,

$$\text{If } P(x_{i+k}|x_i^-, y_i^-) \neq P(x_{i+k}|x_i^-) \qquad \mathcal{Y} \text{ causes } \mathcal{X} \tag{1}$$

where x_i^- and y_i^- are the past time series of the systems prior to the time i and x_{i+k} is the k-step ahead predictor. Using a linear stochastic system, a totally admissible algorithmic choice [2] which has been largely used, based on (1) uses the conditional variance of the prediction error, i.e., $\sigma(x_{i+k}|x_i^-, y_i^-)$ and $\sigma(x_{i+k}|x_i^-)$. The Granger implementation was based on ARMA modelling and has been employed in many data analysis domains, including, econometrics [3] [4] and recently neuroscience [5]. Since for a linear system the autocorrelation and Fourier spectrum are related by the Wiener-Khinchin theorem, we can find implementations of causality measures based on coherence [2]. Recent studies have addressed the nonlinear analysis issue using an extension of Granger causality based on nonlinear predictor modelling [6], [7], [8]. Other approaches using entropy (for a review see [9]) have been developed. In [10] (1) has been called the *measure of deviance from the Markov property* which leads to the concept of Transfer Entropy.

Most of the methods we have listed above for causality discovery containing nonlinear systems make use of a preprocessing stage based on State Space Reconstruction [11]. This approach has become a popular solution in order to learn the underlying dynamical structure which generates the observable time series. We use this technique in this paper to build the reconstructed state space in which we run the modified Lorenz analogues approaches.

Defining *synchronisation* has been less empirical due its historical focus on periodic system such as relaxation oscillators in which phase contains the information needed to infer interdependence and synchronisation. In the last two decades, it has been shown [12] that chaotic systems can identically synchronise and other levels of definition needed to be introduced [13]. One of the reasons was the query of how to define the phase of a chaotic signal. There are different measures of synchronisation and we refer to the literature [14], [15]. In this work, we focus on a class of methods which are related to the concept of *Generalised Synchronisation* [16]. Two systems are synchronised if it is possible to find a smooth function Ψ which relates the two state-spaces of the underlying generators of $\mathcal{X}$ and $\mathcal{Y}$, i.e., $\mathbf{Y} = \Psi(\mathbf{X})$, where $\mathbf{X}$ and $\mathbf{Y}$ are the state spaces of $\mathcal{X}$ and $\mathcal{Y}$. A set of possible measures of generalised synchronisation have been explored using the concept of mutual neighbours in the reconstructed or embedding state spaces which we denote here as $\hat{\mathbf{X}}$ and $\hat{\mathbf{Y}}$. In [17], the author projects the local structure of one reconstructed state space, i.e. $\hat{\mathbf{X}}$, on the other, i.e. $\hat{\mathbf{Y}}$ in order to search for generalised synchronisation. Other similar measures can be found in [18] and [19] in which the *Synchronisation Likelihood* is defined. In [20] the procedure of mutual neighbours is used to build a simple nonlinear predictor in state space.

In this paper, we compare two nonlinear approaches to interdependence analysis motivated by a local nonlinear mapping version of the Lorenz method of analogues. The Lorenz technique was originally a simple way to build a predictor in the reconstructed space. We first consider here a novel version of the Lorenz algorithm which uses a kernel density estimation techniques in order to compensate nonlinearity of the underlying generator of the data. This particular predictor models the conditional distributions required in (1) which we employ to check, using the deviation of the Markov property, the information flow between the two time series generators. In the second part, we investigate two possible ways we found in literature, to obtain the joint information of the past sets x_i^- and y_i^- on the left hand side of (1). The causality literature suggests approaches based on joining the sets x_i^- and y_i^- obtaining a larger dimensionality set $x_i^- \oplus y_i^-$ based on a 'direct sum' generator space. Instead, using an approach emphasized for example in [20] [19], we use a method based on mutual mappings between local neighbours in the two spaces which we denote by $P(x_{i+1}|x_i^-, y_i^-) = P(x_{i+1}|x_i^- \star y_i^-)$ with $\star$ the notation that we use to define the mutual neighbours mapping operator, to be discussed later.

In the following section, we define the approach we use to construct nonparametric predictors in embedding space using the Lorenz method of analogues. We define in the second part the mutual predictors using the two different approaches using $\oplus$, $\star$ operators. We employ a Granger like statistic to check the deviance to the Markov property in (1) as our measure of directed interdependence. We illustrate the approach using synthetic examples where we have control over the coupling.

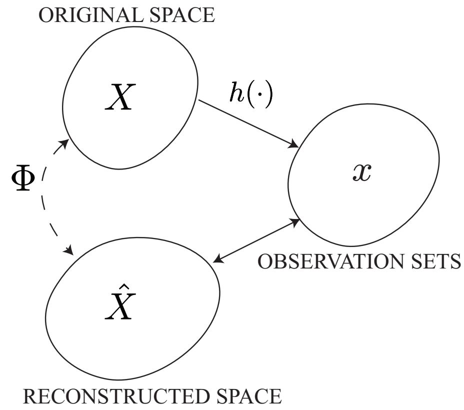


Fig. 1. The dynamics of the underlying generator described by the space $\mathbf{X}$ are projected to the observable subspace x by the function $h(\cdot)$. A space Reconstruction algorithm uses the time series point in x to build an equivalent space $\hat{\mathbf{X}}$ of $\mathbf{X}$ up to a diffeomorphic function $\Phi(\cdot)$.

II. METHODS

A. Nonlinear modelling

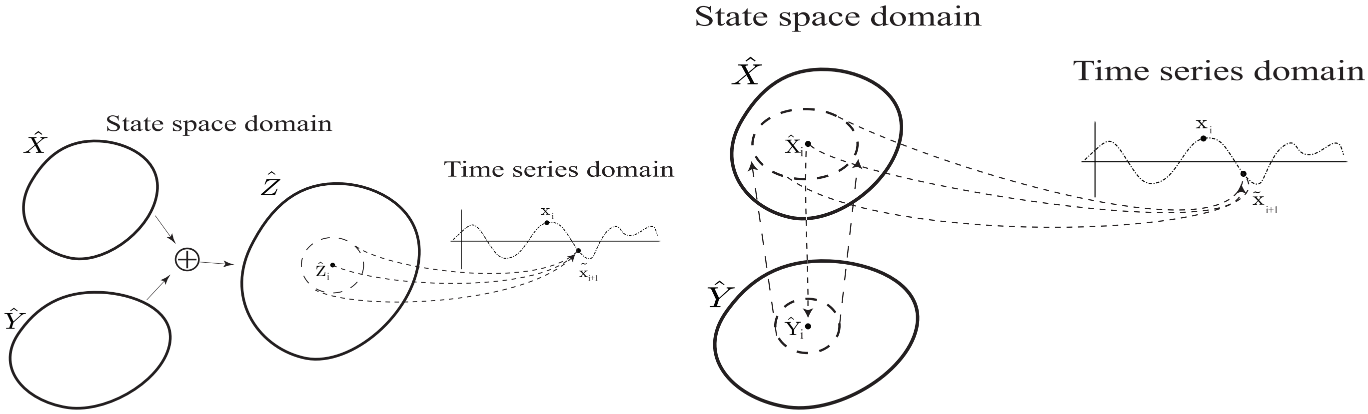
We consider two finite dimensional dynamical systems $\mathcal{X}$ and $\mathcal{Y}$ which describe our underlying generators of the data. Using the state space representations in (2-3), we assume the dynamics of the system is described by stochastic differential equations $f_{\mathcal{X}}(\cdot)$, $f_{\mathcal{Y}}(\cdot)$ and the measure functions $h_{\mathcal{X}}(\cdot)$, $h_{\mathcal{Y}}(\cdot)$ which defines the observable time series x_i and y_i.

$$\mathcal{X} = \begin{cases} \mathbf{X}_{i+1} & = & f_{\mathcal{X}}(\mathbf{X}_i, \mathbf{Y}_i) + \xi_{\mathcal{X}} \\ x_i & = & h_{\mathcal{X}}(\mathbf{X}_i) + \nu_{\mathcal{X}} \end{cases} \tag{2}$$

$$\mathcal{Y} = \begin{cases} \mathbf{Y}_{i+1} & = & f_{\mathcal{Y}}(\mathbf{Y}_i, \mathbf{X}_i) + \xi_{\mathcal{Y}} \\ y_i & = & h_{\mathcal{Y}}(\mathbf{Y}_i) + \nu_{\mathcal{Y}} \end{cases} \tag{3}$$

where ξ and ν are general stochastic terms which we consider additive and uncorrelated with each other. The observable time series are the quantities between which we want to infer interdependence which we consider here to be one-dimensional. As we introduced in the previous section, we use the bold notation, i.e. $\mathbf{X}$, to define a vector space while the indexed bold one, i.e. $\mathbf{X_i}$, to specify an element. The lower notation, i.e. x, indicates the corresponding observable time series.

We proceed to the nonlinear modelling of x and y using a two stage identification procedure. In the first part, we use a *State Space Reconstruction* which we described (1). For each time point i of the time series, we construct an M-dimensional space from the vectors $\hat{\mathbf{X}}_i = (x_i, x_{i+\tau}, \ldots, x_{i+(M-1)\tau})$ and $\hat{\mathbf{Y}}_i = (y_i, y_{i+\tau}, \ldots, y_{i+(M-1)\tau})$ where τ is the delay parameter [21]. In order to simplify the analysis, in this paper we assume that both $\mathcal{X}$ and $\mathcal{Y}$ are embedded in the same dimension M and have common delay τ. For a linear stochastic system, the dimension of the spaces $\hat{\mathbf{X}}_i$ and $\hat{\mathbf{Y}}_i$ are related by the order of the model which can be retrieved for autoregressive representations by the AIC criterion [22], for



(a) Joint Space

(b) Mutual Neighbours

Fig. 2. Graphical comparison between the two approaches we use to extract the conditional joint distribution between the two reconstructed spaces.

instance. In nonlinear time series analysis, Takens [11] showed that under the sufficient condition that $M \geq D+1$, the space $\hat{\mathbf{X}}$ and the space of the vectors $\mathbf{X}$ are an *embedding*, i.e., they are related by a diffeomorphic map Φ. We have that D is the dimension of the limit set of $\mathbf{X}$ or the attractor of (2-3).

The previous property leads to a useful fact for the purpose of building a prediction model which we consider as the second identification stage . The projection over the observable space x of the true joint distribution of the data $p(\mathbf{X}, x)$ can be retrieved using $p(\hat{\mathbf{X}}, x)$ instead, thanks to the diffeomorphic condition between $\hat{\mathbf{X}}$ and $\mathbf{X}$. In this work, we model $p(\hat{\mathbf{X}}, x)$ (the same for $p(\hat{\mathbf{Y}}, y)$) using a class of kernel density estimation methods based on Parzen estimators from which we can derive the following conditional expectation (see [23] for details):

$$\tilde{x}_{i+k} = E(x|\hat{\mathbf{X}}) = \sum_{j=1}^{N} w(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_j) x_j \tag{4}$$

where $\tilde{x}_{i+k}$ is the k-step ahead predictor of the time series point x given by the N-nearest points $\hat{\mathbf{X}}_j$ to $\hat{\mathbf{X}}_i$. The kernel function $w(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_j)$ is imposed to satisfy the summation constraint $\sum_{j=1}^{N} w(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_j) = 1$. If we choose the simplest kernel function as a constant, i.e., $w(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_j) = 1/N$, the approach is equivalent to the well-known *Lorenz method of Analogues* [21] or the *zero-th order* predictor [20].

A different choice of the kernel $w(\cdot, \cdot)$ can be used in order to increase the performance and the complexity of the predictor model. In this paper we continue with the Lorenz method, since the main goal is the comparison between the mutual predictors approaches defined by the operators $\oplus$ and $\star$ which we describe in the next section. We leave the performance analysis for causality discovery using different kernels as future development.

The approach we describe is an example of a *memory-based* method in the pattern recognition literature [23]. From the technique we use to build the reconstructed space, we have a one-to-one relationship between the spaces $\hat{\mathbf{X}}$ and the observable space x that can be considered as a database. We further divided this database in two parts: a training set and a test set. Normally, the training set is used to learn the conditional probability $p(\hat{\mathbf{X}}, x)$ using a minimization algorithm which is then used on the test set to produce the predicted value $\tilde{x}$. Memory-based methods do not learn during the training stage but at each new embedding vector $\hat{\mathbf{X}}_i$ the algorithm needs to find the neighbours $\hat{\mathbf{X}}_j$ and project the neighbourhood structure on the correspond time series points x_j. Computationally, we have that memory based methods are fast in training, since they require only to build the set, but slow in prediction. In order to increase the performance of the algorithm, we need to build the training set using an optimal procedure. Algorithms which build the embedding space using K-D trees [24] can compute the neighbours problem in $O(Klog\,K)$ time compared to an unstructured solution. The complexity of the neighbours problem is $O(dK^2)$ if we need to compute the pairwise distance of K embedding d dimensional vectors.

The analogues algorithm requires parameters to be tuned such as the dimension of the embedding, the delay parameter and the number of neighbours. Using a different kernel function $w(\cdot, \cdot)$ could result in further parameters to choose. In the literature [21], there are several procedures developed to tune the embedding parameters M and τ. In this work, since the purpose of the predictor methods is not forecasting, but method comparison, we can use the available test set to drive the choice of the optimal parametrization of the algorithm. As described in the results section, we collect the time series for an offline interdependence analysis. The test set is known a priori and we consider a *Cross-Validation-Procedure* to choose the parameters which minimise the out-of-sample prediction errors in this set. The choice of using the test set in order to compute the prediction errors avoids the problem of model overfitting of the training data. We define the prediction errors as e_x, e_y.

B. Mutual Predictors

Following (1) and the Granger paradigm, considering one time series, i.e. x, we want to examine if the information gained by incorporating the other time series, i.e., y improves the nonlinear predictability of x. In the section above we modelled the conditional distribution of the right hand side of (1). In this part we are interested in devising a mutual predictor of x which takes into account the information of both embedding spaces $\hat{\mathbf{X}}$,$\hat{\mathbf{Y}}$ in order to model the left hand side of (1) and quantify the deviation from the Markov property.

As described in the introduction, in this paper we investigate two different approaches to gather mutual information from the reconstructed spaces $\hat{\mathbf{X}}$ and $\hat{\mathbf{Y}}$. The first one, which we refer to as the *Joint space approach* is depicted in Fig. (2) and it uses the direct sum space $\hat{\mathbf{Z}} = \hat{\mathbf{X}} \oplus \hat{\mathbf{Y}}$ where $z_i = (x_i, x_{i+\tau}, \ldots, x_{i+(M-1)\tau}, y_i, y_{i+\tau}, \ldots, y_{i+(M-1)\tau})$ is a double-dimensional concatenated embedding space. Using the space $\hat{\mathbf{Z}}$ we model the conditional probabilities $p(\hat{\mathbf{Z}}, x)$ and $p(\hat{\mathbf{Z}}, y)$ using the kernel analogues approach we discussed in the previous section. This mutual predictor for the case of $p(\hat{\mathbf{Z}}, x)$ assumes the following expression

$$(\tilde{x}|xy)_{i+k} = \sum_{j=1}^{N} w(\hat{\mathbf{Z}}_i, \hat{\mathbf{Z}}_j) x_j \tag{5}$$

where $\hat{\mathbf{Z}}_j$ are the N-nearest neighbours of $\hat{\mathbf{Z}}_i$.

For the second approach, the *Mutual neighbours approach*, we employ an alternative notation using the operator $\hat{\mathbf{X}} \star \hat{\mathbf{Y}}$ to extract mutual information from the two reconstructed spaces $\hat{\mathbf{X}}$ and $\hat{\mathbf{Y}}$. Figure (2) depicts this approach. We consider the embedding vector $\hat{\mathbf{X}}_i$ at time point i and we search for the corresponding counterpart $\hat{\mathbf{Y}}_i$ in the partner space. We then obtain the N-nearest neighbours of $\hat{\mathbf{Y}}_i$ in the partner space and we gather the respective set of time points, i.e. $\hat{\mathbf{Y}}(j)$. We choose the mutual set of neighbours back in the original $\hat{\mathbf{X}}$ space corresponding to the $\hat{\mathbf{Y}}$-space neighbours $\hat{\mathbf{Y}}(j)$ and we build the following mapping predictor

$$(\tilde{x}|xy)_{i+k} = \sum_{j=1}^{N} w(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_{\hat{\mathbf{Y}}(j)}) x_{\hat{\mathbf{Y}}(j)} \tag{6}$$

As in the previous section, we compute for (6) and (5) the k-th step prediction out-of-sample errors on the test set. We denote this error as $e_{x|xy}$ and $e_{y|xy}$, respectively.

From the construction of the algorithms, we emphasise that the mutual neighbours approach is computationally faster than the joint space approach. Joining the two spaces needs an additional searching procedure for neighbours in the new higher dimensional space $\hat{\mathbf{Z}}$. Instead, the mutual neighbours technique can re-use the local structure which has been computed for each $\hat{\mathbf{X}}$ and $\hat{\mathbf{Y}}$ separately.

C. Interdependence Analysis

In the final step, we need to employ a measure of interdependence which quantifies the deviation away from the Markov property in (1). As discussed in the introduction, the conditional probabilities are usually substituted by the variances of the prediction errors which in this case are estimated from (4) and (5) or (6) for the conditional probabilities. This choice is valid if we consider the system has been generated by a linear stochastic component. In this work, we model the underlying generator of the data using a local linear kernel mapping which is consistent with this assumption.

The Granger literature, for example [3] [25], has largely discussed different statistical significance tests in order to check the difference between single and mutual predictability. A comparison of significance tests can be found in [26]. In this work we employ the following measure which has been used in [6]:

$$G_{y \to x} = ln \left(\frac{\sigma_x^2}{\sigma_{x|xy}^2} \right). \tag{7}$$

where σ_x^2 and $\sigma_{x|xy}^2$ are respectively the variance of the single and the mutual prediction errors. Similarly, we compute the term $G_{x \to y}$ using σ_y^2 and $\sigma_{y|xy}^2$ to check for the inverse link. A potential difference between $G_{y \to x}$ and the equivalent $G_{x \to y}$ therefore provides a test for directionality of the coupling which we define as the *differential Granger measure* $\Delta G = G_{x \to y} - G_{y \to x}$. In [7], the authors proposed another index for bidirectional coupling $D = \frac{c_2 - c_1}{c_1 + c_2}$ where $c_1 = \sigma_x^2 - \sigma_{x|xy}^2$ and $c_2 = \sigma_y^2 - \sigma_{y|xy}^2$ but we leave this index for future comment.

III. Results

We consider two numerical experiments of nonlinear systems common in the literature to compare interdependence behaviour. We consider here a quantitative study of the two approaches introduced previously. During the simulation, we discard the first 10^4 synthetically generated samples (the 'burn-in' samples) and we divide the rest into 20 sets of 1000 points. On each set we construct the models discussed previously. We further partitioned each set into two subsets of 700 samples for the training and 300 for the test procedures. In the following discussion, we compute the mean and standard deviation of the differential Granger measure ΔG computed over the 20 sets for one-step ahead predictors.

A. Logistic map

We consider first a unidirectional coupled logistic map which has been described in [7]:

$$\mathcal{X}: \quad x_{i+1} = a x_i (1 - x_i) + s \eta_{i+1} \tag{8}$$

$$\mathcal{Y}: \quad y_{i+1} = (1 - C) a y_i (1 - y_i) + C a x_i (1 - x_i) + s \xi_{i+1} \tag{9}$$

where η and ξ are unit variance Gaussian distributed noise and s controls the strength of the noise terms. The parameter $a = 3.8$ has been chosen. The coupling can vary between $C = 0$ and $C = 0.5$ for the noiseless case. In Fig (7), we plot the synchronisation pattern for different coupling strengths

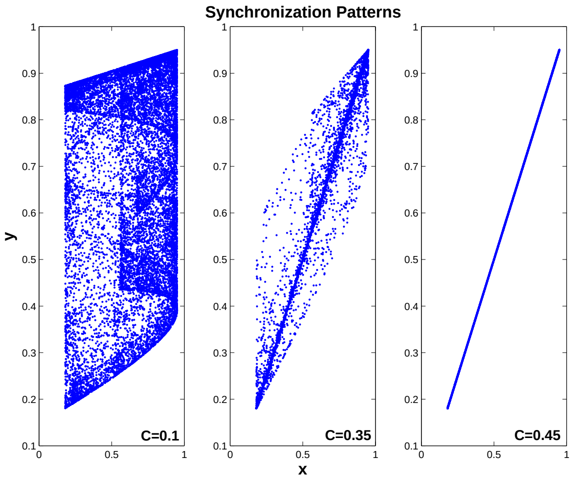


Fig. 3. Synchronisation patterns x vs y for the noiseless coupled logistic map for different values of coupling C. We notice perfect synchronisation for $C = 0.45$

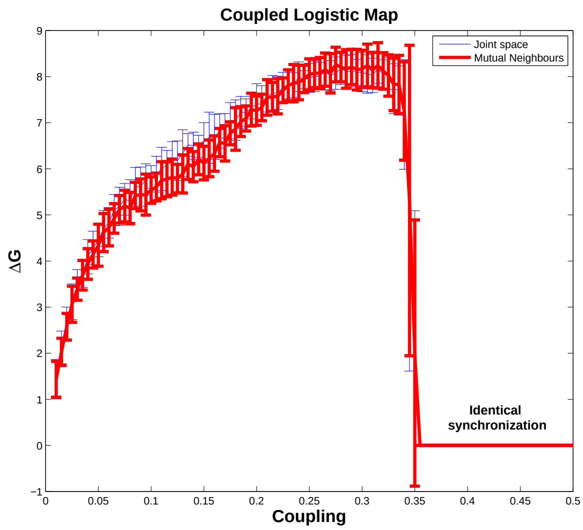


Fig. 4. Values of the directional Granger measure ΔG for different values of the coupling. The bold line is the mutual neighbours approach and the thin the joint space approach. $\Delta G > 0$ indicates the correct information flow from $\mathcal{Y}$ to $\mathcal{X}$

of the map. The case of identical ('perfect') synchronisation occurs for $C > 0.35$, is evident in the figure.

In Fig (4), we consider the measure of ΔG for the two different embedding space treatments. For this particular map we perform a cross validation procedure and we choose $M = 1$ and $\tau = 1$. Moreover, we establish that two neighbours are sufficient to build the optimal predictor model. Both of the methods find directional interdependence from $\mathcal{Y}$ to $\mathcal{X}$ without any particular difference in terms of value. In fact in Fig. (4) the positive value of ΔG indicates the direction of the information flow from $\mathcal{Y}$ to $\mathcal{X}$. Both measures fail to find any directionality when the two systems are in perfect synchronisation, which is intuitively correct. In this situation it should be impossible to distinguish any directional information flow between the systems.

We repeat the analysis for the coupled logistic map including a nonzero $s = 0.01$ random noise contribution. This value has been chosen so as not to distort the convergent evolution of the systems in (8-9). In Fig (5) we do not have, as in the previous case, perfect synchronisation but instead a 'noisy synchronisation' for $C = 0.45$. From the interdependence analysis in Fig (6) we can still find evidence for directionality in both approaches. Different from the zero-noise case, both methods find a non zero directional values for all studied couplings. However, the values of the two methods settle down to a different ΔG for high values of coupling. In the noisy synchronisation stage their behaviours are different.

B. Henon map

Next, we take the case of a 2D discrete system using the numerical simulations of two unidirectionally coupled Hénon maps which has been studied in several papers [20] [27] [18].

$$\mathcal{X} \quad \begin{cases} x_{i+1} &= 1.4 - x_i^2 + 0.3u_i \\ u_{i+1} &= x_i \end{cases} \tag{10}$$

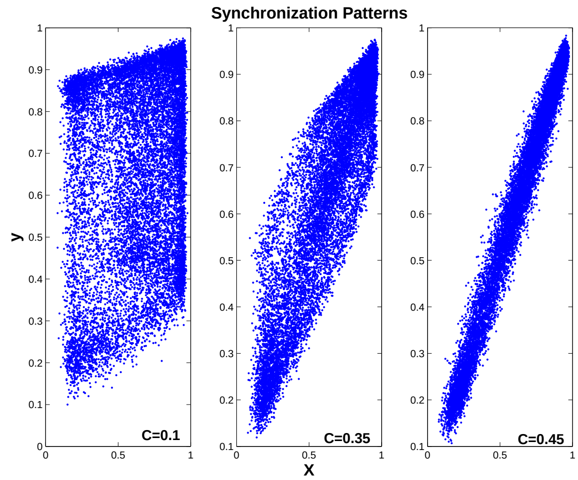


Fig. 5. Same plot of Fig (4) for the case of coupled logistic maps with noise level $s = 0.01$

$$\mathcal{Y} \quad \begin{cases} y_{i+1} &= 1.4 - (Cx_i + (1-C)y_i)y_i + Bv_i \\ v_{i+1} &= y_i \end{cases} \tag{11}$$

Following the literature, the parameter $B = 0.3$ imposes two identical system while $B = 0.1$ is used for non-identical systems [20]. Here, we assume two identical systems. We vary the coupling parameter C between 0 and 1. For $C \simeq 0.8$ we observe the coupled system switching to perfect synchronisation. In Fig. (7) we show the synchronisation pattern for this example.

Using the values $M = 3$, $\tau = 1$ and we construct a predictor using 4 neighbours. These values were selected based on

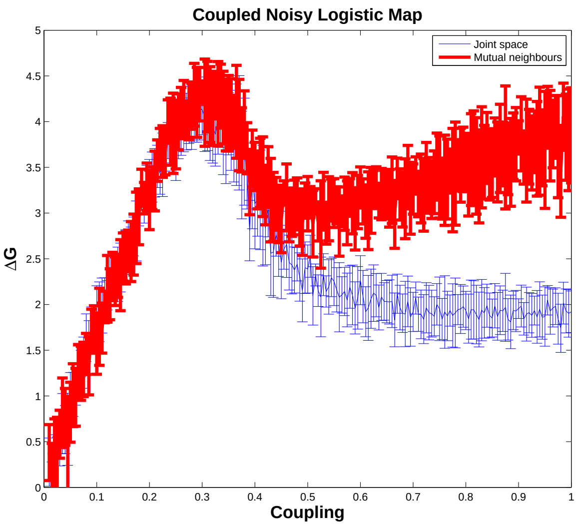


Fig. 6. Granger measure ΔG in the case of coupled logistic noisy map. Directionality can be found in the range of coupling we study. However in this case we notice a difference in the two approaches for high value of coupling.

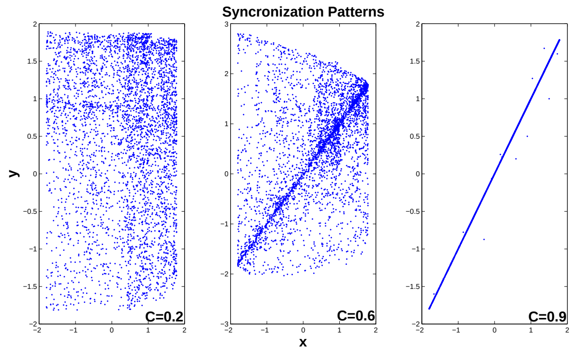


Fig. 7. Henon map Synchronisation pattern x vs y

cross validation. In Fig. (8) we show the differential Granger measure ΔG with respect to the two approaches. In this case, the higher complexity of the map even in the noiseless case, drives a difference in the two approaches for the value reached by ΔG. Importantly, for both cases we can find the correct directionality of the flow.

C. Chain of Tent maps

Our final example of increasing complexity is a chain of L one-dimensional tent maps. The coupling scheme has been studied in [10] and is given by

$$x_{i+1} = f\left(Cx_i^{l-1} + (1-C)x_i^l\right), \tag{12}$$

where l indicates the index of the chosen map in the chain. The tent map is given by

$$f(x) = \begin{cases} 2x & x < 0.5 \\ 2-2x & x \geq 0.5. \end{cases} \tag{13}$$

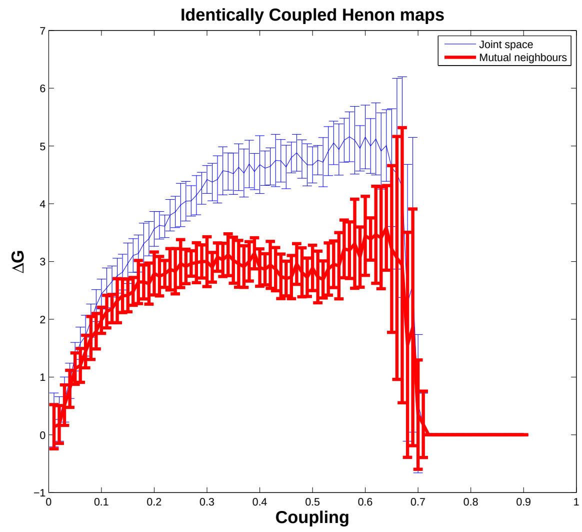


Fig. 8. Directional Granger measure ΔG

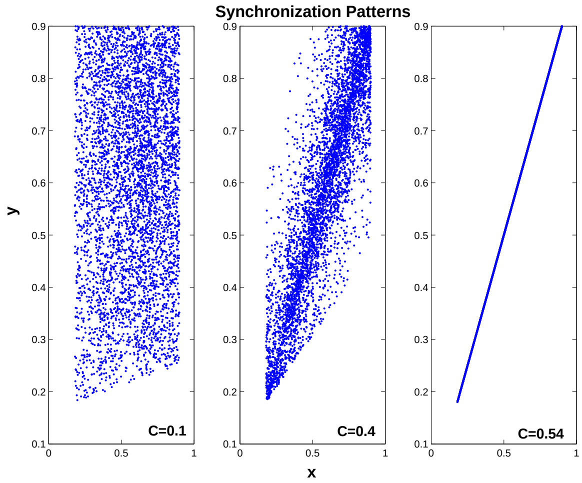


Fig. 9. Chain of tent maps with 2 elements. Synchronisation patterns x vs y.

In the following analysis, we consider the case where the number of maps L is increased from 2 to 5. Using the same paradigm as in the two previous examples, for each value of L we compute the joint space and the mutual neighbours Granger indices between the first and the last elements in the chain. We use $M = 1$, $\tau = 1$ and we construct the Lorenz analogues predictor using 4 neighbours. In Fig. (9), we plot the synchronization patterns between the first and the last element of a chain of $L = 2$ tent elements. We observe a transition to complete synchronization at the value of $C = 0.45$. In Fig. (10), we plot the synchronization pattern for $L = 4$. The transition to the complete synchronization stage is the same for $L = 2, 4$. From the plots in Fig. (9)(10), we notice that the results are less 'sparse' for $L = 2$, prior to complete synchronization.

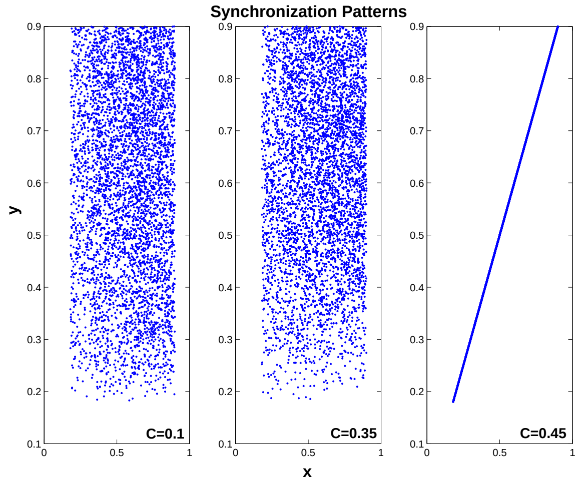


Fig. 10. Chain of tent maps with 4 elements. Synchronisation patterns x vs y.

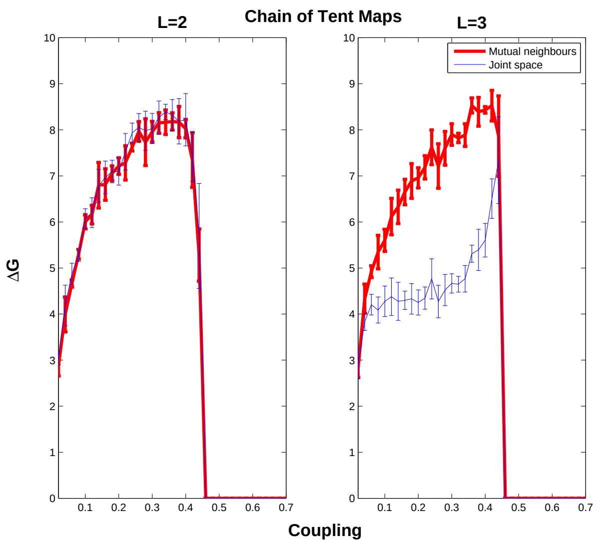


Fig. 11. Directional Granger measure ΔG for the case of $L = 2$ and $L = 3$ tent map elements. Increasing the number of elements results in a decreasing value of the joint space approach Granger index.

In Fig. (11), we show the results of the two differential Granger indices using $L = 2$ and $L = 3$ while Fig. (12) depicts the indices for $L = 4$ and $L = 5$. We notice that the behaviour of the mutual neighbours index does not modify its qualitative behaviour under an increase of the number of tent map elements. On the other hand, the joint space index decreases its value as the length L is changed. Nevertheless, both measures correctly find the true directional interaction in the chain.

IV. Conclusion

The goal of this work was to investigate two different state space approaches to infer dynamical directional interdependence between nonlinear time series of a latent complex system. We based our comparison on the Granger approach given by the mutual predictability using (1). The two techniques we considered were the joint space approach, which is used in the causality literature and the mutual neighbours approach, which has been studied mainly for generalised synchronisation purposes.

In order to check predictability, in this work we employed a model which was equivalent to the Lorenz method of analogues based on a particularly simple form of interpolation of the conditional density estimation. For more general problems, the kernel model will need to be expanded to incorporate more useful kernels leading to an alternative formulation of the analogues approach. We are currently investigating different choices of kernels and their spatial extension in order to study local and global modelling of the reconstructed manifold.

From our qualitative investigation, based on synthetic examples, we conclude that the mutual neighbours approach performs as well as the joint space approach in finding directional interactions but with less computational load. The computational load is due to the choice of the density estimation algorithm which needs a procedure to search for analogues in the reconstructed space. From the examples we studied we have observed differences between the two approaches when the complexity of the underlying system is increased, although both methods correctly estimated the directionality of information flow. An analytical study of the difference between these two approaches is left for a future investigation.

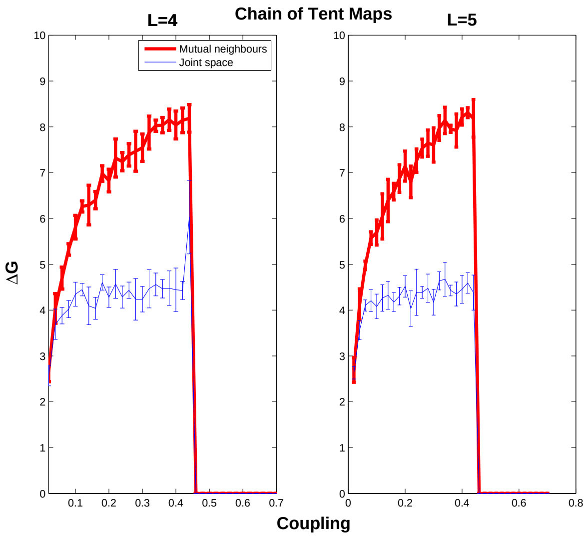


Fig. 12. Same plot as in Fig. (11) with $L = 4$ and $L = 5$.

Acknowledgment

The authors wish to acknowledge EPSRC project EP/D00036X/1 “Exploiting emergent collective behaviours in complex large MEMS systems”, for funding

REFERENCES

[1] C. W. J. Granger, "Testing for causality : A personal viewpoint," *Journal of Economic Dynamics and Control*, vol. 2, no. 1, pp. 329–352, May 1980.

[2] ——, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–38, July 1969.

[3] C. A. Sims, "Money, income, and causality," *American Economic Review*, vol. 62, no. 4, pp. 540–52, September 1972.

[4] Z. A. Ozdemir and E. Cakan, "Non-linear dynamic linkages in the international stock markets," *Physica A: Statistical Mechanics and its Applications*, vol. 377, no. 1, pp. 173–180, Apr. 2007.

[5] B. Gourevitch, R. Bouquin-Jeannes, and G. Faucon, "Linear and non-linear causality between signals: methods, examples and neurophysiological applications," *Biol. Cybern.*, vol. 95, no. 4, pp. 349–369, 2006.

[6] Y. Chen, G. Rangarajan, J. Feng, and M. Ding, "Analyzing multiple nonlinear time series with extended granger causality," *Physics Letters A*, vol. 324, pp. 26–35, Apr. 2004.

[7] N. Ancona, D. Marinazzo, and S. Stramaglia, "Radial basis function approach to nonlinear granger causality of time series," *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, vol. 70, no. 5, p. 056221, 2004.

[8] W. A. Freiwald, P. Valdes, J. Bosch, R. Biscay, J. C. Jimenez, L. M. Rodriguez, V. Rodriguez, A. K. Kreiter, and W. Singer, "Testing non-linearity and directedness of interactions between neural groups in the macaque inferotemporal cortex," *Journal of Neuroscience Methods*, vol. 94, no. 1, pp. 105–119, Dec. 1999.

[9] K. Hlavackova-Schindler, M. Palus, M. Vejmelka, and J. Bhattacharya, "Causality detection based on information-theoretic approaches in time series analysis," *Physics Reports*, vol. 441, no. 1, pp. 1–46, Mar. 2007.

[10] T. Schreiber, "Measuring information transfer," *Phys. Rev. Lett.*, vol. 85, no. 2, pp. 461–464, Jul 2000.

[11] F. Takens, "Detecting strange attractors," in *Lecture Notes in Mathematics*, D. Rand and L. Yang, Eds. Springer, 1981, vol. 898.

[12] L. M. Pecora and T. L. Carroll, "Synchronization in chaotic systems," *Phys. Rev. Lett.*, vol. 64, no. 8, pp. 821–824, Feb 1990.

[13] S. Boccaletti, J. Kurths, G. Osipov, D. Valladares, and C. Zhou, "The synchronization of chaotic systems," *Physics Reports*, vol. 366, pp. 1–101(101), August 2002.

[14] R. Quian Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, "Performance of different synchronization measures in real data: A case study on electroencephalographic signals," *Phys. Rev. E*, vol. 65, no. 4, p. 041903, Mar 2002.

[15] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization : A Universal Concept in Nonlinear Sciences*. Cambridge University Press, April 2003.

[16] N. F. Rulkov, M. M. Sushchik, L. S. Tsimring, and H. D. I. Abarbanel, "Generalized synchronization of chaos in directionally coupled chaotic systems," *Phys. Rev. E*, vol. 51, no. 2, pp. 980–994, Feb 1995.

[17] J. Arnhold, P. Grassberger, K. Lehnertz, and C. Elger, "A robust method for detecting interdependences: application to intracranially recorded eeg," *Physica D*, vol. 134, pp. 419–430(12), 1999.

[18] R. Quiroga, J. Arnhold, and P. Grassberger, "Learning driver-response relationships from synchronization patterns," *Phys. Rev. E*, vol. 61, no. 5, pp. 5142–5148, May 2000.

[19] C. J. Stam and B. W. van Dijk, "Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets," *Phys. D*, vol. 163, no. 3, pp. 236–251, 2002.

[20] S. J. Schiff, P. So, T. Chang, R. E. Burke, and T. Sauer, "Detecting dynamical interdependence and generalized synchrony through mutual prediction in a neural ensemble," *Phys. Rev. E*, vol. 54, no. 6, pp. 6708–6724, Dec 1996.

[21] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge University Press, November 2003.

[22] P. J. Brockwell and R. A. Davis, *Time series: theory and methods*. New York, NY, USA: Springer-Verlag New York, Inc., 1986.

[23] C. M. Bishop, *Pattern Recognition and Machine Learning*, ser. Information Science and Statistics, Springer, Ed. Springer-Verlag New York Inc., 2006.

[24] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms, Second Edition*. The MIT Press, September 2001.

[25] J. Geweke, "Measurement of linear dependence and feedback between multiple time series," *Journal of the American Statistical Association*, vol. 77, no. 378, pp. 304–313, 1982.

[26] J. Geweke, R. Meese, and W. Dent, "Comparing alternative tests of causality in temporal systems : Analytic results and experimental evidence," *Journal of Econometrics*, vol. 21, no. 2, pp. 161–194, Feb. 1983.

[27] T. Kreuz, F. Mormann, R. G. Andrzejak, A. Kraskov, K. Lehnertz, and P. Grassberger, "Measuring synchronization in coupled model systems: A comparison of different approaches," *Physica D: Nonlinear Phenomena*, vol. 225, no. 1, pp. 29–42, Jan. 2007.

Erik Casagrande received a MSc in control engineering in 2003 from Padova university. He completed his studies visiting Lund University and working in image processing and computer vision areas. He then received an MPhil in numerical modelling in 2005 from Padova university. He is currently working towards the PhD degree at Aston university focusing on causality and synchronisation identification analysis. Recent research interest has been in nonlinear time series, dynamical systems, machine learning techniques for biomedical data.

David Lowe had an early background in theoretical physics in the analysis and modelling of extremely small many particle systems at the quantum and semiclassical level, obtaining BSc and PhD degrees from Warwick University, UK. He then worked in signal processing areas in the Defence Research Agency UK and then in statistical pattern processing in areas such as speech, and generic sensor processing. In recent years he has developed expertise in the invention, analysis and application of nonlinear inference machines such as artificial and biological neural networks. He is coinventor of the Radial Basis Function neural network architecture and NeuroScale visualisation machine learning models, and the Shadow Targets optimisation algorithm. Recent research has been in stochastic control, steganography, bioinformatic and biomedical data analysis, brain signal processing, and emergent behaviour in coupled MEMS sensor arrays. He has held a Chair at Aston University, UK since1994.

Avgis Hadjipapas obtained a Medical Degree from the University of Kiel, Germany in 2000, an MS degree in Cognitive Science/Electroencephalography from the Central and Eastern European Center in Cognitive Science (Sofia, Bulgaria) in 2002 and a PhD degree in neuroscience/Magnetoencepahalography from the Wellcome Trust Laboratory for MEG studies (Aston University, UK) in 2005. He has held an appointment as a Research Fellow with the Wellcome Trust Laboratory for MEG studies Aston University, since January 2005. His research interests are in the field of functional neuroimaging in humans, mostly focused in the time series analysis of magneto- and -electroenecephalographic (MEG/EEG) signals, the modeling of the underlying network generators of these macroscopic brain signals and the identification of such generators from measured time series both in health and disease. Avgs Hadjipapas has developed expertise in the noninvasive characterization of local cortical dynamics underlying basic functions of human visual cortex and the assessment of functional /dysfunctional long-range coordination the brain.