35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Decomposing the Transfer Entropy to Quantify Lag-Specific Granger Causality in Cardiovascular Variability*

Luca Faes, *Member, IEEE,* and Giandomenico Nollo, *Member, IEEE*

***Abstract*— We present a modification of the well known transfer entropy (TE) which makes it able to detect, besides the direction and strength of the information transfer between coupled processes, its exact timing. The approach follows a decomposition strategy which identifies –according to a lag-specific formulation of the concept of Granger causality– the set of time delays carrying significant information, and then assigns to each of these delays an amount of information transfer such that the total contribution yields the overall TE. We propose also a procedure for the practical estimation from time series data of the relevant delays and lag-specific TE in both bivariate and multivariate settings. The proposed approach is tested in simulations and in real cardiovascular time series, showing the feasibility of lag-specific TE estimation, the ability to reflect expected mechanisms of cardiovascular regulation, and the necessity of using the multivariate TE to properly assess time-lagged information transfer in the presence of multiple interacting systems.**

I. Introduction

The transfer entropy (TE) is a well known measure of directional information transfer between coupled dynamical systems [1]. The popularity of TE for detecting directional dependencies from time series data stems from its solid foundation in information theory, its sensitivity to both linear and nonlinear interactions, and its close connection with the ubiquitous concept of Granger causality [2]. Thanks to these features, the TE is widely used to assess the transfer of information in physiological systems, though mostly following a bivariate approach whereby only the time series from the two investigated systems are considered [3,4]. However, bivariate TE analysis may lead to wrong inference of Granger causality when the two analyzed systems are potentially connected to other interacting systems [2,5]. Only recently, the introduction of data-efficient procedures tackling the problem of high-dimensional entropy estimation has made it possible to pursue fully multivariate approaches to the computation of TE [6,7].

A limitation of the TE is that it is not lag-specific, i.e., it quantifies the flow of information between systems without detecting the timing through which information flows. Evaluating the timing of information transfer may be of great importance to understand the function of complex networks such as those sub-serving physiological interactions. The present study introduces an approach to modify the TE in order to make it able to detect the delay of Granger causal interactions between coupled processes. The approach is based on decomposing the TE into a sum of terms that quantify the information transfer at the specific time lags for which Granger causality exists between the two analyzed processes. We also present a procedure for the empirical estimation of lag-specific Granger causality and TE in both bivariate and multivariate settings. The overall approach is first validated on simulations of multiple stochastic processes, and then applied to cardiovascular and cardiorespiratory time series measured from healthy humans during a head-up tilt testing protocol.

*Research supported in part by IRCS PAT funding.

L. F. is with the BIOtech, Department of Physics, University of Trento, 38123 Mattarello (TN), Italy (corresponding author phone: +39-0461-282773; fax: +39-0461-283091; e-mail: luca.faes@unitn.it).

G. N. is with the BIOtech, Department of Industrial Engineering, University of Trento, and IRCS PAT-FBK, Trento, 38123 Mattarello (TN), Italy (e-mail: nollo@science.unitn.it).

II. Methods

A. Lag-Specific Granger Causality and Transfer Entropy

Let us consider two dynamical systems X and Y, visiting states described by the discrete time stationary processes x and y, and denote as x_n, y_n the corresponding random variables at time n, and as $X_n^-=(x_{n-1},x_{n-2},...)$, $Y_n^-=(y_{n-1},y_{n-2},...)$ the sets of all variables describing the past of the processes. Then, according to the original definition of Granger causality [8], G-causality from x to y, $X_n^- \rightarrow y_n$, exists if X_n^- contains information that helps predicting y_n above and beyond the information contained in Y_n^-. A more general formulation of this notion is based on conditional probabilities, and can be formulated in the information domain in terms of the well known transfer entropy (TE) [1] stating that $X_n^- \rightarrow y_n$ if and only if $TE_{X\rightarrow Y}>0$ [2], where

$$TE_{X\rightarrow Y}=I(y_n, X_n^- | Y_n^-). \quad (1)$$

The TE, which is formulated in (1) as a conditional mutual information (CMI), can be interpreted as the reduction of uncertainty about y_n when learning the past of x, if the past of y is already known. The TE measures aggregate Granger causal influence of x at past lags, and thus is not lag-specific.

In order to characterize Granger causal influences between processes for specific time lags, the definition of G-causality can be intuitively itemized as follows: G-causality exists from x to y at lag u, $x_{n-u} \rightarrow y_n$, if x_{n-u} contains information that helps predicting y_n above and beyond the information contained in Y_n^- and in $X_n^- \backslash x_{n-u}$ (where \ denotes subtraction from a set). In the information domain, this definition can be formulated stating that $x_{n-u} \rightarrow y_n$ if and only if $I_{X\rightarrow Y}(u)>0$, where the CMI $I_{X\rightarrow Y}(u)$ is defined as

$$I_{X\rightarrow Y}(u)=I(y_n, x_{n-u} | Y_n^-, X_n^- \backslash x_{n-u}). \quad (2)$$

The CMI in (2) can be interpreted as the reduction of uncertainty about y_n when learning the past of x at lag u, if the past of x at all other lags and the whole past of y are already known. While the condition $I_{X\rightarrow Y}(u)>0$ allows

assessing lag-specific G-causality, the modulus of $I_{X\to Y}(u)$ cannot be directly related to the information transfer measured by the TE. To get a lag-specific measure of information transfer we propose to decompose the TE as

$$TE_{X\to Y} = \sum_{u_k\in\Lambda} I\left(y_n, x_{n-u_k}\middle|Y_n^-, x_{n-u_{k+1}},\ldots,x_{n-u_L}\right) = \sum_{k=1}^{L} TE_{X\to Y}(u_k) \quad (3)$$

where the sum is extended to the set of lags $\Lambda=(u_1,\ldots u_L)$ for which lag specific G-causality exists (i.e., $x_{n-u}\to y_n$ if and only if $u\in\Lambda$). The decomposition in (3) puts in evidence lag-specific information transfers, defined in a way such that their aggregate contribution yields the overall TE. In Sect. II.*B* we illustrate a procedure for estimating the set Λ of the lags bearing G-causality according to (2), and then the overall $TE_{X\to Y}$ as well as the individual contributions $TE_{X\to Y}(u)$ according to (3).

While the above description refers to bivariate systems, it is well known that the presence of other systems potentially connected with X and Y might change the interpretation of G-causality and TE computed as in (1) [2,5,8]. Thus, when measurements from these other systems, collected in the vector $\boldsymbol{Z}$, are available in the form of the process $\boldsymbol{z}$, a multivariate form of the TE, $TE_{X\to Y|\boldsymbol{Z}}=I(y_n, X_n^-|Y_n^-, \boldsymbol{Z}_n^-)$, needs to be used to rule out the information shared between X and Y that could be triggered by their common interaction with $\boldsymbol{Z}$. Accordingly, in the multivariate case the presence of G-causality is assessed checking for nonzero values of the CMI $I_{X\to Y|\boldsymbol{Z}}(u)=I(y_n, x_{n-u}|Y_n^-, \boldsymbol{Z}_n^-, X_n^-\backslash x_{n-u})$, and the decomposition in (3) is modified considering the quantities $TE_{X\to Y|\boldsymbol{Z}}(u_k)= I(y_n, x_{n-u_k}|Y_n^-, \boldsymbol{Z}_n^-, x_{n-u_{k+1}},\ldots, x_{n-u_L})$ as lag-specific information transfers composing the multivariate index $TE_{X\to Y|\boldsymbol{Z}}$.

B. Estimation Approach

In this Section we propose a unified approach for estimating from time series data all measures defined above (i.e., both lag-specific and aggregate TE, in either the bivariate or multivariate formulation). The approach is based on recognizing that any CMI measure can be expressed as the difference between two conditional entropy (CE) terms; e.g., the multivariate aggregate TE can be expressed as $TE_{X\to Y|\boldsymbol{Z}}=I(y_n, X_n^-|Y_n^-, \boldsymbol{Z}_n^-)=H(y_n|Y_n^-, \boldsymbol{Z}_n^-)-H(y_n|X_n^-,Y_n^-, \boldsymbol{Z}_n^-)$, where $H(y|\boldsymbol{v})$ is the CE measuring the entropy of the scalar variable y conditioned to the vector variable $\boldsymbol{v}$. Then, we compute the CE according to a non-uniform conditioning scheme [7] that follows a sequential procedure whereby the conditioning vector $\boldsymbol{v}$ is updated progressively, taking all relevant processes into consideration at each step and selecting the components that better reduce the uncertainty about the target variable y.

Specifically, a set of initial candidate components is first defined including the past of all relevant processes up to a maximum lag L_{max}, i.e., $\Omega^b=(x_{n-1},\ldots, x_{n-L_{max}},y_{n-1},\ldots,y_{n-L_{max}})$ for the bivariate analysis and $\Omega=(\Omega^b, \boldsymbol{z}_{n-1},\ldots, \boldsymbol{z}_{n-L_{max}})$ for the multivariate analysis. Then, we compute the entropy of the target variable y_n conditioned to Ω starting from an empty conditioning vector, $\boldsymbol{v}_0=(\cdot)$, and proceeding as follows: at each step $k\geq 1$, form the candidate vector $(w,\boldsymbol{v}_{k-1})$, where $w\in\Omega$, $w\notin\boldsymbol{v}_{k-1}$, and compute the CE $H(y_n|w,\boldsymbol{v}_{k-1})$; repeat the computation for all possible candidates, and then retain the candidate for which the estimated CE is minimum, i.e., set $\boldsymbol{v}_k=(\hat{w},\boldsymbol{v}_{k-1})$ such that $\hat{w}=\arg\min_w H(y_n|w,\boldsymbol{v}_{k-1})$; terminate the procedure when an irrelevant component has been selected, i.e., when the CMI $I(y_n,\hat{w}|\boldsymbol{v}_{k-1})=H(y_n|\boldsymbol{v}_{k-1})-H(y_n|\boldsymbol{v}_k)$ is not statistically significant. The significance of the reduction in the CE brought by the candidate $\hat{w}$ selected at step k was assessed empirically using surrogate data. Specifically, the CMI $I(y_n,\hat{w}|\boldsymbol{v}_{k-1})$ was compared with a threshold taken as the $100(1-\alpha)^{\text{th}}$ percentile of the distribution of the CMI computed after shifting the realizations of $\hat{w}$ of a randomly selected lag with respect to y_n and $\boldsymbol{v}_{k-1}$; if the original CMI was above the threshold, the component $\hat{w}$ was included in the conditioning vector, otherwise it was discarded and the procedure terminated including $k-1$ components in the final vector $\boldsymbol{v}=\boldsymbol{v}_{k-1}$.

After termination of the sequential estimation procedure, the conditioning vector is composed as $\boldsymbol{v}=(\boldsymbol{v}^x, \boldsymbol{v}^y, \boldsymbol{v}^z)$, where $\boldsymbol{v}^x$, $\boldsymbol{v}^y$, and $\boldsymbol{v}^z$ denote the components of $\boldsymbol{v}$ belonging respectively to X, Y, and $\boldsymbol{Z}$. Then, computing the CMI between the lagged components of X and the target variable y_n as $I_{X\to Y|\boldsymbol{Z}}(u)=H(y_n|\boldsymbol{v}\backslash x_{n-u})-H(y_n|\boldsymbol{v})$, we have that G-causality at lag u from X to Y is detected only when x_{n-u} is selected by the conditioning procedure, because $I_{X\to Y|\boldsymbol{Z}}(u)>0$ if $x_{n-u}\in\boldsymbol{v}^x$ and $I_{X\to Y|\boldsymbol{Z}}(u)=0$ if $x_{n-u}\notin\boldsymbol{v}^x$. Therefore, our estimate of the set of lags bearing G-causality from X to Y is $\Lambda=\{u_k|x_{n-u_k}\in\boldsymbol{v}^x\}$. As a consequence, the lag-specific information transfer is estimated as $TE_{X\to Y|\boldsymbol{Z}}(u)=0$ if $u\notin\Lambda$ (i.e., if $x_{n-u}\notin\boldsymbol{v}^x$), and as $TE_{X\to Y|\boldsymbol{Z}}(u)=H(y_n|\boldsymbol{v}^y,\boldsymbol{v}^z,x_{n-u_{k+1}},\ldots,x_{n-u_L})-H(y_n|\boldsymbol{v}^y,\boldsymbol{v}^z,x_{n-u_k},\ldots,x_{n-u_L})$ if $u=u_k\in\Lambda$. The aggregate TE results either summing up all lag-specific terms or computing $TE_{X\to Y|\boldsymbol{Z}}=H(y_n|\boldsymbol{v}^y,\boldsymbol{v}^z)-H(y_n|\boldsymbol{v})$. Note that estimation of all measures for pure bivariate systems results simply as a particular case in which the components of $\boldsymbol{Z}$ do not appear in the conditioning procedure (i.e., $\Omega=\Omega^b$, $\boldsymbol{v}^z=(\cdot)$). In this study, practical estimation of the CE from time series data was performed using the classical histogram-based method, that consists in coarse-graining the observed dynamics using Q quantization levels, and computing entropies by approximating probability distributions with the frequencies of occurrence of the quantized values [9].

III. Verification on Simulated Processes

To test our approach on simulations reproducing oscillations and interactions typical of short-term vascular, cardiac and respiratory variability, we considered three systems X, Y and Z described by the stochastic processes

$$\begin{aligned} x_n &= 2\rho_x\cos(2\pi f_x)x_{n-1} - \rho_x^2 x_{n-2} + 0.4z_{n-d_1} + u_n \\ y_n &= 0.2x_{n-d_2} + c\,z_{n-d_3} + v_n \\ z_n &= 2\rho_z\cos(2\pi f_z)z_{n-1} - \rho_z^2 z_{n-2} + w_n \end{aligned} \quad (4)$$

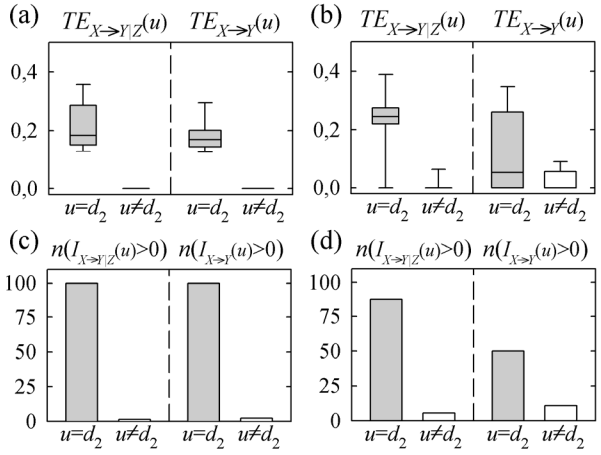

Figure 1. Distribution over 100 realizations of (4) of the estimated bivariate and multivariate lag-specific TE ($TE_{X\to Y}(u)$, $TE_{X\to Y|Z}(u)$), and corresponding number of realizations for which lag-specific G-causality was detected ($n(I_{X\to Y}(u)>0)$, $n(I_{X\to Y|Z}(u)>0)$), computed for the coupled lags ($u=d_2$, gray) and for the uncoupled lags ($u\neq d_2$, white) in the absence (c=0; panels a,c) and in the presence (c=0.4; panels b,d) of coupling from Z to Y.

where u_n, v_n, and w_n are independent white noise processes with zero mean and unit variance, and x and z are described as second order autoregressive processes oscillating at the frequencies f_x=0.1 and f_z=0.3 (ρ_x=0.9, ρ_z=0.95). The imposed lag-specific G-causality relations are $z_{n-d_1}\to x_n$, $x_{n-d_2}\to y_n$, and $z_{n-d_3}\to y_n$ (modulated by the coupling parameter c). We considered two coupling situations, the first with z affecting x but not y (c=0), and the second with z affecting both x and y (c=0.4). In each case, we generated 100 realizations of (4), each lasting 300 samples; at each realization, the coupling delays d_1, d_2, and d_3 were randomly chosen between 1 and 5.

The analysis was performed separately following the bivariate and the multivariate approach, and focusing on G-causality from X to Y. The conditioning procedure was run including L_{max}=5 components from each series in the set of initial candidates, setting α=0.05 as statistical significance level for candidate selection, and using Q=6 quantization levels for histogram-based entropy estimation. For each process realization, we detected a true positive (TP) information transfer when x_{n-d_2} was selected in the conditioning procedure, and a false positive (FP) when the component x_{n-u}, with $u\neq d_2$, was selected. False negative (FN) and true positive (TP) detections occurred when x_{n-d_2} was not selected and when x_{n-u} was not selected at $u\neq d_2$, respectively. Results are presented in Fig. 1, showing the distributions over all realizations of the estimated bivariate and multivariate TE obtained separately for $u=d_2$ and for $u\neq d_2$, together with their corresponding relative frequency of detected G-causality for $u=d_2$ (i.e., the percentage of TP) and for $u\neq d_2$ (i.e., the percentage of FP). We see that, in the absence of common driving of Z on both X and Y (parameter c=0), bivariate and multivariate TE perform equally well in picking up the correct interaction delays, as documented by the large TE values for $u=d_2$ and null TE values for $u\neq d_2$ (Fig. 2a), and by the high detection rate combined with a low FP rate (Fig. 2c, indicating 100% sensitivity and 98% specificity in both cases). On the contrary, when Z affected

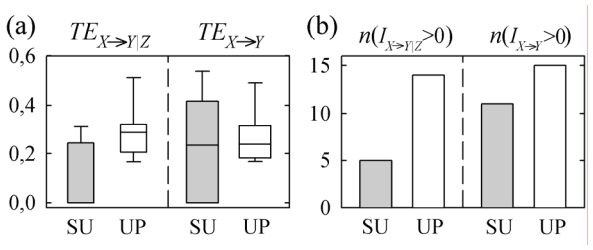

Figure 2. Distribution over 15 subjects of the estimated bivariate and multivariate TE from SAP to HP ($TE_{X\to Y}$, $TE_{X\to Y|Z}$; a), and corresponding number of realizations for which G-causality was detected ($n(I_{X\to Y}>0)$, $n(I_{X\to Y|Z}>0)$; b), computed in the supine (SU) and upright (UP) position.

both X and Y (c=0.4) only the multivariate TE performed well, showing $TE_{X\to Y|Z}(d_2)$ well separated from $TE_{X\to Y|Z}(u)$ for $u\neq d_2$ (Fig. 2b) with high TP rate (sensitivity=88%) and low FP rate (specificity=95%) (Fig. 2d, left), while the bivariate TE exhibited partially overlapping distributions of $TE_{X\to Y}$ over coupled and uncoupled lags (Fig. 2b) and degraded accuracy in the identification of the interaction delays (Fig 2d, right, indicating sensitivity=50%, specificity=89%).

IV. Application to Cardiovascular Variability

The proposed approach was applied to cardiovascular and cardiorespiratory short-term interactions, considering the vascular, cardiac and respiratory systems respectively as systems X, Y and Z, and the systolic arterial pressure (SAP), heart period (HP) and respiratory flow (RF) as corresponding processes x, y, and z. Stationary realizations of the three processes, each lasting 300 samples, were obtained from 15 young healthy subjects in the resting supine position (SU), and in the 60° upright position (UP) after passive head-up tilt [10]. The analysis was performed normalizing each time series to zero mean and unit variance, and then computing bivariate and multivariate TE from X to Y (analysis parameters: L_{max}=10, α=0.05, Q=6). Since, according to the convention adopted to measure the time series, zero-lag effects of RF and SAP on HP are physiologically meaningful [11], the non-delayed components x_n and z_n were included in the candidate sets Ω and Ω^b considered for describing the target variable y_n [5].

Results of the analysis of overall information transfer from SAP to HP are depicted in Fig. 2. We found that in the supine position the bivariate approach assessed higher transfer than the multivariate approach, in terms of both TE values (Fig. 2a) and number of subjects with significant detected G-causality (Fig. 2b), while in the upright position the two approaches performed comparably. The trends observed for the multivariate TE confirm previous results suggesting an intensified transfer of information from SAP to HP with the supine-to-upright transition, probably reflecting an enhancement of the baroreflex control of heart rate [10]. On the contrary, the bivariate TE omitting to consider the effects of RF in SAP-HP causality analysis seems unable to highlight this expected activation of the baroreflex.

The lag-specific analysis reported in Fig. 3 confirmed the overall TE analysis as regards the comparison between bivariate and multivariate approaches. Indeed, the observed

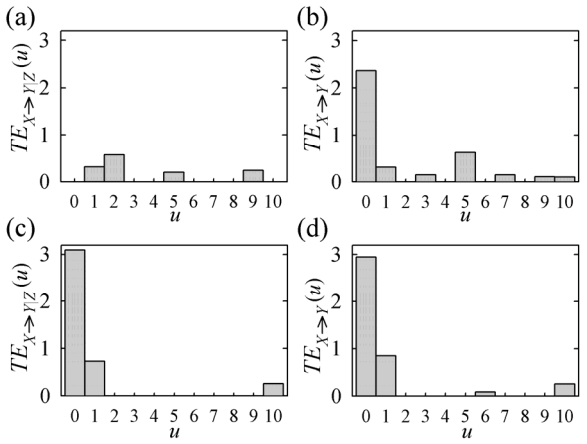

Figure 3. Cumulative lag-specific TE from SAP to HP obtained summing over 15 subjects the values of the multivariate TE ($TE_{X\to Y|Z}(u)$,a,c) and of the bivariate TE ($TE_{X\to Y}(u)$, b,d) estimated in the supine position (a,b) and in the upright position (c,d).

patterns of lag-specific causality were different between bivariate and multivariate TE in the supine position (Fig. 3a,b), and very similar in the upright position (Fig. 3c,d). Our interpretation for this behavior is that the RF is known to affect strongly both SAP and HP in the resting condition [12], thus likely acting as a confounder in a bivariate SAP-HP analysis, while it is less effective in driving HP during the orthostatic stress induced by tilt [13], thus exerting a reduced interference over the coupling between SAP and HP. Another interesting result of the entropy decomposition is the clear emergence of G-causality at small delays from SAP to HP with the tilting transition, documented by the big amount of multivariate TE measured at lag 0 in the upright position (Fig. 3c), compared with the absence of information transfer at the same lag in the supine position (Fig. 3a). This result suggests that the tilt-induced activation of the baroreflex is mainly manifested by means of fast, within-beat effects of SAP on HP variability, and is consistent with physiological evaluations reporting values lower than one second for the latency of this reflex [14].

V. Conclusions

The most distinctive features of the approach proposed in this study to assess the strength and the delay of causal interactions between coupled processes are that it is designed: (i) to work in the model-free framework provided by information theory; (ii) to estimate nonzero information transfer only at the time lags for which significant causality exists according to a lag-specific formulation of the Granger notion [8]; and (iii) to quantify the information transfer in a way such that the sum of all contributions at the different time lags yields exactly the well-defined TE measure [1]. The reported simulation examples demonstrated the ability of the approach to recover the true interaction delays, and consequently to properly allocate the overall information transfer across time, in short process realizations mimicking the conditions typical of physiological time series analysis. The subsequent application to cardiovascular and cardiorespiratory series suggested the capability of the approach to detect both strength and timing of the cardiovascular information transfer in agreement with the expected behavior of important physiological control mechanisms such as the baroreflex [12-14]. Moreover, the comparison between bivariate and multivariate analyses confirmed that the proper assessment of Granger causal measures, like the TE and our lag-specific version, can be properly accomplished in cardiovascular variability analysis only compensating for the confounding effects of respiration on the joint variability of heart rate and arterial pressure [15].

References

[1] T. Schreiber, "Measuring information transfer," *Phys. Rev. Lett.*, vol. 85, pp. 461-464, 2000.

[2] P. O. Amblard and O. J. Michel, "The relation between Granger causality and directed information theory: a review," *Entropy*, vol. 15, no. 1, pp. 113-143, 2013.

[3] R. Vicente, M. Wibral, M. Lindner, and G. Pipa, "Transfer entropy-a model-free measure of effective connectivity for the neurosciences," *J. Comput. Neurosci.*, vol. 30, no. 1, pp. 45-67, Feb. 2011.

[4] L. Faes, S. Erla, and G. Nollo "Compensating for instantaneous signal mixing in transfer entropy analysis of neurobiological time series," in *Proc. 34th Annu. Int. Conf. IEEE-EMBS*, San Diego, 2012, pp. 3672–3675.

[5] L. Faes, G. Nollo, and A. Porta, "Compensated transfer entropy as a tool for reliably estimating information transfer in physiological time series," *Entropy*, vol. 15, no. 1, pp. 198-219, 2013.

[6] J. Runge, J. Heitzig, V. Petoukhov, and J. Kurths, "Escaping the Curse of Dimensionality in Estimating Multivariate Transfer Entropy," *Phys. Rev. Lett.*, vol. 108, pp. 258701, 2012.

[7] L. Faes, G. Nollo, and A. Porta, "Information-based detection of nonlinear Granger causality in multivariate processes via a nonuniform embedding technique," *Phys. Rev. E*, vol. 83, no. 5-1, pp. 051112, May 2011.

[8] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37 pp. 424-438, 1969.

[9] L. Paninski, "Estimation of entropy and mutual information," *Neur. Comput.*, vol. 15 pp. 1191-1254, 2003.

[10] L. Faes, G. Nollo, and A. Porta, "Non-uniform multivariate embedding to assess the information transfer in cardiovascular and cardiorespiratory variability series," *Comput. Biol. Med.*, vol. 42, pp. 290-297, 2012.

[11] L. Faes and G. Nollo, "Extended causal modeling to assess partial directed coherence in multiple time series with significant instantaneous interactions," *Biol. Cybern.*, vol. 103, pp. 387-400, 2010.

[12] M. A. Cohen and J. A. Taylor, "Short-term cardiovascular oscillations in man: measuring and modelling the physiologies," *J. Physiol.*, vol. 542, no. Pt 3, pp. 669-683, Aug.2002.

[13] L. Faes, G. Nollo, and A. Porta, "Information domain approach to the investigation of cardio-vascular, cardio-pulmonary, and vasculo-pulmonary causal couplings," *Front. Physiol.*, vol. 2, no. 90, pp. 1-13, 2011.

[14] T. G. Pickering and J. Davies, "Estimation of the conduction time of the baroreceptor-cardiac reflex in man," *Cardiovasc. Res.*, vol. 7, no. 2, pp. 213-219, Mar. 1973.

[15] A. Porta, T. Bassani, V. Bari, G. D. Pinna, R. Maestri, and S. Guzzetti, "Accounting for Respiration is Necessary to Reliably Infer Granger Causality From Cardiovascular Variability Series," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 3, pp. 832-841, Mar. 2012.