An Approach to Decompose the Information Shared between Cardiovascular and Respiratory Physiological Systems

Chiara Barà[1], Yuri Antonacci[1], Michal Javorka[2], Luca Faes[1]

[1] Department of Engineering, University of Palermo, Palermo, Italy
[2] Department of Physiology, Comenius University in Bratislava, Jessenius Faculty of Medicine in Martin, Martin, Slovak Republic

Abstract

The framework of Partial Information Decomposition (PID) quantifies the redundant and synergistic information that two source variables share with a third variable within a network of interacting systems. While the PID has been formulated so far either for discrete or continuous Gaussian variables, this work introduces a computation strategy, based on nearest-neighbor entropy estimation, to decompose the information that a discrete variable shares with two continuous sources. The approach is first validated in a Boolean logic gate producing exclusive synergy, and then applied in the context of cardiovascular physiology to decompose the information shared by the binary variable mapping the respiratory phase and the continuous time series of heart period and systolic arterial pressure variability. We show how cardiac and vascular regulatory mechanisms are individually related to respiration at rest and during postural stress, and jointly interact in maintaining cardiovascular and cardiorespiratory homeostasis during the inspiratory and expiratory breathing phases.

1. Introduction

In recent years, the field of Network Physiology has emerged as a result of the growing interest in studying the complex networks of physiological systems that interact to ensure several homeostatic functions in the human organism [1]. In fact, physiological regulatory mechanisms do not operate exclusively through the response of individual organ systems to a stimulus, but also encompass complex functional interactions among multiple systems. In this context, the concepts of redundancy and synergy have emerged as fundamental tools to describe how multiple physiological "source" variables share information with a "target" variable, highlighting modes of interaction that cannot be revealed looking at the pairwise interactions between each source and the target [2].

Among the approaches introduced to measure these high-order interactions, the Partial Information Decomposition (PID) is the most promising since it provides non-negative decomposition terms with specific meaning related to unique, redundant and synergistic information [3]. However, the practical computation of the PID terms is typically restricted to discrete random variables [4] or to continuous variables with Gaussian distribution [5], thus limiting its applicability to categorical or quantized data, or to linear systems. To overcome this limitation, this work introduces a mixed approach considering a discrete target and two continuous source variables. This model-free approach is implemented using nearest-neighbor entropy estimators, and tested over simulated data with known ground truth. Then, it is applied to cardiovascular time series measuring the variability of heart period and arterial pressure to study their unique, redundant and synergistic contributions to the phase of respiration in healthy subjects studied at rest and during postural stress.

2. Non-parametric PID formulation

In the PID framework, the mutual information (MI) between two source variables X_1 and X_2 and a target variable Y is expressed as follows [3]:

$$\begin{aligned} I(Y;X_1,X_2) = U(Y;X_1) + U(Y;X_2) \\ + R(Y;X_1,X_2) + S(Y;X_1,X_2), \end{aligned} \tag{1}$$

where $U(Y;X_1)$ and $U(Y;X_2)$ denote the information exclusively shared with Y by each of the sources (*unique information*), $R(Y;X_1,X_2)$ the information shared both by X_1 and X_2 (*redundancy*), and $S(Y;X_1,X_2)$ the information emerging only when X_1 and X_2 are observed jointly (*synergy*). Starting from the so-called redundancy lattice, an iterative approach is used to compute the decomposition terms starting from the MI between Y and the sources and from the redundancy, i.e.,

$$\begin{aligned} &U(Y;X_1) = I(Y;X_1) - R(Y;X_1,X_2), \\ &U(Y;X_2) = I(Y;X_2) - R(Y;X_1,X_2), \\ &S(Y;X_1,X_2) = I(Y;X_1,X_2) - I(Y;X_1) \\ &- I(Y;X_2) + R(Y;X_1,X_2). \end{aligned} \tag{2}$$

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.351

Considering two continuous sources with outcomes $x_j \in D_{X_j} \subseteq \mathbb{R}$ and a discrete variable Y with possible outcomes $i \in A_Y = \{1,\ldots,Q\}$, each with probability $p(i)$, the redundancy term is formulated in [3] as the expected value of the minimum information that the sources share with each specific outcome of Y, i.e.,

$$R(Y;X_1,X_2) = \sum_{i\in A_Y} p(i) \min_{X_j, j=1,2} I(Y=i;X_j). \quad (3)$$

The probability distributions of each source when considering, for the variable Y, respectively, all possible outcomes, i.e., $p(x_j)$, and a specific state i, i.e., $p(x_j|i)$, are used to define the *specific information* shared between the outcome $\{Y=i\}$ and the source X_j, as:

$$I(Y=i;X_j) = \int_{x_j \in D_{X_j}} p(x_j|i) \log \frac{p(x_j|i)}{p(x_j)} dx_j. \quad (4)$$

In this work, we propose to estimate the probability mass of Y in (3) with the frequentistic approach, and the specific MI in (4) with the Kraskov-Stögbauer-Grassberger (KSG) estimation strategy [6]. Specifically, defining the searching distances d to the kth neighbor of the sample $s=\{x_1,x_2\}$ associated with the sample $\{Y=i\}$, i.e., $\{s\}_i$, and searching for samples in all the exploited spaces using the same range d, the following estimator is derived:

$$\begin{aligned}\hat{I}(Y=i;X_j) &= \psi(N) - \psi(N_i) \\ &+ 1/N_i \sum_{\{x_j\}_i} \Big(\psi(m_j^i+1) - \psi(m_j+1)\Big),\end{aligned} \quad (5)$$

where $\psi(\cdot)$ is the digamma function, N is the total number of samples, N_i the number of samples associated to $\{Y=i\}$, and m_j and m_j^i the points counted at distances smaller than $d/2$ in the space X_j when considering samples for any Y, and for the specific values $\{Y=i\}$, respectively. After estimating the redundant term according to (3), and the shared information between Y and the sources both separately and jointly, as:

$$\hat{I}(Y;X_j) = \sum_{i\in A_Y} \hat{p}(i)\hat{I}(Y=i;X_j), \quad (6)$$

$$\begin{aligned}\hat{I}(Y;X_1,X_2) = \sum_{i\in A_Y} \hat{p}(i)\Big(&\psi(N) - \psi(N_i) \\ &+ \psi(k) - 1/N_i \sum_{\{s\}_i} \psi(m_{12}+1)\Big),\end{aligned} \quad (7)$$

where m_{12} quantifies the neighbors within distance $d/2$ in the space $[X_1X_2]$ independently of Y, the unique and the synergistic contributions are finally computed using (2).

Surrogate data analysis was performed to assess the statistical significance of each source contribution to a specific state of the variable Y as well as of each term of the decomposition. Specifically, by randomly shuffling the variable Y, the correspondence of each sources sample with the attributed Y was destroyed. Iterating this shuffling led to obtain a distribution of all decomposition terms on surrogate data, which can then be compared with the value computed on the original data. A non-parametric percentile test was used to determine the threshold significance as the 95*th* percentile of the surrogate distribution, above which the measure was deemed as significant.

3. Simulation study

The proposed PID estimator was tested on the Boolean logic XOR-gate which is known to be characterized by a synergistic behaviour [4]. Considering two independent uniformly distributed continuous variables, i.e., $X_1 \in \mathcal{U}\{a,b\}$ and $X_2 \in \mathcal{U}\{c,d\}$, a third variable is defined as $Y = \Theta(X_1 \cdot X_2)$, where the Heaviside function is $\Theta(W)=1$ if $W>0$ and $\Theta(W)=0$ otherwise. Fixing the parameters of X_1 as $b=0.5$ and $a=b-1$, and varying those of X_2 with d in the range $[0.05:0.1:0.95]$ and $c=d-1$, the theoretical values of the joint probability distributions reported in Table 1 are obtained, and from these the marginal probabilities and the true values of the decomposition terms are computed.

Table 1. Theoretical values of $p(Y,X_1,X_2)$ as function of the parameters b and d.

X_1	X_2	$p(Y=1,X_1,X_2)$	$p(Y=0,X_1,X_2)$
<0	<0	$(b-1)(d-1)$	0
<0	>0	0	$(1-b)d$
>0	<0	0	$b(1-d)$
>0	>0	bd	0

Figure 1a shows the theoretical and estimated trends (one-hundred 300-sample realizations) of the decomposition terms of the information shared by X_1 and X_2 with Y. Regarding the estimation process, the neighbors search was performed by using $k=5$. For $d=0.5$, the simulated system appears fully synergistic, with $I(Y;X_1,X_2)$ equals to $S(Y;X_1,X_2)$ and the other terms of the decomposition equal to zero. Indeed, the value of both sources must be known to determine the state of Y [4]. By varying d, even if $I(Y;X_1,X_2)$ remains constant, the synergistic contribution of the sources decreases and the unique contribution of X_1 increases. These results are expected as, considering the extreme case of $d=1$, the only variable influencing the state of Y is X_1. Estimated quantities follow the theoretical trends, with negative bias for the measures obtained searching neighbors in the higher dimensional space, i.e., $I(Y;X_1,X_2)$ and $S(Y;X_1,X_2)$.
Figure 1b shows the trends of the specific information

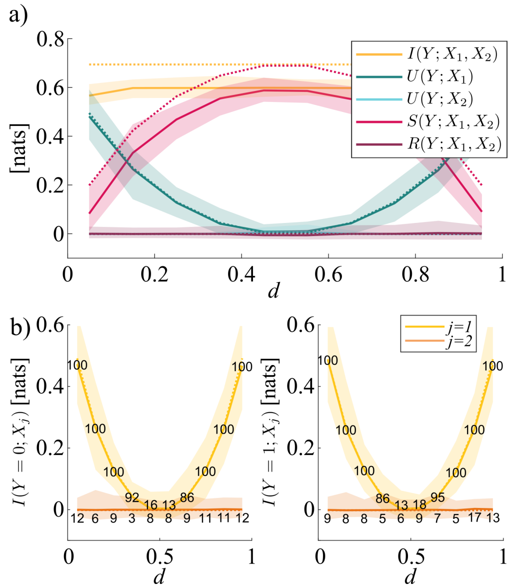

Figure 1. In (a), estimated (shadow areas with average values in continuous lines) and theoretical (dotted lines) trends of the PID terms for the logic XOR-gate between $X_1 \in \mathcal{U}\{a,b\}$ and $X_2 \in \mathcal{U}\{c,d\}$ by varying d. In (b), the specific information terms are depicted along with the percentage of significance for the estimated values.

shared by the sources and each outcome of Y and its statistical significance evaluated on one-hundred surrogate data. The results suggest that the only source sharing information with the variable Y is X_1, and that, shifting d from 0.5, the contribution of X_1 increases equally for the two possible states of Y. The surrogate data analysis confirms the absence of statistically significant values for $I(Y=i;X_1)$ for $d \sim 0.5$, and for $I(Y=i;X_2)$ for all values of d, considering both $i=0$ and $i=1$.

4. Application to physiological time series

The data analyzed in this work consist on synchronously acquired electrocardiographic (ECG; CardioFax ECG-9620, NihonKohden, Japan), arterial blood pressure (ABP; Finometer Pro, FMS, Netherlands) and respiratory volume (RespiTrace, NIMS, USA) signals (sampling frequency of 1 kHz). Sixty-one young healthy subjects (37 females, 17.5±2.4 years old) underwent a protocol involving (i) a supine resting state (REST) and (ii) an head-up tilt phase (HUT) during a 45 degree tilt to evoke postural stress. The study was approved by the ethical committee of the Jessenius Faculty of Medicine, Comenius University, Slovakia. From these signals, stationary and artifact-free variability time series of heart period, systolic arterial pressure and respiration of 300 samples were measured respectively as the time distances between consecutive ECG R peaks (series H), the local maxima of the ABP signal within each detected heart period (series S), and the values of the respiratory signal sampled at each detected ECG R peak (series R). Further details on data acquisition and preprocessing are reported in [7].

To exploit the proposed PID estimation approach for investigating the interplay between the cardiovascular dynamics and the breathing phase, the time series R was binarized defining $\bar{R}_n = 0$ if $R_n > R_{n+1}$ and $\bar{R}_n = 1$ if $R_n \leq R_{n+1}$; in this way, the expiration and inspiration phases were coded respectively with the values 0 and 1. The source time series H and S were left continuous and normalized to zero mean and unit variance.

After computing the terms of the partial information decomposition with $X_1 = [H_n\ H_{n+1}]$ and $X_2 = [S_n\ S_{n+1}]$ taken as source variables and $Y = \bar{R}_n$ taken as target variable, and fixing the number of neighbors k to 5, a statistical comparison between the two acquisition phases, i.e., REST vs HUT, was performed across subjects by using a paired Student's t-test with 5% significance. Moreover, the statistical significance of each decomposition terms for the information shared by H and S with the ventilatory activity was assessed generating one-hundred surrogate data.

Figure 2a shows mean and individual values of the unique and redundant/synergistic contributions of the information exchanged by H and S with R during REST and HUT. While an overall significant decrease of the interaction between cardiovascular and respiratory systems is detected with the postural stress, the unique contributions of cardiac and vascular dynamics have opposite trends. Besides a prevalence of cardiorespiratory over vascular-respiratory dynamics at REST, decreasing values $U(R;H)$ and increasing values of $U(R;S)$ are observed with HUT. Physiologically, these results are related to the response of the cardiorespiratory and cardiovascular control mechanisms to the orthostatic challenge. Specifically, the direct influence of the central respiratory drives on the sinus atrial node, known as respiratory sinus arrhythmia, decreases [8]. On the other hand, the increase of the information exchanged between respiration and systolic arterial pressure is related to an enhancement of the cyclic intrathoracic pressure changes and a suppressed baroreflex buffering occurring during ventilation [9]. Cardiorespiratory and vasculo-respiratory dynamics appear predominantly redundant, probably due to the influence of the ventilatory activity on both cardiac and vascular dynamics with similar latency [8, 9]. In panel b, the information exchanged individually by H and S with R and during the specific ventilatory phases are depicted. We find that the reduction of $I(R;H)$ during HUT is related to both ventilatory phases, with also a reduction of about 15% of the significant values. Conversely, the information shared by R and S increases significantly only during inspiration.

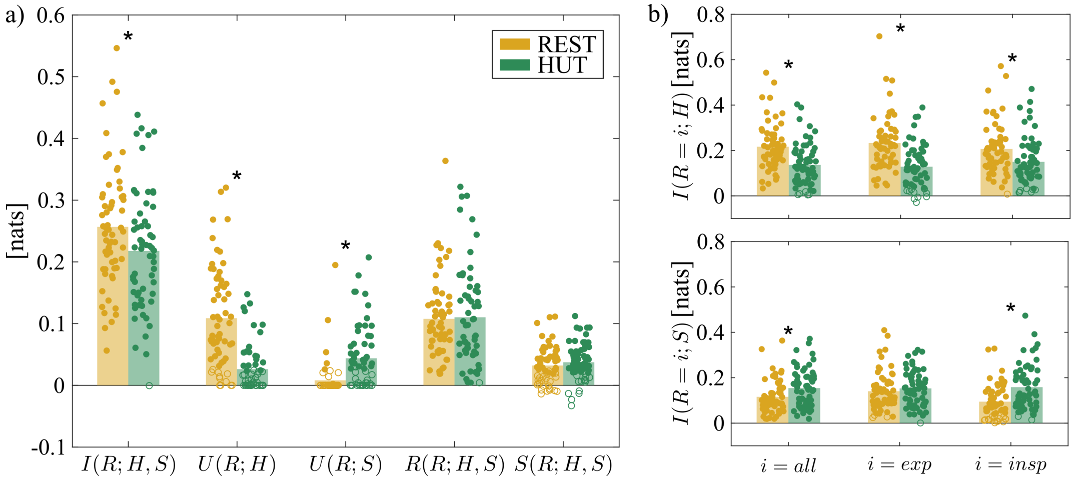

Figure 2. Mean and individual values of the PID terms considering heart period H and systolic arterial pressure S as sources and respiration R as third variable (panel a), and of the information exchanged by R and each source overall and in the ventilation phases of expiration (i.e., *exp*) and inspiration (i.e., *insp*) (panel b). Empty dots represent non-significant values with surrogate data analysis. Statistical test: *, paired Student's t-test $p < 0.05$, REST vs HUT.

5. Conclusions

Understanding how multiple control systems contribute to a pathophysiological state is crucial. Our results evidenced how the PID framework allows to distinguish the contribution of each source system to the state of the target, as well as to highlight behaviors emerging from source interactions even when the overall information shared within the network is unchanged. The proposed estimator followed accurately the ground truth values of the information terms in the simulated system, and returned unique and redundant information terms reflecting expected physiological responses to the postural stress. Further investigations are needed to assess the bias and variance of the proposed estimators as a function of data length and estimation parameters, as well as to refine the approach followed to perform surrogate data analysis.

Acknowledgments

This research was supported by the project "HONEST - High-Order Dynamical Networks in Computational Neuroscience and Physiology: an Information-Theoretic Framework", Italian Ministry of University and Research (funded by MUR, PRIN 2022, code 2022YMHNPY, CUP: B53D23003020006), and the grant VEGA no.1/0283/21.

References

[1] Ivanov PC. The New Field of Network Physiology: Building the Human Physiolome. Frontiers in Network Physiology 2021;1:711778.

[2] Faes L, Pernice R, Mijatovic G, Antonacci Y, Krohova JC, Javorka M, Porta A. Information Decomposition in the Frequency Domain: a new Framework to Study Cardiovascular and Cardiorespiratory Oscillations. Philosophical Transactions of the Royal Society A 2021;379(2212):20200250.

[3] Williams PL, Beer RD. Nonnegative Decomposition of Multivariate Information. arXiv e prints 2010;arXiv–1004.

[4] Timme N, Alford W, Flecker B, Beggs JM. Synergy, Redundancy, and Multivariate Information Measures: an Experimentalist's Perspective. Journal of computational neuroscience 2014;36:119–140.

[5] Barrett AB. Exploration of Synergistic and Redundant Information Sharing in Static and Dynamical Gaussian Systems. Physical Review E 2015;91(5):052802.

[6] Kraskov A, Stögbauer H, Grassberger P. Estimating Mutual Information. Physical review E 2004;69(6):066138.

[7] Javorka M, Krohova J, Czippelova B, Turianikova Z, Lazarova Z, Wiszt R, Faes L. Towards Understanding the Complexity of Cardiovascular Oscillations: Insights From Information Theory. Computers in biology and medicine 2018;98:48–57.

[8] Eckberg DL. Point: Counterpoint: Respiratory Sinus Arrhythmia is due to a Central Mechanism vs. Respiratory Sinus Arrhythmia is due to the Baroreflex Mechanism. Journal of applied physiology 2009;.

[9] Draghici AE, Taylor JA. The Physiological Basis and Measurement of Heart Rate Variability in Humans. Journal of physiological anthropology 2016;35:1–8.

Address for correspondence:

Chiara Barà, chiara.bara@unipa.it

Department of Engineering, University of Palermo, Viale delle Scienze, Bldg. 9, Palermo, 90128, Italy