A cooperation index based on correlation matrix spectrum and Rényi entropy

Marco Righero*, Oscar De Feo†, Mario Biey*
*Electronic Engineering Dept., Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129, Torino, Italy
E-mail:{marco.righero, mario.biey}@polito.it
†Solianis Monitoring AG, Leutschenbachstrasse 46, CH-8050, Zürich, Switzerland
E-mail: oscar.de-feo@a3.epfl.ch

***Abstract*— Indices based on correlation or more subtle strategies are among the standard ways to infer dependencies (*i.e.,* exchange of information or coupling) in aggregations of different systems observed in the time domain. We propose a new index based on Rényi entropy and confront it with other indices, studying if some of these techniques can recognize when we are observing the same system twice, even when the observation conditions are bad. It turns out that our index gives better results than the other examined ones. Moreover, we notice that those indices based on data processed with state space reconstruction and filtered with Principal Component Analysis are, generally, less sensitive to bad observations. However, state space reconstruction by itself is not enough to obtain good performances when the data are very noisy, and a Principal Component Analysis filter is needed to improve the results.**

I. Introduction

Given multivariate measurements in the time domain, the issue of inferring cooperation (flow of information or coupling) among them has attracted a lot of attention of researchers (see for example [1]–[6]), who proposed many different indices to address it. Here we propose a new index, a simple extension of another one already present in literature [3], and we give some numerical examples which highlight the pros of this novelty and the importance of state space reconstruction with time-delay technique and of Principal Component Analysis (PCA) [7]. In the first block of experiments, a chaotic system is observed twice, with different observation matrices, and noise is added. Several cooperation indices are then computed and compared to see if they are robust against observation differences and noise. Loosely speaking, we confront how close they get to the ideal value of 1, corresponding to the maximum of cooperation—which, theoretically, should always be obtained, as we are looking at the same system, just from different angles. In the second block of experiments, still at a preliminary stage, we look at two coupled systems, to see if the considered indices can scale with respect to the coupling strength.

II. Cooperation estimators

In the following we use this notation: for $t \in [0, T]$ let $\mathbf{x}(t) \in \mathbb{R}^U$ the trajectory of a dynamical system whose evolution is described by the ordinary differential system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$. An observation $y(t) \in \mathbb{R}$ is obtained from $\mathbf{x}(t)$ through the matrix $\mathbf{C} \in \mathbb{R}^{1\times U}$ as $y(t) = \mathbf{C}\mathbf{x}(t)$, $t \in [0, T]$ and a sampled observation $s(\ell)$ with sampling time δt is given by $s(\ell) = y(\ell\delta t)$, $\ell \in \{0, 1, \ldots, L\}$. As we consider more than one experiment, we use a superscript to denote each one, so that $\mathbf{x}^n(t)$ is the trajectory of the experiment n and $\mathbf{x}^n_u(t)$ is its component u, with $n \in \{1, 2, \ldots, N\}$ and $u \in \{1, 2, \ldots, U\}$. We are interested in estimating the cooperation between two different trajectories $\mathbf{x}^1(t)$ and $\mathbf{x}^2(t)$ given two sampled observation of them $s^1(\ell)$ and $s^2(\ell)$ obtained with two unknown and maybe different matrices $\mathbf{C}^1$ and $\mathbf{C}^2$. Without loss of generality, we suppose the observations normalized to have zero mean and unitary variance. As shown in [8], [9], given a scalar sampled observation $s(\ell)$ of a multidimensional system on an attractor $\mathcal{M}$, we can construct a space $\mathbb{R}^V$ such that $\mathcal{M}$ is embedded into $\mathbb{R}^V$ with the technique of time delay. With time delay τ and dimension V, the multidimensional series reconstructed in this way is $\mathbf{z}(\ell)$, where each element $\mathbf{z}(\ell)$ is given by $\mathbf{z}(\ell) = [s(\ell)\ s(\ell + \tau)\ \ldots\ s(\ell + (V - 1)\tau)]$ for $\ell \in \{1, \ldots, L - (V - 1)\tau\}$. Determining good values for the parameters V and τ is a difficult task. We choose to use, respectively, the first zero of the mutual information function and the false nearest neighbor technique [9] as implemented in the TISEAN routine [10]. Given $s^n(\ell), \ldots, s^N(\ell)$ N sampled observations (time series) let $\mathbf{P} \in \mathbb{R}^{N\times N}$ the correlation matrix whose elements $\mathbf{P}_{ij}$ are

$$\mathbf{P}_{ij} = \frac{\operatorname{Cov}\left(s^i, s^j\right)}{\sqrt{\operatorname{Cov}\left(s^i, s^i\right) \cdot \operatorname{Cov}\left(s^j, s^j\right)}}, \tag{1}$$

where Cov () is the covariance.

A. *S and its time-delay variants*

In [2], [3], the authors introduce an index to measure cooperation among cluster of subsystems from time series. We briefly summarize how it works. Let $\{\lambda_1, \ldots, \lambda_N\}$ the eigenvalues of the correlation matrix $\mathbf{P}$ Eq. (1) of N time series whose cooperation we are interested in. The Shannon entropy-like quantity

$$I = -\sum_{n=1}^{N} \frac{\lambda_n}{N} \log\left(\frac{\lambda_n}{N}\right) \tag{2}$$

is inversely proportional to the amount of cooperation among the N time series and ranges in $[0, \log(N)]$, so that

$$S = 1 - I/\log(N) \tag{3}$$

is a normalized index of cooperation among $s^1(\ell), \dots, s^N(\ell)$. The index S can be extended to work with delay-embedded time series too, with an appropriate partitioning of the correlation matrix and some minor modifications [2], [3]. In the following we consider just two time series, $(N = 2)$ and denote by S the index applied to these time series and with S_{Emb} the S index applied to their delay-embedded version. With S_{PCA} we denote the index S applied to the delay-embedded version of the two time series where the reconstructed space is firstly increased (we take $2\tau V$) with delay fixed to 1 and then reduced using the PCA technique [7] as a filtering procedure.

B. The C index and some considerations on S

As a comparison, we consider a simple cooperation index between two time series $s^1(\ell)$ and $s^2(\ell)$ given by the absolute value of the correlation and call it C. As we consider normalized time series, we have that $C(s^1, s^2) = \left|\mathrm{Cov}\left(s^1, s^2\right)\right|$. Differently from S, this index can neither deal with clusters of more than two indices nor with delay-embedded time series. We can explicitly write the relation of C and S with the covariance of s^1 and s^2, say $\rho = \mathrm{Cov}\left(s^1, s^2\right)$, in an analytic way. In Fig. 1 we plot S and C as a function of ρ. As both C and S are even, we consider only the values $\rho \in [0, 1]$. We notice than the index S is very steep near the value 1, so that when dealing with highly correlated time series every little difference on the ρ axis is amplified. Moreover, the value of S is always lower than C. We would like to consider an index with properties similar to those of S (ability to analyze clusters and delay-embedded time series) but with values nearer to C, to fully exploit its intrinsic information.

C. S^α: S with Rényi entropy

The S index of N time series $s^1(\ell), \dots, s^N(\ell)$ is based on Shannon entropy of the spectrum of the correlation matrix $\mathbf{P}$ Eq. (1). We generalize it using Rényi entropy I^α [11], defining, for the given $\lambda_1, \dots, \lambda_N$ eigenvalues of $\mathbf{P}$,

$$I^\alpha = \frac{1}{1-\alpha} \log\left(\sum_{n=1}^{N} \left(\frac{\lambda_n}{N}\right)^\alpha\right). \tag{4}$$

As Rényi entropy of order α converges to Shannon entropy for α going to 1, we have that I^α converges to I defined in Eq. (2). The quantity defined in Eq. (4) is inversely proportional to the amount of cooperation among the N time series and ranges in $[0, \log(N)]$, so that

$$S^\alpha = 1 - I^\alpha/\log(N) \tag{5}$$

is a normalized index of cooperation among $s^1(\ell), \dots, s^N(\ell)$. The extra degree of freedom provided by α allows us to fine tune the index to be nearer the value

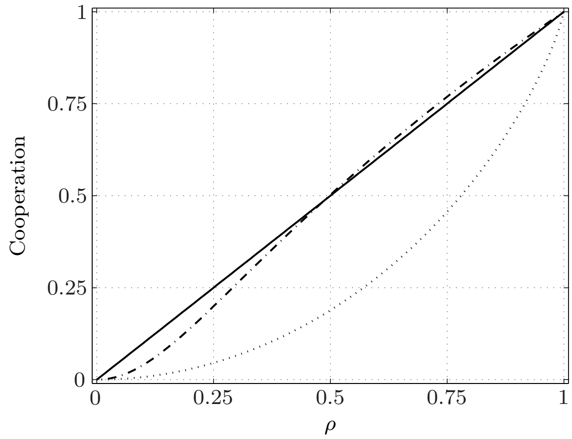

Fig. 1. The indices S (*dotted line*, Eq. (3)), C (*solid line*) and S^6 (*dash-dotted line*, Eq. (5)) as functions of ρ, the correlation between two time series.

of C. For $\alpha = 6$, analytically computing S^α as a function of ρ, as we have previously done for S, we obtain the plot shown in Fig. 1 with a dash-dotted line. The value $\alpha = 6$ has been obtained by empirical trial and error and we fix this value for all the following experiments, even if further analyses on the influence of this parameter are needed. This index can be extended to delay-embedded time series as S, so we denote with S^α_{Emb} and S^α_{PCA} the counterpart of S_{Emb} and S_{PCA} using Rényi entropy of order α instead of Shannon entropy.

D. R, a mutual information based index

Indices based on correlation are generally easy to compute, involving highly optimized matrix operations, but consider just second order statistics. On the opposite, mutual information [12], albeit difficult to estimate, exploits all the moments. As a comparison, we consider also the normalized redundancy R of two time series $s^1(\ell)$ and $s^2(\ell)$ defined as

$$R(s^1, s^2) = \frac{MI(s^1, s^2)}{\min\left(H(s^1), H(s^2)\right)}, \tag{6}$$

where $H(s)$ is the entropy of s and $MI(s^1, s^2) = H(s^1) + H(s^2) - H(s^1, s^2)$ is the mutual information between s^1 and s^2 [12]. We denote with $R_{Emb}(s^1, s^2)$ and R_{PCA} the redundancy index applied to delay-embedded and delay-embedded and PCA filtered time series, respectively.

III. Experimental setup

A. Observing the same system twice

The first experiment we perform is to observe the same chaotic system, whose trajectory is $\mathbf{x}(t)$, through two different matrices, to corrupt the observed time series with noise and to apply the previous indices to study cooperativeness between them. In an ideal setting, we should obtain the maximum value (*i.e.,* 1) regardless of the matrices and the noise, as we are observing the same system. To study how, in reality, the observation setup affects the performance of the indices, we observe the system with a fixed matrix $\mathbf{C}^1$ and then with a matrix $\mathbf{C}^2$ depending on a parameter $p \in [0,1]$ such that for $p=1$ $\mathbf{C}^2 = \mathbf{C}^1$, while for $p \neq 1$ $\mathbf{C}^1 \neq \mathbf{C}^2$. We focus on three systems: The Colpitts oscillator, [13] Eq. (7), the Lorenz system, [14] Eq. (8) left, and the Rössler system, [15] Eq. (8) right

$$\begin{cases} \dot{\mathbf{x}}_1 = \dfrac{g}{Q(1-k)}\mathbf{x}_3 + \dfrac{\alpha g}{Q(1-k)}\left(1-\mathrm{e}^{-\mathbf{x}_2}\right), \\ \dot{\mathbf{x}}_2 = \dfrac{g}{Qk}\mathbf{x}_3 + \dfrac{-(1-\alpha)g}{Qk}\left(1-\mathrm{e}^{-\mathbf{x}_2}\right), \\ \dot{\mathbf{x}}_3 = -\dfrac{Qk(1-k)}{g}(\mathbf{x}_1+\mathbf{x}_2) - \dfrac{1}{Q}\mathbf{x}_3, \end{cases} \tag{7}$$

with $k=0.5$ and $\alpha = 0.996$ and the parameters Q, g uniformly chosen in an interval of relative amplitude 10% around the values $10^{0.15}$ and $10^{0.625}$, respectively;

$$\begin{cases} \dot{\mathbf{x}}_1 = \sigma(\mathbf{x}_2 - \mathbf{x}_1), \\ \dot{\mathbf{x}}_2 = \rho\mathbf{x}_1 - \mathbf{x}_2 - \mathbf{x}_1\mathbf{x}_3, \\ \dot{\mathbf{x}}_3 = \mathbf{x}_1\mathbf{x}_2 - \beta\mathbf{x}_3, \end{cases} \quad \begin{cases} \dot{\mathbf{x}}_1 = \mathbf{x}_2 - \mathbf{x}_3, \\ \dot{\mathbf{x}}_2 = \mathbf{x}_1 + a\mathbf{x}_2, \\ \dot{\mathbf{x}}_3 = b + \mathbf{x}_1(\mathbf{x}_1 - c), \end{cases} \tag{8}$$

with ρ, σ and β uniformly chosen in an interval with relative amplitude of 10% around the values $10, 27$ and $8/3$, respectively, and a, b and c uniformly chosen in an interval with relative amplitude of 10% around the values $0.1, 0.1$ and 14, respectively. For each system, 100 histories are generated in the following way: (a) Initial conditions and system parameters are randomly generated; (b) The system is integrated using MatLab standard routines. Transient ($[0, 200\pi]$ for Colpitts system, $[0, 100]$ for Lorenz system and $[0, 500]$ fo Rössler) is discarded; (c) The system is integrated again, using the last point of the previous step as the new initial condition, and data are sampled to generate time series of 1000 points. Time span and sampling time are, respectively, $[0, 20\pi], 2\pi/100$ for the Colpitts systems, $[0, 10], 0.01$ for the Lorenz systems and $[0, 5], 0.05$ for the Rössler systems. Each sampled trajectory is then observed through the matrices $\mathbf{C}^1 = [1\ 0\ 0], \mathbf{C}^2 = [p\ (1-p)\ 0]$; $\mathbf{C}^1 = [1\ 0\ 0], \mathbf{C}^2 = [p\ 0\ (1-p)]$; $\mathbf{C}^1 = [0\ 0\ 1], \mathbf{C}^2 = [0\ p\ (1-p)]$; with $p \in [0,1]$. The time series $s^1 = \mathbf{C}^1\mathbf{x}$ and $s^2 = \mathbf{C}^2\mathbf{x}$ are then corrupted with gaussian noise. We examine the cases where the SNR in dB is $40, 20, 12$, and 6. The cooperation between the two time series is evaluated with the 10 indices previously described. We summarize the results in Tab. (I) and Tab. (II), where we present the mean and the standard deviation of the indices against the SNR. To compute these figures, we merged together the observations obtained from histories generated as described in (a) – (c) for the different systems of Eq. (7) and Eq. (8) using different matrices and different values of p, in order to simulate a blind setup.

B. Observing two coupled systems

In this second block of experiments, we consider a Rössler system linearly driving a Lorenz one, so that the overall system $[\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3\ \mathbf{x}_4\ \mathbf{x}_5\ \mathbf{x}_6]'$ ($'$ denotes transpose) obeys

$$\begin{cases} \dot{\mathbf{x}}_1 = T\left[\mathbf{x}_2 - \mathbf{x}_3\right],\ \dot{\mathbf{x}}_2 = T\left[\mathbf{x}_1 + a\mathbf{x}_2\right], \\ \dot{\mathbf{x}}_3 = T\left[b + \mathbf{x}_1(\mathbf{x}_1 - c)\right],\ \dot{\mathbf{x}}_4 = \sigma(\mathbf{x}_5 - \mathbf{x}_4), \\ \dot{\mathbf{x}}_5 = \rho\mathbf{x}_4 - \mathbf{x}_5 - \mathbf{x}_4\mathbf{x}_6 + \xi(\mathbf{x}_2 - \mathbf{x}_5), \\ \dot{\mathbf{x}}_6 = \mathbf{x}_4\mathbf{x}_5 - \beta\mathbf{x}_6, \end{cases} \tag{9}$$

with ξ coupling parameter ranging in $[0,1]$ $\rho = 10$, $\sigma = 27$, $\beta = 8/3$, $a = 0.1$, $b = 0.1$, $c = 14$, and $T = 6$ to adjust the different speed of the Rössler system compared to the Lorenz one. For different random initial conditions, we integrate Eq. (9) using MatLab standard routines. The transient ($[0, 100]$) is discarded and then the system is integrated again and data are sampled to generate time series of 1000 points. Time span is $[0, 10]$ and sampling time 0.01. We generate 20 histories for each of 6 equally spaced values of ξ in $[0,1]$ and then observe the trajectories through the matrices $\mathbf{C}^1 = [0\ 1\ 0], \mathbf{C}^2 = [1-p\ p\ 0]$; $\mathbf{C}^1 = [0\ 1\ 0], \mathbf{C}^2 = [0\ p\ 1-p]$; $\mathbf{C}^1 = [1\ 0\ 0], \mathbf{C}^2 = [p\ 0\ 1-p]$; with $p \in \{0, 0.5, 1\}$. The time series $s^1 = \mathbf{C}^1[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]'$ and $s^2 = \mathbf{C}^2[\mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6]'$ are then corrupted with gaussian noise such that the SNR in dB is 6. The cooperation between the two time series is evaluated with the indices C, S, S^6 and R previously described. In this preliminary experiment, we focus on non embedded time series, just to see if the indices can scale with the coupling or not. The embedding procedure brings other aspects to be further studied. We summarize the results in Fig. 2, where we present the mean of the indices against the coupling, merging together the observations for different matrices and different values of p to simulate a blind setup.

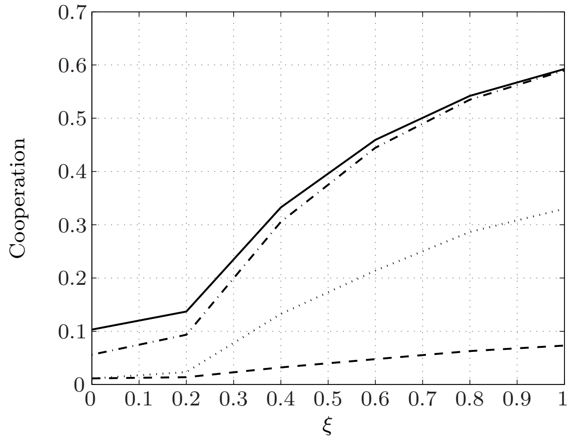

Fig. 2. Mean of the indices S (*dotted line*, Eq. (3)), C (*solid line*), S^6 (*dash-dotted line*, Eq. (5)) and R (*dashed line*, Eq. (6)) with respect to the coupling ξ.

IV. Results and discussion

A. Observing the same system twice

From Tab. (I) and Tab. (II) we see that the indices based on Rényi entropy can obtain better values than

SNR (dB)	C	S	S_{Emb}	S_{PCA}	R	R_{Emb}	R_{PCA}	S^6	S^6_{Emb}	S^6_{PCA}
40	0.798448	0.666621	0.655491	0.744656	0.397295	0.875920	0.922469	0.798674	0.845025	0.890319
20	0.790627	0.644281	0.434230	0.651020	0.228836	0.619423	0.761979	0.791520	0.732159	0.850599
12	0.751163	0.555550	0.295870	0.547210	0.135950	0.378541	0.539220	0.754895	0.632231	0.798816
6	0.638738	0.372483	0.206693	0.450592	0.070684	0.212562	0.391126	0.645859	0.538725	0.745480

TABLE I

Mean of cooperation indices for different SNR (dB)

SNR (dB)	C	S	S_{Emb}	S_{PCA}	R	R_{Emb}	R_{PCA}	S^6	S^6_{Emb}	S^6_{PCA}
40	0.282172	0.336523	0.230555	0.224007	0.225174	0.081680	0.055770	0.295574	0.184844	0.157606
20	0.279412	0.322632	0.165724	0.212705	0.104591	0.084770	0.131117	0.293323	0.172231	0.160062
12	0.265442	0.272811	0.106298	0.198019	0.058904	0.066050	0.158097	0.281776	0.162027	0.162371
6	0.225070	0.178392	0.072710	0.168480	0.030152	0.044441	0.133228	0.245935	0.145833	0.161864

TABLE II

Standard deviation of cooperation indices for different SNR (dB)

those based on Shannon entropy and comparable or better than those based on mutual information, which require a bigger computational effort to be evaluated. The indices working with delay-embedded time series without filtering have poor performances as the SNR decreases, probably due to the fact that the reconstructed space is made up of uncorrelated noise for a great part. However, when coupled with PCA filtering, the delay-embedding technique can improve the results and, at the end, we see that the index S^6_{PCA} is able to recognize a high value of cooperation with low standard deviation even with SNR of 6 dB, outperforming all the other indices.

B. Observing two coupled systems

From Fig. 2, even if it is a preliminary and incomplete result, we see that the index S^6 is able to scale with respect the coupling strength in a way more similar to the correlation than the index S. As an advantage on the simple correlation, on the other hand, it can process delay-embedded time series and clusters of systems. Delay-embedded version and other values for the parameter α need to be tested.

V. Conclusion

We presented a slightly modified version of a cooperation index known in literature as S estimator [2], [3]. Our version exploits the generalization of Shannon entropy to Rényi entropy to behave in a way more similar to the simple correlation, which is indeed the starting point for both indices. As an advantage on the correlation, however, it can process time-embedded time series and work with clusters of systems like the original S index.

References

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

[2] O. De Feo and C. Carmeli, "Identifying dependencies among multivariate time series," in *Preceeding of the 2004 NOLTA, Fukuoka (JAPAN)*, Nov. 29, Dec. 3 2004, pp. 203 – 206.

[3] C. Carmeli, M. G. Knyazeva, G. Innocenti, and O. De Feo, "Assessment of EEG synchronization based on state-space analysis," *Neuroimage*, vol. 25, pp. 339 – 354, 2005.

[4] S. Frenzel and B. Pompe, "Partial mutual information for coupling analysis of multivariate time series," *Physical Review Letters*, vol. 99, no. 20, pp. 204 101/(1 – 4), Nov. 2007.

[5] M. Jalili, S. Lavoie, P. Deppen, R. Meuli, K. Q. Do, M. Cuénod, M. Hasler, O. De Feo, and M. G. Knyazeva, "Dysconnection topography in schizophrenia revealed with state-space analysis of EEG," *PLoS ONE*, vol. e1059, no. 10, pp. 1 – 13, Oct. 2007.

[6] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Physical Review E*, vol. 69, no. 6, p. 066138, June 2004.

[7] I. T. Jolliffe, *Principal Component Analysis*. Springer, NY, 2002.

[8] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, Warwick 1980*. Springer Berlin / Heidelberg, 1981, vol. 898, pp. 366 – 381.

[9] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, ser. Cambridge Nonlinear Science Series. Cambridge University Press, 2004, no. 7.

[10] R. Hegger, H. Kantz, and T. Schreiber, "Practical implementation of nonlinear time series methods: The TISEAN package," *Chaos*, vol. 9, no. 413, pp. 413–435, 1999.

[11] A. Rényi, "On measures of information and entropy," in *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, 1960, pp. 547 – 561.

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Edition*. Wiley New york, 2006.

[13] O. De Feo, G. M. Maggio, and M. P. Kennedy, "The Colpitts oscillators: Families of periodic solutions and their bifurcations," *International Journal of Bifurcation and Chaos*, vol. 10, no. 2, pp. 935 – 958, 2000.

[14] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of Atmospheric Sciences*, vol. 20, no. 2, pp. 130 – 141, 1963.

[15] O. E. Rössler, "An equation for continuous chaos," *Physics Letters A*, vol. 57, no. 5, pp. 397 – 398, 1976.