Chapter 1
Comparison of Resampling Techniques for the Non-Causality Hypothesis

Angeliki Papana, Catherine Kyrtsou, Dimitris Kugiumtzis and Cees G.H. Diks

Abstract Different resampling schemes for the null hypothesis of non-causality are assessed. As test statistic the partial transfer entropy (PTE), an information and model-free measure, is used. Two resampling methods, 1) the time shifted surrogates and 2) the stationary bootstrap, are combined with the following three independence settings (giving in total six resampling schemes), all consistent to the null hypothesis of non-causality: A) only the driving variable is resampled, B) both the driving and response variable are resampled, and C) both the driving and response variable are resampled while also the dependence of the future of the response variable and the vector of its past values is destroyed. The empirical null distribution of the PTE as the surrogate and bootstrapped time series become more independent is examined along with the size and power of the respective tests.

1.1 Introduction

Resampling techniques are utilized for the construction of the empirical null distribution of a test statistic, when the asymptotic distribution cannot be established. We

Angeliki Papana
University of Macedonia, 156 Egnatias street, 54006 Thessaloniki, Greece, e-mail: angeliki.papana@gmail.com

Catherine Kyrtsou
University of Macedonia, Greece; University of Strasbourg, BETA; University of Paris 10, Economix, ISC-Paris, Ile-de-France, e-mail: ckyrtsou@uom.gr

Dimitris Kugiumtzis
Aristotle University of Thessaloniki, Department of Electrical and Computer Engineering, 54124 Thessaloniki, Greece, e-mail: dkugiu@auth.gr

Cees G.H. Diks
University of Amsterdam, Center for Nonlinear Dynamics in Economics and Finance (CeNDEF), Roetersstraat 11, NL - 1018 WB Amsterdam, The Netherlands, e-mail: C.G.H.Diks@uva.nl

are concerned with the inter-dependence structure of multivariate time series. The generated resampled time series have to capture statistical properties of the original time series but also satisfy the corresponding null hypothesis, H_0, of no inter-dependence between two time series in the presence of the other variables [7]. Bootstrapping, first introduced in [1], aims at estimating the properties of a test statistic when sampling from an approximating distribution. For time series, bootstraps must be carried out in a way that they suitably capture the dependence structure of the data generation process consistent to the H_0, and be otherwise random e.g. [6, 12, 11]. Similarly, randomization methods preserve the dependence structure consistent to H_0 when randomly shuffling the time series [15, 13, 3].

The transfer entropy (TE) is a non-parametric measure that quantifies the amount of directed transfer of between two random processes [14]. The TE is a non-symmetrical measure defined in terms of transition probabilities that provides information about the direction of the dependencies. The TE from a process X to another process Y is the amount of additional information (reduction of uncertainty) about the future values of Y provided by knowing past values of X and Y instead of past values of Y alone. The advantages of TE are that it is model-free, makes no assumptions about the distribution of the data and is effective in case of non-linear signals. The partial transfer entropy (PTE) is a multivariate extension of the TE [16, 9].

Resampling techniques are utilized for the H_0 of non-causality, i.e. no causal effects from one variable (driver) to another one (response), in the presence of the remaining observed variables (confounding variables). A suitable statistic, sensitive to the inter-dependence of the time series, is the PTE. The causality test is actually a significance test for the PTE and in the absence of asymptotic distribution for the PTE, resampling is required.

The appropriateness of six resampling schemes for the null hypothesis H_0 of non-causality is examined. Specifically, we combine two resampling methods: 1) the time shifted surrogates [13] and 2) the stationary bootstrap [11], with three independence settings of the time series adapted for the non-causality test: A) resampling only the time series of the driving variable, B) resampling separately the driving and the response time series, and C) resampling separately the driving and the response time series, while destroying the dependence of the future and past of the response variable. The properties of the test for the six resampling schemes, i.e. the empirical distribution of PTE, the size and power of the test, are assessed in a simulation study.

The structure of the paper is as follows. In Sec. 1.2, the PTE is briefly discussed and in Sec. 1.3 the resampling methods and the independence settings are presented. In Sec. 1.4, the resampling schemes are evaluated with means of simulations on different coupled and uncoupled multivariate systems. The conclusions are drawn in Sec. 1.5.

1.2 Partial Transfer Entropy

The partial transfer entropy (PTE) is a multivariate information measure [16, 9], introduced as an extension of the bivariate measure of transfer entropy (TE) [14]. The TE quantifies the amount of information explained in a response variable Y at h time steps ahead from the state of a driving variable X accounting for the concurrent state of Y. Let $\{x_t, y_t\}$, $t = 1,\ldots,n$ be the observed time series of two variables, and $\mathbf{x}_t = (x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau})'$ and $\mathbf{y}_t = (y_t, y_{t-\tau}, \ldots, y_{t-(m-1)\tau})'$ the reconstructed state space vectors, where m is the embedding dimension and τ is the time delay. The TE from X to Y is the conditional mutual information $I(y_{t+h}; \mathbf{x}_t | \mathbf{y}_t)$:

$$\begin{aligned}\mathrm{TE}_{X\to Y} &= I(y_{t+h}; \mathbf{x}_t | \mathbf{y}_t) = \sum p(y_{t+h}, \mathbf{x}_t, \mathbf{y}_t) \log \frac{p(y_{t+h}|\mathbf{x}_t, \mathbf{y}_t)}{p(y_{t+h}|\mathbf{y}_t)} \\ &= H(\mathbf{x}_t, \mathbf{y}_t) - H(y_{t+h}, \mathbf{x}_t, \mathbf{y}_t) + H(y_{t+h}, \mathbf{y}_t) - H(\mathbf{y}_t),\end{aligned} \tag{1.1}$$

where TE is given either based on probability distributions ($p(x)$ is the probability mass function of the discretized variable x) or entropy terms ($H(\mathbf{x}) = -\int f(\mathbf{x}) \log f(\mathbf{x}) \mathrm{d}\mathbf{x}$ is the differential entropy of the vector variable $\mathbf{x}$ with probability density function $f(\mathbf{x})$).

The PTE accounts for the direct coupling of X to Y conditioning on the confounding variables of a multivariate system, collectively denoted Z. It is given by

$$\begin{aligned}\mathrm{PTE}_{X\to Y|Z} &= I(y_{t+h}; \mathbf{x}_t | \mathbf{y}_t, \mathbf{z}_t) \\ &= H(\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t) - H(y_{t+h}, \mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t) + H(y_{t+h}, \mathbf{y}_t, \mathbf{z}_t) - H(\mathbf{y}_t, \mathbf{z}_t).\end{aligned} \tag{1.2}$$

The estimation of the TE and PTE relies upon the estimation of the probability density functions. Different types of estimators exist, such as histogram-based (e.g. by discretizing the state space to equidistant intervals at each axis), kernel-based and using correlation sums. Here, we use the nearest neighbor estimator [5], which is proved to be effective especially for high-dimensional data [17].

Theoretically, the PTE (and TE) should be zero in the case of no causal effects. However, a bias can be present due to various reasons, e.g. the estimation method for the entropies and subsequently densities, the selection of the embedding parameters, the finite sample size and the noise level as well [8]. In order to determine whether a PTE value indicates a weak coupling or whether it is not statistically significant, a resampling method should be used to assess its statistical significance.

1.3 Resampling Techniques

We first present the two resampling methods of time shifted surrogates and stationary bootstrap, and then the three independence settings. Time shifted surrogates preserve the dynamics of a time series $\{x_1, \ldots, x_n\}$, while the couplings are destroyed [13]. They are formed by cyclically time shifting the components of

$\{x_1,\ldots,x_n\}$. To formulate them from a time series with length n, an integer d is randomly chosen ($d < n$) and the d first values of the time series are moved to the end: $\{x_t^*\} = \{x_{d+1},\ldots,x_n,x_1,\ldots,x_d\}$. For testing $X \to Y$ in a bivariate time series, the pair $\{x_t^*, y_t\}$ is consistent with the H_0 of non-causality.

The stationary bootstrap has been proposed for the calculation of standard errors and the construction of confidence intervals for a statistic based on weakly dependent stationary observations [11]. The bootstrap series are generated by resampling blocks of random size, where the length of each block has a geometric distribution. For a fixed probability p, block lengths L_i are generated with probability $p(L_i = k) = (1-p)^{(k-1)}$ for $k = 1,2,\ldots$. The starting time points of the blocks I_i are drawn from the discrete uniform distribution on $\{1,\ldots,n-k\}$. A bootstrap time series $\{x_t^*\}$ is formed by first starting with a random block as defined above $B_{I_1,L_1} = \{x_{I_1}, x_{I_1+1},\ldots,x_{I_1+L_1-1}\}$, and blocks are added until length n is reached.

Three independence settings are considered for both resampling methods, all consistent with the H_0 of non-causality from X to Y conditioned on Z. The first setting, denoted A, is to resample only the time series of the driving variable X. This is the standard approach for surrogate test for the significance of causality measures [13, 2, 10]. The intrinsic dynamics of the variable X is preserved in the resampled time series $\{x_t^*\}$ but the coupling between X^* and Y is destroyed, so that H_0 is fulfilled and $\mathrm{PTE}_{X^* \to Y|Z} = 0$. The variables X and Y as well as X and Z are independent, however the pair of variables (Y,Z) preserves its interdependence.

The second setting, denoted B, suggests to randomize both the driving variable X and the response Y, i.e. resampled time series $\{x_t^*\}$ and $\{y_t^*\}$ are generated. Again, the intrinsic dynamics of both X and Y are preserved but the coupling between them is destroyed, H_0 is fulfilled and $\mathrm{PTE}_{X^* \to Y^*|Z} = 0$. In this case, independence holds for all variable pairs (X,Y), (Y,Z) and (X,Z). However, there is still no complete independence between all arguments in the definition of PTE, as y_{t+h} preserves by construction of $\{y_t^*\}$ its dependence on $\mathbf{y}_t$.

Finally, we consider the third setting of complete independence of all variables involved in the definition of PTE, denoted C, i.e. in addition to the resampling of X and Y, also y_{t+h} is resampled separately. Thus all terms in PTE, i.e. y_{t+h}, $\mathbf{x}_t$, $\mathbf{y}_t$ and $\mathbf{z}_t$ are independent, and H_0 is again fulfilled.

Combining the two resampling methods (time shifted surrogates and stationary bootstraps) and the three independence settings (A, B and C), six resampling schemes are formulated that are utilized to test the null hypothesis of no causal effects among the variables of multivariate systems.

1.4 Simulation study

In the simulation study we apply the significance test for the PTE with the six resampling schemes to multiple realizations of different simulation systems. Specifically, we estimate the PTE from 100 realizations per simulation system. For each realization and each resampling scheme, $M = 100$ resampled time series are generated. Let us denote q_0 the PTE value from one realization of a system and $q_1, q_2, \dots, q_M$ the PTE values from the resampled time series for this particular realization and for a specific resampling scheme. The rejection of H_0 of no causal effects is decided by the rank ordering of the PTE values computed on the original time series, q_0, and the resampled time series, $q_1, q_2, \dots, q_M$. For the one-sided test, if r_0 is the rank of q_0 when ranking the list $q_0, q_1, \dots, q_M$ in ascending order, the p-value of the test is $1-(r_0-0.326)/(M+1+0.348)$, by applying the correction in [19].

The simulation systems that have been used in this study are the following:

1. Three coupled Hénon maps, with nonlinear couplings ($X_1 \to X_2$, $X_2 \to X_3$) (System 5 in [10]) with equal coupling strengths c for $X_1 \to X_2$ and $X_2 \to X_3$. We set $c = 0$ (uncoupled case), $c = 0.3$ and $c = 0.5$ (strong coupling). The Hénon map is a well-known discrete-time dynamical system that exhibits chaotic behavior [4].
2. A vector autoregressive process of 4 variables and order 5, VAR(5), with linear couplings ($X_1 \to X_3$, $X_2 \to X_1$, $X_2 \to X_3$, $X_4 \to X_2$) (Eq.(12) in [18]).
3. Five coupled Hénon maps, with nonlinear couplings ($X_1 \to X_2$, $X_2 \to X_3$, $X_3 \to X_4$, $X_4 \to X_5$) defined similarly to system 1. We consider again equal coupling strengths c, and set $c = 0$ (uncoupled case), $c = 0.2$ and $c = 0.4$ (strong coupling).

We consider time series lengths $n = 512$ and 2048. To estimate the PTE, we set the embedding dimension m to appropriate values for each system, i.e. $m = 2$ for system 1 and 3, $m = 5$ for system 2, the delay time $\tau = 1$ and the time step ahead $h = 1$ (as defined in [14]). The number of nearest neighbors for the estimation of the probability distributions is 10.

In terms of presentation, we focus on the sensitivity of the PTE (percentage of rejection of H_0 when there is true direct causality), as well as the specificity of the PTE (percentage of no rejection of H_0 when there is no direct causality), at the significance level $\alpha = 0.05$. The notation $X_2 \to X_1|Z$ denotes the Granger causality from X_2 to X_1, accounting for the presence of confounding variables $Z = X_3, \dots$. For brevity, we use the notation $X_2 \to X_1$ instead of $X_2 \to X_1|Z$, implying the conditioning on the confounding variables. The notation of Granger causality for other pairs of variables is analogous.

System 1. The mean PTE values are negatively biased in the uncoupled case ($c = 0$). For $c = 0.3$ and $c = 0.5$, they are much larger when direct couplings exist ($X_1 \to X_2$, $X_2 \to X_3$) than the rest of the directions, and increase with n. Regarding the indirect coupling $X_1 \to X_3$, PTE increases with n for $c = 0.5$ (mean $PTE_{X_1 \to X_3} = -0.0002$ for $n = 512$ and 0.0075 for $n = 2048$).

We evaluate how the null distribution of the PTE from the six resampling schemes differs with respect to the original PTE values. For $c = 0$, all the resampling schemes correctly indicate the absence of couplings; the percentages of significant PTE values vary from 0% to 12% (see Table 1.1). Considering $c = 0.3$, the schemes B and C indicate correctly the couplings, while scheme A indicates the spurious one $X_2 \to X_1$ and the indirect one $X_1 \to X_3$. The percentage of erroneously rejected H_0 for non-existing or indirect couplings tends to increase with c and the time series length for all resampling schemes, the most robust being 1C and 2C.

Table 1.1 Percentage of significant PTE values for system 1 for $n = 512$ / 2048, for the six resampling schemes. The directions of true couplings are highlighted. A single number is displayed when the same percentage corresponds to both n.

$c=0$	$X_1 \to X_2$	$X_2 \to X_1$	$X_2 \to X_3$	$X_3 \to X_2$	$X_1 \to X_3$	$X_3 \to X_1$
1A	2 / 3	3 / 9	3 / 5	3 / 5	6 / 4	4 / 10
1B	4 / 1	5 / 12	4 / 5	3 / 4	4 / 6	5 / 10
1C	1 / 0	0	2 / 0	0	1 / 0	1 / 0
2A	2 / 1	3 / 7	3 / 5	1 / 2	6 / 5	3 / 12
2B	2 / 0	1	1 / 0	1 / 0	4 / 1	1 / 3
2C	0	0	0	0	2 / 0	0
$c=0.3$	$\mathbf{X_1 \to X_2}$	$X_2 \to X_1$	$\mathbf{X_2 \to X_3}$	$X_3 \to X_2$	$X_1 \to X_3$	$X_3 \to X_1$
1A	**100**	11 / 30	**100**	14 / 13	15 / 36	5 / 4
1B	**100**	9 / 31	**100**	3 / 2	8 / 7	3 / 4
1C	**100**	3 / 0	**87 / 100**	0	1 / 0	0
2A	**100**	9 / 26	**100**	8 / 11	9 / 27	4 / 3
2B	**100**	3 / 11	**100**	1 / 0	2	2 / 0
2C	**100**	2 / 0	**100**	0	0	0
$c=0.5$	$\mathbf{X_1 \to X_2}$	$X_2 \to X_1$	$\mathbf{X_2 \to X_3}$	$X_3 \to X_2$	$X_1 \to X_3$	$X_3 \to X_1$
1A	**100**	8 / 32	**100**	11 / 14	32 / 95	8
1B	**100**	2 / 25	**100**	3 / 0	6 / 68	8 / 1
1C	**100**	0	**100**	0	2 / 32	0
2A	**100**	4 / 24	**100**	9 / 11	23 / 93	6 / 4
2B	**100**	1 / 9	**100**	1 / 0	4 / 57	2 / 1
2C	**100**	0	**100**	0	1 / 33	0

It turns out that when the resampled time series become more independent, the percentage of spurious couplings decreases. The most independent resampling schemes 1C and 2C give smallest rejection rate since the null distribution for the test is more spread and displaced to the right as the resampling changes from the least independent scheme (scheme A) to the most independent one (C) (see Fig. 1.1).

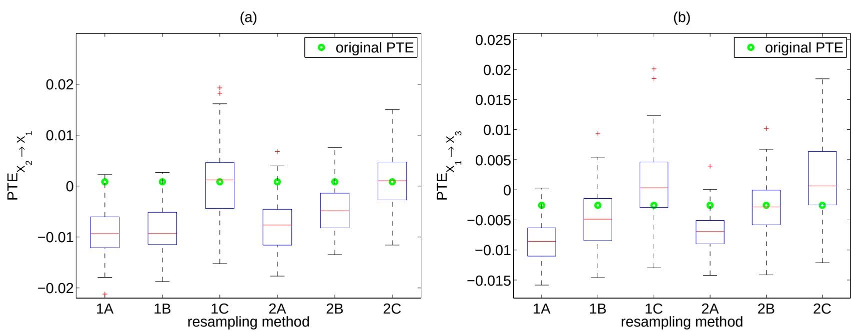

Fig. 1.1 Boxplots of surrogate / bootstrap PTE values and original PTE value from one realization of system 1 for $c = 0.3$ and $n = 2048$, for the directions (a) $X_2 \to X_1$ and (b) $X_1 \to X_3$.

System 2. The mean PTE values from 100 realizations of the second system for the directions of the true couplings are larger than for the other directions and increase with n, with the exception of $X_2 \rightarrow X_3$ that is at a lower level and does not increase with n (see Table 1.2). Concerning the uncoupled directions, the mean PTE values vary from 0.0013 to 0.0095 for both n and they decrease with n (the three largest mean PTE values across all non-direct couplings are reported in Table 1.2).

Table 1.2 Mean PTE values and percentage of significant PTE values from 100 realizations of system 2. The format of the table is as for Table 1.1.

mean PTE	$\mathbf{X_2 \rightarrow X_1}$	$\mathbf{X_1 \rightarrow X_3}$	$\mathbf{X_4 \rightarrow X_2}$	$\mathbf{X_2 \rightarrow X_3}$	$X_2 \rightarrow X_4$	$X_3 \rightarrow X_4$	$X_1 \rightarrow X_4$
$n = 512$	**0.0920**	**0.0772**	**0.0998**	**0.0060**	0.0095	0.0071	0.0067
$n = 2048$	**0.1221**	**0.0965**	**0.1355**	**0.0059**	0.0061	0.0042	0.0034
$n = 512$ / 2048	$\mathbf{X_2 \rightarrow X_1}$	$\mathbf{X_1 \rightarrow X_3}$	$\mathbf{X_4 \rightarrow X_2}$	$\mathbf{X_2 \rightarrow X_3}$	$X_2 \rightarrow X_4$	$X_3 \rightarrow X_4$	$X_1 \rightarrow X_4$
1A	**100**	**100**	**100**	**22 / 66**	3	1 / 0	1 / 0
1B	**100**	**99/ 100**	**100**	**0**	3 / 1	1 / 0	0
1C	**100**	**100**	**100**	**4 / 6**	0	0	0
2A	**100**	**100**	**100**	**14 / 60**	1 / 3	1 / 0	1 / 0
2B	**100**	**100**	**100**	**0**	3 / 4	2 / 0	1 / 0
2C	**100**	**100**	**100**	**5 / 14**	0	0	0

The true couplings $X_2 \rightarrow X_1$, $X_1 \rightarrow X_3$, $X_4 \rightarrow X_2$ are well established by the significance test (see Table 1.2), while no spurious causalities are identified (percentage of significant PTE vary from 0% to 6% at the uncoupled directions). The weak coupling $X_2 \rightarrow X_3$ is detected only by the scheme A, with a power of the test increasing with n. We note again that the surrogate / bootstrap PTE values increase as the resampled time series become more independent (see Figure 1.2).

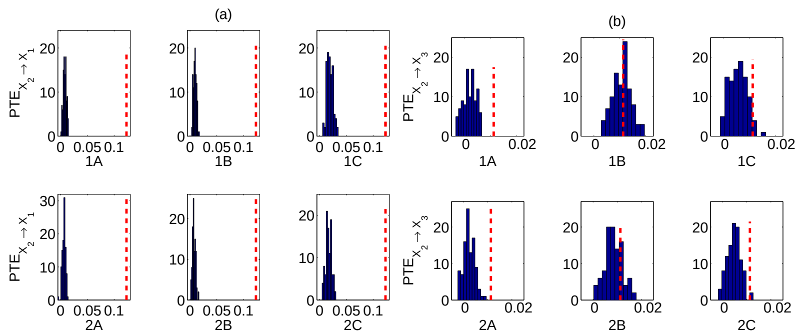

Fig. 1.2 Distribution of surrogate / bootstrap PTE values and original PTE value (vertical dotted line) from one realization of system 2 with $n = 2048$, for the directions a) $X_2 \rightarrow X_1$ and b) $X_2 \rightarrow X_3$.

System 3. No couplings are noted in the uncoupled case ($c = 0$) for system 3; the percentage of significant PTE values range from 0% to 11% for all the resampling schemes and both time series lengths. The PTE is also effective for $c = 0.2$ (see Table 1.3). As resampled time series become more independent, a loss in the power of the test is observed for $n = 512$. For the strong coupling strength $c = 0.4$, indirect and spurious couplings are observed for $n = 2048$ based on the resampling scheme A, e.g. we obtained for scheme 1A: 49% for $X_1 \rightarrow X_3$, 60% for $X_2 \rightarrow X_4$, 64% for $X_3 \rightarrow X_5$, 19% for $X_2 \rightarrow X_1$, 18% for $X_3 \rightarrow X_2$, 22% for $X_4 \rightarrow X_3$ and 27% for $X_5 \rightarrow X_4$. Similar results are observed for scheme 2A. Scheme B indicates the spurious coupling $X_2 \rightarrow X_4$, but at a lower percentage than scheme A. Only the true couplings are indicated using the resampling methods C (see Table 1.3).

Table 1.3 Percentage of significant PTE values from 100 realizations of system 3 for $n = 512$ / 2048, for the true couplings, an indirect coupling ($X_2 \rightarrow X_4$) and an uncoupled case ($X_5 \rightarrow X_4$).

$c = 0.2$	$\mathbf{X_1 \rightarrow X_2}$	$\mathbf{X_2 \rightarrow X_3}$	$\mathbf{X_3 \rightarrow X_4}$	$\mathbf{X_4 \rightarrow X_5}$	$X_2 \rightarrow X_4$	$X_5 \rightarrow X_4$
1A	**54 / 100**	**58 / 100**	**63 / 100**	**54 / 100**	8 / 5	10 / 7
1B	**54 / 100**	**57 / 100**	**59 / 100**	**52 / 100**	6 / 2	9 / 5
1C	**31 / 100**	**22 / 98**	**18 / 99**	**16 / 99**	1 / 0	0
2A	**53 / 100**	**62 / 100**	**67 / 100**	**59 / 100**	8 / 7	14 / 9
2B	**48 / 100**	**60 / 100**	**64 / 100**	**57 / 100**	7 / 0	11 / 1
2C	**29 / 100**	**29 / 100**	**28 / 100**	**26 / 99**	1 / 0	2 / 0
$c = 0.4$	$\mathbf{X_1 \rightarrow X_2}$	$\mathbf{X_2 \rightarrow X_3}$	$\mathbf{X_3 \rightarrow X_4}$	$\mathbf{X_4 \rightarrow X_5}$	$X_2 \rightarrow X_4$	$X_5 \rightarrow X_4$
1A	**100**	**100**	**98 / 100**	**100**	15 / 60	18 / 27
1B	**100**	**100**	**99 / 100**	**99 / 100**	6 / 21	3 / 6
1C	**100**	**83 / 100**	**86/ 100**	**84 / 100**	1	1 / 0
2A	**100**	**100**	**100**	**100**	21 / 65	22
2B	**100**	**100**	**100**	**100**	9 / 25	10 / 8
2C	**96 / 100**	**96 / 100**	**96 / 100**	**96 / 100**	1	2 / 1

1.5 Discussion

The importance of assessing the correct statistical significance for the partial transfer entropy (PTE) has been explored in a simulation study. Concerning the resampled time series, by definition, the mutual information of X and Y conditioned on Z should be in theory zero, i.e. $I(Y;X|Z) = 0$. The formulation of more independent resampled data (schemes B and C) compared to the standard technique (scheme A) seems to improve the bias of the test statistic and helps prevent false indications of couplings in the case of the nonlinear coupled systems. The size and the power of the test are improved, especially if strong couplings exist. However, when the couplings are linear, scheme A seems to be more efficient in identifying weak couplings.

It turns out that when the PTE is estimated for an increasing level of randomness in the resampled time series, the estimated PTE values also increase, while the distribution of PTE from the resampled time series gets wider and less spurious couplings are thus detected. This higher specificity comes at the cost of lower sensitivity, and vice versa. Thus, none of the six resampling schemes turns out to be optimal, but it becomes clear that the significance test for the PTE gets more conservative as resampling is more random.

The aforementioned resampling schemes can be utilized for any test statistic in order to examine the null hypothesis of no causal effects. Since the efficiency of a causality measure is determined in terms of the corresponding resampling technique that is used, the usefulness of each of the examined resampling schemes will be further investigated for different causality measures.

Acknowledgements The research project is implemented within the framework of the Action ”Supporting Postdoctoral Researchers” of the Operational Program ”Education and Lifelong Learning” (Action’s Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

References

1. B. Efron. Bootstrap methods: another look at the jackknife. *Ann. Stat.*, 7(1):1–26, 1979.
2. L. Faes, A. Porta, and G. Nollo. Mutual nonlinear prediction as a tool to evaluate coupling strength and directionality in bivariate time series: comparison among different strategies based on k nearest neighbors. *Phys. Rev. E*, 78:026201, 2008.
3. L. Faes, A. Porta, and G. Nollo. Testing frequency-domain causality in multivariate time series. *IEEE Trans. Biomed. Eng.*, 57(8):1897–1906, 2010.
4. M. Hénon. A two-dimensional mapping with a strange attractor. *Commun. Math. Phys.*, 50(1):69–77, 1976.
5. A. Kraskov, H. Stögbauer, and P. Grassberger. Estimating mutual information. *Phys. Rev. E*, 69(6):066138, 2004.
6. H.R. Künsch. The jackknife and the bootstrap for general stationary observations. *Ann. Stat.*, 17:1217–1241, 1989.
7. S.N. Lahiri. *Resampling Methods for Dependent Data*. Springer, 2003.
8. A. Papana, D. Kugiumtzis, and P.G. Larsson. Reducing the bias of causality measures. *Physical Review E*, 83:036207, 2011.
9. A. Papana, D. Kugiumtzis, and P.G. Larsson. Detection of direct causal effects and application to electroencephalogram analysis. *Int. J. Bifurcat. Chaos*, 22(9):1250222, 2012.
10. A. Papana, C. Kyrtsou, D. Kugiumtzis, and C. Diks. Simulation study of direct causality measures in multivariate time series. *Ent.*, 15(7):2635–2661, 2013.
11. D. Politis and J.P. Romano. The stationary bootstrap. *J. Am. Stat. Assoc.*, 89(428):1303–1313, 1994.
12. D. Politis, J.P. Romano, and T. Lai. Bootstrap confidence bands for spectra and cross-spectra. *IEEE Trans. Sig. Proc.*, 40:1206–1215, 1992.
13. 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*, 65:041903, 2002.
14. T. Schreiber. Measuring information transfer. *Phys. Rev. Lett.*, 85(2):461–464, 2000.
15. J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J.D. Farmer. Testing for nonlinearity in time series: the method of surrogate data. *Phys. D*, 58:77–94, 1992.
16. V.A. Vakorin, O.A. Krakovska, and A.R. McIntosh. Confounding effects of indirect connections on causality estimation. *Journal of Neuroscience Methods*, 184:152–160.
17. I. Vlachos and D. Kugiumtzis. Non-uniform state space reconstruction and coupling detection. *Phys. Rev. E*, 82:016207, 2010.
18. M. Winterhalder, B. Schelter, W. Hesse, K. Schwab, L. Leistritz, D. Klan, R. Bauer, J. Timmer, and H. Witte. Comparison of linear signal processing techniques to infer directed interactions in multivariate neural systems. *Signal Proc.*, 85(11):2137–2160, 2005.
19. G.H. Yu and C.C. Huang. A distribution free plotting position. *Stoch. Env. Res. Risk Assess.*, 15(6):462–476, 2001.