



OPEN ACCESS

EDITED BY

Ulrich Parlitz,
Max Planck Institute for Dynamics and Self-
Organization, Germany

REVIEWED BY

Jens C. Pruessner,
McGill University, Canada
Yutaka Osada,
Tohoku University, Japan

*CORRESPONDENCE

Luca Faes,
✉ luca.faes@unipa.it

[†]These authors have contributed equally to
this work

RECEIVED 16 August 2025

REVISED 06 November 2025

ACCEPTED 13 November 2025

PUBLISHED 06 January 2026

CITATION

Sparacino L, Pinto H, Barà C, Antonacci Y,
Pernice R, Rocha AP and Faes L (2026)
Quantifying coupling and causality in dynamic
bivariate systems: a unified framework for time-
domain, spectral, and information-
theoretic analysis.
Front. Netw. Physiol. 5:1687132.
doi: 10.3389/fnetp.2025.1687132

COPYRIGHT

© 2026 Sparacino, Pinto, Barà, Antonacci,
Pernice, Rocha and Faes. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Quantifying coupling and causality in dynamic bivariate systems: a unified framework for time-domain, spectral, and information-theoretic analysis

Laura Sparacino^{1†}, Helder Pinto^{2†}, Chiara Barà¹, Yuri Antonacci¹,
Riccardo Pernice¹, Ana Paula Rocha² and Luca Faes^{1,3*}

¹Biosignals and Information Theory Laboratory, Department of Engineering, University of Palermo, Palermo, Italy, ²CMUP LASI, Departamento de Matemática, Faculdade de Ciências, Universidade do Porto, Porto, Portugal, ³Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia

Understanding the underlying dynamics of complex real-world systems, such as neurophysiological and climate systems, requires quantifying the functional interactions between the system units under different scenarios. This tutorial paper offers a comprehensive description to time, frequency and information-theoretic domain measures for assessing the interdependence between pairs of time series describing the dynamical activities of physical systems, supporting flexible and robust analyses of statistical dependencies and directional relationships. Classical time and frequency domain correlation-based measures, as well as directional approaches derived from the notion of Granger causality, are introduced and discussed, along with information-theoretic measures of symmetrical and directional coupling. Both linear model-based and non-linear model-free estimation approaches are thoroughly described, the latter including binning, permutation, and nearest-neighbour estimators. Special emphasis is placed on the description of a unified framework that establishes a connection between causal and symmetric, as well as spectral and information-theoretic measures. This framework enables the frequency-specific representation of information-theoretic metrics, allowing for a detailed investigation of oscillatory components in bivariate systems. The practical computation of the interaction measures is favoured by presenting a software toolbox and two exemplary applications to cardiovascular and climate data. By bridging theoretical concepts with practical tools, this work enables researchers to effectively investigate a wide range of dynamical behaviours in various real-world scenarios in Network Physiology and beyond.

KEYWORDS

correlation, causality, information dynamics, spectral integration property, parametric autoregressive models, network physiology

1 Introduction

A central challenge in the broad field of Network Science consists in deciphering how complex, system-wide dynamics emerge from the interactions among the units of a distributed network (Bashan et al., 2012; Barabási, 2013; Bassett and Sporns, 2017). A key strategy to tackle this problem involves quantifying bivariate interactions, that is,

measuring how pairs of individual components influence each other over time. Numerous data-driven methodologies for network inference have been developed to investigate interactions emerging from time-resolved data, enabling researchers to map the functional interdependencies that drive the dynamic behaviours of the investigated systems. For instance, in neuroscience, functional connectivity between pairs of brain regions has often been assessed via correlation-based measures, capturing co-activation patterns linked to cognition and altered in pathological conditions such as Alzheimer and schizophrenia (Buckner et al., 2005; Bassett and Sporns, 2017). In regulatory physiology, the dynamic activity of cardiovascular, cardiorespiratory and cerebrovascular systems has been widely investigated using dynamic measures of complexity and causality in different experimental conditions and patho-physiological states, evidencing well-known behaviours including frequency-specific responses along the baroreflex and the pressure-to-flow link [see, e.g., Schulz et al. (2013); Bari et al. (2016); Porta and Faes (2015); Pernice et al. (2022b); Sparacino et al. (2024a); Cairo et al. (2025)]. Also in very different fields such as earth system science, causal approaches have revealed the presence and importance of directional links, e.g., between sea-surface temperature and fish populations (Sugihara et al., 2012), as well as between atmospheric patterns and air circulation (Runge et al., 2019).

In the context of data-driven network modelling, bivariate interactions have been inferred from pairs of time series describing the dynamic activities of the two investigated systems, and generally assessed by measures of coupling and causality in the time, frequency and/or information-theoretic domains (Pereda et al., 2005; Porta and Faes, 2015; Schulz et al., 2013; Edinburgh et al., 2021; Cliff et al., 2021). Specifically, *non-directional coupling* relations between time series refer to associations which do not specify the direction of influence, rather look for symmetrical statistical dependencies between them (Faes et al., 2012a; Faes and Nollo, 2013). A common and simple approach to compute non-directional coupling in the time domain is cross-correlation, which measures the similarity between time series as a function of the time lag, capturing synchronous or time-shifted dependencies (Reinsel, 2003; Pereda et al., 2005). While straightforward and computationally efficient, cross-correlation assumes linearity, can be sensitive to noise or temporal misalignment (Cliff et al., 2021) and does not assume causality between the time series. Nevertheless, the principle of *causality* is fundamental to identify driver-response (i.e., time-lagged) relations between the series. In the linear signal processing framework, this principle has been explored with reference to the concept of Granger causality (GC), originally developed by Wiener (Wiener, 1956) and then made operative by Granger in the context of linear regression models (Granger, 1969). In particular, GC relates the presence of a cause-effect relation to two aspects: the cause must precede the effect in time and must carry unique information about the present value of the effect. This relationship is not symmetrical and can be bidirectional, thus enabling the detection of both directional and reciprocal influences (Bressler and Seth, 2011; Porta and Faes, 2015). Differently from non-directional measures, causality approaches exploiting this concept have allowed focusing on specific directional pathways of interaction within the investigated network, as widely done, e.g., in neurophysiology (Ding et al.,

2006; Porta and Faes, 2015; Seth et al., 2015; Koutlis et al., 2021; Charleston-Villalobos et al., 2023; Pernice et al., 2022a; Difrancesco et al., 2023; Pichot et al., 2024), or in climate (Smirnov and Mokhov, 2009; McGraw and Barnes, 2018; Runge et al., 2019) and ecological (Freeman, 1983; Sugihara et al., 2012) sciences.

A limitation of traditional time domain measures of coupling and causality is their lack of frequency resolution, as they capture overall dependencies between two time series without isolating contributions from specific oscillatory components. This represents an issue in fields such as cardiovascular analysis, where physiological signals including heart rate and blood pressure exhibit distinct rhythmic patterns, especially within the low-frequency (LF, 0.04 – 0.15 Hz) and high-frequency (HF, 0.15 – 0.4 Hz) bands (Cohen and Taylor, 2002). To overcome this issue, representations of coupling and causality in the frequency domain are often desirable to examine oscillatory interactions and identify the individual rhythmic components in the measured data. For this purpose, linear parametric model-based and non-parametric approaches, the latter based on the definition of the cross power spectral density as the Fourier transform of the cross-correlation function (Priestley, 1981), have been widely used to represent signals in the frequency domain. For instance, power spectral density estimates and spectral coherence (Kay, 1988), which is the frequency domain counterpart of the time domain cross-correlation, have been adopted to study a variety of patho-physiological conditions in brain (Möller et al., 2001; Coben et al., 2008; Milde et al., 2010; Wang et al., 2015), cardiovascular (Orini et al., 2011; Daoud et al., 2018) and climate (Antonacci et al., 2021b) networks. Additionally, spectral causality measures including directed coherence (Saito and Harashima, 1981; Baccalá et al., 1998) and linear feedback (Geweke, 1982) have been defined to explore frequency-specific directional patterns, and applied to a variety of neurophysiological data (Brovelli et al., 2004; Faes et al., 2005; 2012a; Barrett et al., 2012; Porta and Faes, 2013; Pernice et al., 2022b; Clemson et al., 2022).

Statistical dependencies among real-world time series can be evaluated using information-theoretic tools. Entropy-based measures of mutual information, mutual information rate and transfer entropy have been largely exploited to assess the overall information shared between two interdependent systems (Shannon, 1948; Cover, 1999), the dynamic interdependence between two systems per unit of time (Gelfand and Yaglom, 1959; Duncan, 1970; Faes et al., 2022; Barà et al., 2023; Antonacci et al., 2025; Pinto et al., 2025a), and the dynamic information transferred to a selected target system from the other connected system (Schreiber, 2000; Wibral et al., 2014; Faes et al., 2015b; 2016a; Shao et al., 2023), respectively. Remarkably, the mutual information rate represents a dynamic version of the mutual information and, as such, reflects the strength of the symmetrical statistical association between two dynamic systems. It has been decomposed into terms with meaningful physical interpretations corresponding to the well-known conditional entropy and transfer entropy measures (Barà et al., 2023; Pinto et al., 2025a). These metrics are closely related to the notions of complexity of individual systems and causality between pairs of systems, and widely exploited in various applications to real data (Faes et al., 2013c; Runge et al., 2019; Martins et al., 2020; Silini et al., 2023; Barà et al., 2023).

Information-theoretic measures have the main advantage of generality, being defined in terms of probability distributions, and can thus be stated in a fully model-free formulation (Ash, 2012). Examples of model-free estimators include the k -nearest neighbour (Kozachenko and Leonenko, 1987; Kraskov et al., 2004), permutation-based (Bandt and Pompe, 2002; Barà et al., 2023) and binning (Darbellay and Vajda, 1999; Cellucci et al., 2005; Azami et al., 2023) approaches. These methods enable the detection of non-linear and consequently more complex relationships, although they require trade-offs in terms of dimensionality, estimation accuracy, computational complexity and sensitivity to parameter choices. Nevertheless, information measures can also be expressed in terms of predictability improvement under the two key assumptions of linearity and joint Gaussianity; if this is the case, their computation relies on parametric autoregressive models (Lütkepohl, 2005; Barnett et al., 2009), whereby concepts of prediction error and conditional entropy, GC and transfer entropy, or spectral coherence and mutual information rate, have been linked to each other in the time and frequency domains (Geweke, 1982; Barnett et al., 2009; Barrett et al., 2010; Chicharro, 2011; Faes et al., 2015b; Porta and Faes, 2015; Faes et al., 2016b). Remarkably, the information-theoretic and spectral formulations are tightly connected thanks to the fulfillment of the spectral integration property, which is essential to allow quantification of these measures with reference to specific oscillatory components contained within spectral bands of interest.

In the wide context of bivariate time series analysis, this work presents a coherent theoretical framework in which measures of coupling and causality in the time, frequency, and information-theoretic domains are thoroughly reviewed and described, emphasizing properties and relations across domains (Section 1). The practical implementation of the measures is favoured by the exploitation of parametric autoregressive models (Section 2), which establish a connection between information-theoretic and spectral formulations under the assumptions of linearity and joint Gaussianity, and model-free approaches (Section 3), including techniques such as coarse-graining of embedding spaces (k -nearest neighbours estimator) or symbolic representations of the observed dynamics (binning and permutation estimators). The software and the codes relevant to the practical computation and statistical validation of the bivariate interaction measures are presented in Section 4, and collected in the BIM (Bivariate information measures) Matlab toolbox available for free download from <https://github.com/laurasparacino>. To illustrate the behaviour of the discussed measures and to showcase their implementation allowed by the BIM toolbox, illustrative examples are finally reported regarding benchmark applications to cardiovascular and climate time series (Section 5).

2 Framework for the analysis of interactions in bivariate systems

This section first introduces basic concepts of probability related to static and dynamic systems (Section 1.1) (Papoulis and Pillai, 2002). Then, it illustrates well-known correlation-related measures in the time and frequency domains (Section 1.2), as well as measures that implement the concepts of coupling and causality applied to

random variables and processes in the field of information-theory (Section 1.3). A schematic overview of these measures, emphasizing the nature (coupling vs. causality) and computation domain (time, frequency, information-theoretic) of the measures as well as their conceptual and mathematical links, is given in Figure 1.

2.1 Basic notions of probability

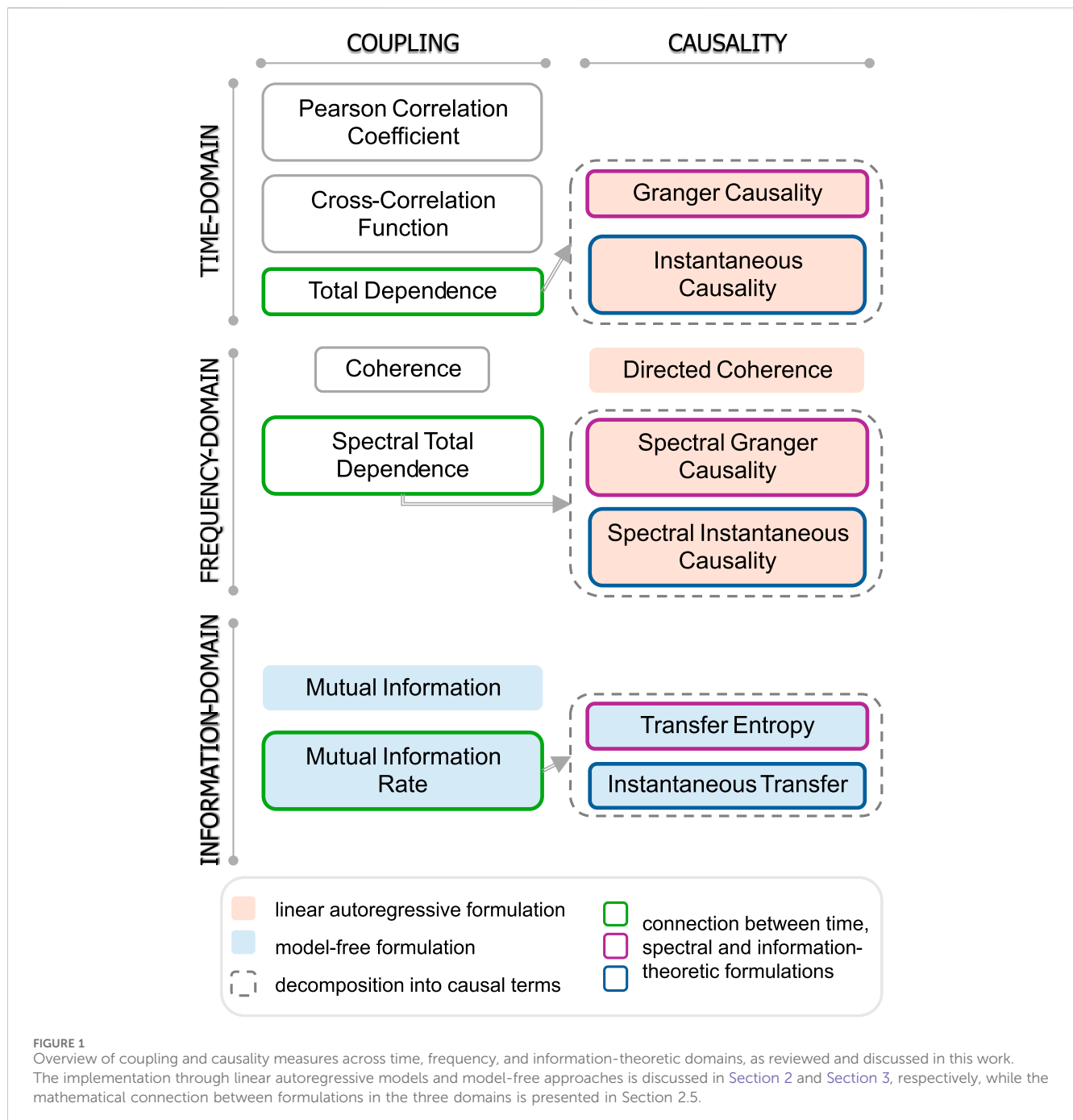
2.1.1 Static analysis of random variables

A *random variable* is a mathematical variable whose value is subject to variations due to chance. Specifically, continuous random variables can take values inside an infinite-dimensional set usually denoted as the domain. The generic scalar random variable V with domain D_V is characterized by its distribution function, which assigns a probability to each measurable subset of D_V . Formally, the probability for the variable V of taking values within the interval $[a, b] \subseteq D_V$ is determined by the integral $Pr\{a \leq V \leq b\} = \int_a^b p_V(v) dv = F_V(b) - F_V(a)$, where p_V is the *probability density function* of the variable and F_V is its cumulative distribution function. The cumulative distribution quantifies the probability that the variable V has v as its upper bound, $F_V(v) = Pr\{V \leq v\}$, while the probability density is mathematically defined as the derivative of the cumulative distribution, in a way such that $F_V(v) = \int_{-\infty}^v p_V(u) du$. These definitions extend in a straightforward way to the generic 2-dimensional variable $\mathbf{V} = \{V_1, V_2\}$ by defining the *joint probability density function* $p_{V_1, V_2}(v_1, v_2)$ and performing multiple integration over the domain of each scalar component to get the cumulative distribution. Moreover, the *conditional probability density function* of, e.g., V_1 given V_2 expresses the probability of observing the value v_1 for V_1 given that the value v_2 has been observed for V_2 : $p_{V_1|V_2}(v_1|v_2) = \frac{p_{V_1, V_2}(v_1, v_2)}{p_{V_2}(v_2)}$.

The bivariate interactions between the two variables V_1, V_2 can be investigated by means of a *static analysis* of pairs of realizations of these variables available in the form of two sequences of data. Static analysis implicitly disregards temporal correlations, assuming that all samples of a data sequence are observation of the same single random variable and thus taking into account only zero-lag effects between the two data sequences analysed.

2.1.2 Dynamic analysis of random processes

Contrarily to static systems, dynamic systems take values over diverse states at different instants of time, thus being explicitly dependent on the flow of time. The evolution over time of these systems can be only described in probabilistic terms using *random processes*, which can be thought of as sequences of random variables ordered according to time. Formally, the states visited by a generic dynamic system over time are described as a stochastic process $Y = \{Y_n\}$, $n = 1, 2, \dots$, where the random variable Y_n describes the n^{th} state assumed by the system at the n^{th} time step. Then, a realization of the stochastic process Y is the time series $y = \{y(1), \dots, y(L)\}$, containing the values of Y collected over L time points. Setting a temporal reference frame in which n represents the present time, we denote as Y_n the random variable describing the present state of Y , and as $Y_n^- = [Y_{n-1}, Y_{n-2}, \dots]$ the random variable that samples the process over the whole past history. In general, the operation of separating the present from



the past allows to consider the flow of time and to study the causal interactions within and between processes by looking at the statistical dependencies among these variables. In fact, the dynamic properties of a system are studied in statistics introducing the concept of *transition probability*, which is the probability associated with the transition of the system from its past states to its present state. Thus, the state transition of the history of Y relevant to the present state Y_n is described by the conditional probability density $p_{Y_n|Y_n^-}(y_n|y_n^-)$.

A useful property of stochastic processes is *wide-sense stationarity* (WSS), which defines the time-invariance of any joint probability density taken from the process. When the process is

stationary, its composing variables are identically distributed, meaning that the probability density is the same at all times; in practice, this allows to pool together the observations measured across time order to estimate the densities, thus enabling the estimation of probabilities from an individual realization of the process, i.e., a single time series (Faes et al., 2016b). For a stationary stochastic process, also the transition probabilities are time-independent, i.e., $p_{Y_n|Y_n^-}(y_n|y_n^-) = p_Y(y_n|y_n^-)$. An important class of dynamic processes is represented by *Markov processes*, for which the present depends on the past only through a finite number of time steps. Specifically, the process Y is a Markov process of order q if its transition probability function satisfies the condition

$p_Y(y_n|y_n^-) = p_Y(y_n|y_{n-1}, y_{n-2}, \dots, y_{n-q})$. With this notation, we define as $Y_n^q = [Y_{n-1}, \dots, Y_{n-q}]$ the random variable that samples the process over the past q lags, with $Y_n^- = \lim_{q \rightarrow \infty} Y_n^q$.

The bivariate interactions between two generic processes $\mathbf{Y} = \{Y_1, Y_2\}$ can be investigated by means of a *dynamic analysis*, where each process Y_i ($i = 1, 2$) describes the dynamic activity at the i^{th} node. We remark that, in the context of bivariate time series analysis, when studying two dynamic processes (e.g., heart rate and blood pressure in physiology), researchers often look at how the oscillatory amplitudes in one signal relate to oscillatory amplitudes in another. Standard bivariate methods might use cross-correlation or coherence, which look at magnitude relationships without explicitly considering directional/phase aspects of the oscillation. In this context, it is worth mentioning the complex/rotary perspective, which usually refers to a way of representing oscillatory signals using complex numbers or rotary (circular) components, rather than just their real-valued time series (Sykulski et al., 2017). From a *complex* perspective, the signal $Y(t)$, with t the temporal variable, is represented in the complex plane, usually via the analytic signal $Y_a(t)$ obtained from its Hilbert transform $\hat{Y}(t)$, i.e., $Y_a(t) = Y(t) + j\hat{Y}(t)$, with $j = \sqrt{-1}$. This captures both amplitude (magnitude) and phase (angle), giving a full oscillatory description of the signal; coupling metrics can be computed if treating each signal as complex separately, and take into account phase relationships, directionality, and rotations in phase space rather than just co-fluctuations in time (Rosenblum and Kurths, 1998). On the other hand, the *rotary* perspective looks at the overall bivariate oscillation a rotating vector (phasor) in the complex plane. Bivariate signals (e.g., $Y_1(t)$ and $Y_2(t)$) are often viewed together as a single complex signal, i.e., $Z(t) = Y_1(t) + jY_2(t)$. The signal $Z(t)$ is decomposed into rotary components, namely, a positive-frequency (counter-clockwise) component and a negative-frequency (clockwise) component, to capture directionality of rotation, lead-lag relationships, and phase differences between the two signals that standard linear correlations would miss (Sykulski et al., 2017). Overall, these two perspectives can reveal directionality in oscillatory processes, e.g., one physiological variable leading or lagging another, forming loops in the phase plane. In spite of its usefulness and broad applicability, this area of research will not be deepened in the present work. Rather, we assume that the processes are *real-valued*, defined at discrete time ($Y_i = \{Y_{i,n}\}$; e.g., are sampled versions of the continuous time processes $Y_i(t)$, taken at the times $t_n = nT$, with T the sampling period) and have zero mean ($\mathbb{E}[Y_{i,n}] = 0$, where $\mathbb{E}[\cdot]$ is the statistical expectation operator).

2.2 Correlation-based measures in the time and frequency domain

In the time domain, the simplest way to identify symmetric statistical dependencies between signals is through *correlation*. The commonly used static approach disregards temporal dependencies and considers the zero-lag interaction between two random variables V_1 and V_2 whose observations are taken simultaneously (i.e., at the same time instant) from the two analysed time series. This interaction is typically quantified using the *Pearson correlation*

coefficient (PCC) ρ_{V_1, V_2} , defined as the ratio between the covariance $\text{cov}(\cdot, \cdot)$ of the two variables and the product of their standard deviations σ :

$$\rho_{V_1, V_2} = \frac{\text{cov}(V_1, V_2)}{\sigma_{V_1} \sigma_{V_2}}, \quad (1)$$

where $\text{cov}(V_1, V_2) = \mathbb{E}[(V_1 - m_{V_1})(V_2 - m_{V_2})]$; $m_V = \mathbb{E}[V]$ is the mean of V . Varying within the range $[-1, 1]$, ρ_{V_1, V_2} quantifies the strength of the linear relationship between the variables, but cannot fully capture non-linear associations (Benesty et al., 2009).

To account for time-lagged dependencies, the time series are considered as realizations of two random processes Y_1 and Y_2 , from which the *cross-correlation function* (CCF) computed as a function of the time lag k as (Kay, 1988; Rodgers and Nicewander, 1988).

$$R_{Y_1 Y_2}(k) = \mathbb{E}[Y_{1,n} Y_{2,n-k}]; \quad (2)$$

while the CCF in Equation 2 generally a function also of the time instant n , such dependence is omitted in Equation 2 under the hypothesis of WSS processes. The CCF captures how past values of Y_2 relate to current values of Y_1 in a linear signal processing framework, enabling the detection of lead-lag relationships. When estimated from data, this function helps identifying temporal dependencies and is often used as a precursor to more complex dynamic measures.

The two observed random processes can be studied in the frequency domain in terms of the power spectral density (PSD) matrix of the stationary bivariate random process $\mathbf{Y} = \{Y_1, Y_2\}$. The PSD matrix, denoted as $\mathbf{P}_Y(\bar{f})$, is a 2×2 matrix which contains the individual PSD of the process Y_i , $P_{Y_i}(\bar{f})$, as i^{th} diagonal element and the cross-PSD between the processes Y_i and Y_j , $P_{Y_i Y_j}(\bar{f})$, as off-diagonal elements in the position $i - j$ ($i, j = 1, 2$); $\omega \in [-\pi, \pi]$ is the normalized angular frequency ($\omega = 2\pi \frac{f}{f_s} = 2\pi \bar{f}$ with $f \in [-\frac{f_s}{2}, \frac{f_s}{2}]$; $\bar{f}$ normalized frequency, f_s sampling frequency):

$$\mathbf{P}_Y(\bar{f}) = \begin{bmatrix} P_{Y_1}(\bar{f}) & P_{Y_1 Y_2}(\bar{f}) \\ P_{Y_2 Y_1}(\bar{f}) & P_{Y_2}(\bar{f}) \end{bmatrix}. \quad (3)$$

The link between the time and frequency domain representations is provided by the Fourier Transform (FT) $\mathfrak{F}\{\cdot\}$. Specifically, the PSD of the process Y_i , ($i = 1, 2$), is defined as $P_{Y_i}(\bar{f}) = \mathfrak{F}\{R_{Y_i}(k)\}$, with $R_{Y_i}(k) = \mathbb{E}[Y_{i,n} Y_{i,n-k}]$ the autocorrelation function of Y_i . Similarly, the cross-PSD between the two processes is defined as the FT of their CCF (Kay, 1988; Priestley, 1981).

$$P_{Y_1 Y_2}(\bar{f}) = \mathfrak{F}\{R_{Y_1 Y_2}(k)\} = \sum_{k=-\infty}^{\infty} R_{Y_1 Y_2}(k) e^{-j2\pi \bar{f} k}, \quad (4)$$

where $j = \sqrt{-1}$; the PSD of Equation 2 represents a fundamental tool in frequency domain analysis quantifying the extent to which two jointly stationary stochastic processes Y_1 and Y_2 co-vary as a function of frequency. By capturing both the magnitude and phase relationships of their frequency components, the cross-PSD provides insights into the presence, strength, and timing of linear interactions across different spectral bands, making it especially valuable in the analysis of oscillatory coupling in complex systems such as physiological (Faes et al., 2022; Sparacino et al., 2025a) or neural networks (Antonacci et al., 2021b). The cross-PSD is generally complex-valued, with its

magnitude describing the co-oscillatory power at frequency $\bar{f}$, and its phase indicating the relative timing or phase delay between the two signals at that frequency.

The *coherence* (Coh) between the two processes Y_1 and Y_2 can be defined as the ratio between the cross-spectrum $P_{Y_1 Y_2}(\bar{f})$ and the squared root of the product between the autospectra of Y_1 and Y_2

$$\Gamma_{Y_1 Y_2}(\bar{f}) = \frac{P_{Y_1 Y_2}(\bar{f})}{\sqrt{P_{Y_1}(\bar{f})P_{Y_2}(\bar{f})}} \quad (5)$$

Since this function is complex-valued, its squared modulus is commonly used to measure the strength of coupling in the frequency domain. Indeed, the magnitude squared coherence, $|\Gamma_{Y_1 Y_2}(\bar{f})|^2$, has a meaningful physical interpretation, since it measures the strength of the linear, non-directional coupled interactions between the processes Y_1 and Y_2 as a function of frequency. The magnitude squared coherence is a normalized measure of coupling, being 0 at a given frequency in case of uncoupling and 1 in case of deterministic linear dependence. Another popular, non-normalized spectral measure of coupling between two processes Y_1 and Y_2 is the logarithmic measure of *total dependence* (TD) defined by Geweke (1982) as

$$f_{Y_1 Y_2}(\bar{f}) = \log \frac{P_{Y_1}(\bar{f})P_{Y_2}(\bar{f})}{|P_Y(\bar{f})|}; \quad (6)$$

exploiting Equation 5, it can be shown that the squared Coh is related to the logarithmic TD measure through the relation

$$f_{Y_1 Y_2}(\bar{f}) = -\log(1 - |\Gamma_{Y_1 Y_2}(\bar{f})|^2). \quad (7)$$

In practice, estimation of the PSD from finite-length time series can be achieved through both parametric (Section 2.4) and non-parametric (Section 4) methods, each with its own strengths and weaknesses (Kay, 1988; Pinna et al., 1996; Zhao and Gui, 2019). The choice of the method depends on the characteristics of the signal, the available data, and the specific requirements of the analysis, being often a trade-off between frequency resolution, variance reduction, and computational complexity.

2.3 Information-theoretic measures of coupling and causality

In this section, we present the well-known information-theoretic measures used for analyzing undirected and directed interactions in bivariate systems. Each of the measures is described with detail; the characterization of their mathematical relationships allows to highlight how they capture distinct aspects of statistical dependence.

2.3.1 Mutual information and mutual information rate

The most popular measure of coupling derived in the frame of information theory is the *mutual information* (MI). The MI is a symmetric measure quantifying the amount of information shared by two random variables V_1 and V_2 , intended as the uncertainty about one of the variables that is resolved by knowing the other. Formally, the MI is defined as (Shannon, 1948; Cover, 1999):

$$I(V_1; V_2) = \mathbb{E} \left[\log \frac{p_{V_1 V_2}(v_1, v_2)}{p_{V_1}(v_1)p_{V_2}(v_2)} \right], \quad (8)$$

where $p(\cdot, \cdot)$ and $p(\cdot)$ represent the joint and marginal probability distributions, respectively. The MI defined in Equation 8 can be equivalently expressed in terms of Shannon entropies as:

$$I(V_1; V_2) = H(V_1) + H(V_2) - H(V_1, V_2), \quad (9)$$

where $H(\cdot)$ denotes entropy and $H(\cdot, \cdot)$ denotes joint entropy (Shannon, 1948; Cover, 1999). The MI is intimately linked to the PCC (Equation 1) when the two random variables V_1 and V_2 have a joint Gaussian distribution; if this is the case, the MI between V_1 and V_2 can be expressed analytically as (Jwo et al., 2023):

$$I(V_1; V_2) = -\frac{1}{2} \log(1 - \rho_{V_1 V_2}^2); \quad (10)$$

Equation 10 establishes a direct and quantitative link between linear correlation and information-theoretic dependence. It is worth stressing that the MI, like the PCC, measures the static interaction between two random variables.

The *mutual information rate* (MIR) generalizes the MI between random variables by quantifying the dynamic coupling between the two stochastic processes Y_1 and Y_2 . Specifically, the MIR measures the amount of information shared per unit of time between the processes, and is defined as (Duncan, 1970; Geweke, 1982; Sparacino et al., 2025a):

$$\begin{aligned} I_{Y_1 Y_2} &= \lim_{q \rightarrow \infty} \frac{1}{q} I(Y_{1,n}^q; Y_{2,n}^q) \\ &= I(Y_{1,n}, Y_{1,n}^-; Y_{2,n}, Y_{2,n}^-) - I(Y_{1,n}^-, Y_{2,n}^-), \end{aligned} \quad (11)$$

where $I(\cdot; \cdot)$ denotes MI. In essence, the MIR quantifies the overall coupling between the two processes elaborating the MI between all their constituent variables; note that the MI computed between a sequence of random variables composing two processes needs to be divided by the number of these variables to yield a non-diverging quantity. As the MI, the MIR is a symmetric measure (i.e., $I_{Y_1 Y_2} = I_{Y_2 Y_1}$) and is widely used to characterize bivariate dynamic interactions across various scientific domains [see, e.g., Baptista and Kurths (2008); Pinto et al. (2025d); Sparacino et al. (2025a)]. The MIR can be straightforwardly decomposed into information-theoretic quantities that offer valuable insights into the complex dynamics of the individual components within a bivariate system:

$$I_{Y_1 Y_2} = H_{Y_1} + H_{Y_2} - H_{Y_1 Y_2}, \quad (12)$$

where $H_{Y_1} = H(Y_{1,n} | Y_{1,n}^-)$ and $H_{Y_2} = H(Y_{2,n} | Y_{2,n}^-)$ denote the entropy rates of Y_1 and Y_2 , respectively, and $H_{Y_1 Y_2} = H(Y_{1,n}, Y_{2,n} | Y_{1,n}^-, Y_{2,n}^-)$ their joint entropy rate (with $H(\cdot | \cdot)$ denoting conditional entropy) (Cover, 1999; Chicharro, 2011). Note that the MIR expressed as in Equation 12 is formally equivalent to the MI expressed as in Equation 9, with the use of entropy rate terms in place of standard entropy terms. Another important meaningful decomposition of the MIR is derived in the next subsection.

We anticipate that, in the case of stationary Gaussian processes, the MIR is closely connected to a well-known time-domain measure of non-directional dynamic coupling related to the concept of TD developed in the context of linear regression models (Geweke, 1982); the relevant derivations will be provided in Section 2.

2.3.2 Causal and instantaneous information transfer

Inferring directional interactions between time series is a fundamental task in the analysis of dynamical systems. Generally, the two random processes representing the dynamic activity of the units interact in a closed-loop manner, i.e., through bidirectional and reciprocal influences which allow to identify asymmetrical driver-response patterns (Porta and Faes, 2015).

In the field of information theory, a widely used measure for assessing such interactions is *transfer entropy* (TE), formally defined as (Schreiber, 2000):

$$\begin{aligned} T_{Y_1 \rightarrow Y_2} &= I(Y_{2,n}; Y_{1,n}^- | Y_{2,n}^-) = H(Y_{2,n} | Y_{2,n}^-) - H(Y_{2,n} | Y_{1,n}^-, Y_{2,n}^-) \\ &= H(Y_{2,n}, Y_{2,n}^-) - H(Y_{2,n}^-) - H(Y_{2,n}, Y_{1,n}^-, Y_{2,n}^-) + H(Y_{1,n}^-, Y_{2,n}^-), \end{aligned} \quad (13)$$

where $I(\cdot; \cdot | \cdot)$ denotes conditional MI; herein, Y_1 is assumed as the *driver* process, while Y_2 as the *target* process.

Remarkably, the MIR can be formulated comparing the sum of the 2 TEs from Y_1 to Y_2 and from Y_2 to Y_1 (Barà et al., 2023):

$$I_{Y_1, Y_2} = T_{Y_1 \rightarrow Y_2} + T_{Y_2 \rightarrow Y_1} + I_{Y_1, Y_2}, \quad (14)$$

where $T_{Y_1 \rightarrow Y_2}$ and $T_{Y_2 \rightarrow Y_1}$ represent the directional information flow from Y_1 to Y_2 and from Y_2 to Y_1 , respectively. The term I_{Y_1, Y_2} , commonly referred to as *instantaneous transfer* (IT) (Chicharro, 2011), captures the instantaneous, bidirectional exchange of information between Y_1 and Y_2 and is defined as:

$$\begin{aligned} I_{Y_1, Y_2} &= I(Y_{1,n}; Y_{2,n} | Y_{1,n}^-, Y_{2,n}^-) \\ &= H(Y_{1,n} | Y_{1,n}^-, Y_{2,n}^-) - H(Y_{1,n} | Y_{2,n}, Y_{1,n}^-, Y_{2,n}^-). \end{aligned} \quad (15)$$

We remark that the TE is a well-known measure of directional information transfer related to the concept of Granger causality (GC) originally developed by Wiener (Wiener, 1956) and then made operative by Granger in the context of linear regression models (Granger, 1969), while the IT is a symmetric measure related to the concept of instantaneous causality (IC) (Chicharro, 2011).

3 Implementation through linear autoregressive models

This section presents the definition and practical implementation of the time, spectral and information-theoretic measures of directed and undirected coupling obtained through linear regression methods making use of univariate and bivariate autoregressive (AR) models. Linear AR models are ubiquitously used to assess dynamics and interactions in time series data, especially in the field of Network Physiology (Bressler and Seth, 2011; Seth et al., 2015; Porta and Faes, 2015; Faes et al., 2016b; Sparacino et al., 2024b; Vakitbilir et al., 2025). Here, we begin discussing three issues that have theoretical relevance and practical implications in the use of AR models for the computation of interaction measures.

The first observation is that fitting linear AR models on time series assumes linearity in the modelled interactions, but not in the processes to be analysed: indeed, while the model assumes linearity in its structure, this does not necessarily imply that the underlying time series must be linear (Barnett and Seth, 2023). Indeed, Wold's decomposition theorem (Wold, 1938) guarantees that any

stationary process can be decomposed into a linear model, although this model may be of infinite order and thus providing a non-parsimonious representation of an underlying nonlinear process (Hannan, 1979). Therefore, linear AR models may in principle be able to describe also the dynamics of processes with nonlinear generating mechanisms.

Another relevant observation is that, while linear AR models can be used to describe time series with any probability distribution, when the underlying processes are jointly Gaussian distributed the measures derived in the time domain (Barrett et al., 2010; Faes et al., 2016b) and in the spectral domain (Chicharro, 2011; Faes et al., 2021; Antonacci et al., 2021b; Sparacino et al., 2025a) have a striking information-theoretic interpretation. In fact, the parametric implementation exploits the knowledge that linear regression models capture all of the entropy differences relevant to the various information measures when the observed processes have a joint Gaussian distribution (Barrett et al., 2010; Faes et al., 2016b).

The third point regards the fact that linear AR models typically limit to past values only the possible influences of one process to another, thereby excluding instantaneous effects (i.e., effects occurring within the same lag) from the model structure (Lütkepohl, 2005). The consequence of this is that the model residuals (prediction errors) are correlated whenever instantaneous effects are present between the analysed time series. On the other hand, the absence of instantaneous effects, typically denoted as *strict causality* of the process (Korhonen et al., 1996; Baselli et al., 1997) implies that the covariance matrix of the residuals is diagonal. While strict causality is often assumed in the computation of causality measures, the presence of instantaneous effects has an impact on the derived measures. In the following subsections we will discuss such an impact and mention how bivariate measures coupling and causality measures can be corrected to account for instantaneous effects.

3.1 Formulation of linear parametric models of bivariate time series

The linear formulation leading to compute coupling measures requires identification of a bivariate AR model composed by two so-called full *auto- and cross-regressive* (ARX) models, from which restricted *autoregressive* (AR) models are derived to compute causality measures. Full ARX models feature two model equations, where the present states of the two processes are written as linear combinations of the past states of both processes weighted by a set of model coefficients plus the residuals. Assuming that $\mathbf{Y}$ is the generic vector process comprising the two scalar processes $\{Y_1, Y_2\}$, the following ARX model can be identified (Lütkepohl, 2005):

$$Y_{1,n} = \sum_{k=1}^p a_{Y_1 Y_1, k} Y_{1, n-k} + a_{Y_1 Y_2, k} Y_{2, n-k} + U_{Y_1, n} \quad (16a)$$

$$Y_{2,n} = \sum_{k=1}^p a_{Y_2 Y_1, k} Y_{1, n-k} + a_{Y_2 Y_2, k} Y_{2, n-k} + U_{Y_2, n} \quad (16b)$$

where $\mathbf{Y}_n = [Y_{1,n}, Y_{2,n}]^\top$ is the 2-dimensional vector collecting the present state of the two processes, $\mathbf{A}_{Y,k} = \begin{bmatrix} a_{Y_1 Y_1, k} & a_{Y_1 Y_2, k} \\ a_{Y_2 Y_1, k} & a_{Y_2 Y_2, k} \end{bmatrix}$ is

the 2×2 matrix of the model coefficients relating the present with the past of the two processes at lag k , and $\mathbf{U}_{Y,n} = [U_{Y_1,n}, U_{Y_2,n}]^\top$ a vector of 2 zero-mean white noises with 2×2 positive definite covariance matrix $\Sigma_{\mathbf{U}_Y} = \begin{bmatrix} \sigma_{U_{Y_1}}^2 & \sigma_{U_{Y_1}Y_2}^2 \\ \sigma_{U_{Y_2}Y_1}^2 & \sigma_{U_{Y_2}}^2 \end{bmatrix}$. Note that $\sigma_{U_{Y_1}Y_2}^2 = \sigma_{U_{Y_2}Y_1}^2 = 0$ in the case of strict causality, implying the absence of instantaneous interactions. The process $\mathbf{Y}_n$ has a 2×2 covariance matrix $\Sigma_Y = \mathbb{E}[\mathbf{Y}_n \mathbf{Y}_n^\top]$, where the diagonal elements represent the variances of the scalar processes $\{Y_1, Y_2\}$, i.e., $\sigma_{Y_1}^2, \sigma_{Y_2}^2$.

3.1.1 Model identification

The identification procedure of the ARX model in Equations 16a, b is typically performed by means of estimation methods based on minimizing the prediction error, i.e., the difference between actual and predicted data (Kay, 1988; Lütkepohl, 2005). While several approaches have been proposed throughout the years (Schlögl, 2006; Antonacci et al., 2020), the most common estimator is the multivariate version of the *ordinary least-squares* (OLS) method (Lütkepohl, 2005). Briefly, defining the past history of $\mathbf{Y}$ truncated at p lags as the $2p$ -dimensional vector $\mathbf{Y}_n^p = [\mathbf{Y}_{n-1}^\top, \dots, \mathbf{Y}_{n-p}^\top]^\top$ and considering L consecutive time steps, a compact representation of Equations 16a, b can be defined as $\mathbf{Y} = \mathbf{A}_Y \mathbf{Y}^p + \mathbf{U}_Y$, where $\mathbf{A}_Y = [\mathbf{A}_{Y,1}, \dots, \mathbf{A}_{Y,p}]$ is the $2 \times 2p$ matrix of unknown coefficients, $\mathbf{Y} = [\mathbf{Y}_{p+1}, \dots, \mathbf{Y}_L]$ and $\mathbf{U}_Y = [\mathbf{U}_{Y,p+1}, \dots, \mathbf{U}_{Y,L}]$ are $2 \times (L - p)$ matrices, and $\mathbf{Y}^p = [\mathbf{Y}_{p+1}^p, \dots, \mathbf{Y}_L^p]$ is a $2p \times (L - p)$ matrix collecting the regressors. The method estimates the coefficient matrices through the OLS formula, $\hat{\mathbf{A}}_Y = \mathbf{Y}(\mathbf{Y}^p)^\top [\mathbf{Y}^p(\mathbf{Y}^p)^\top]^{-1}$. The innovation process is estimated as the residual time series $\hat{\mathbf{U}}_Y = \mathbf{Y} - \hat{\mathbf{A}}_Y \mathbf{Y}^p$, whose covariance matrix $\hat{\Sigma}_{\mathbf{U}_Y}$ is an estimate of $\Sigma_{\mathbf{U}_Y}$. After identification, the model in Equations 16a, b can be analyzed in the frequency domain.

As regards the selection of the model order p , several criteria exist for its determination [see, e.g., Lütkepohl (2005); Karimi (2011)]. One commonly used approach is to set the order according to the *Akaike Information Criterion* (AIC) (Akaike, 1974), or the *Bayesian Information Criterion* (BIC) (Schwarz, 1978). The primary difference between AIC and BIC lies in how they penalize model complexity and their underlying theoretical foundations. AIC is based on information theory and aims to minimize the information lost when using a model to approximate the true process. It focuses on predictive accuracy and is more likely to select models that perform well for future data. The penalty for the number of parameters is more moderate with AIC, which indeed balances goodness of fit with model complexity but places less emphasis on the number of parameters. On the other hand, BIC approximates the posterior probability of a model given the data, and seeks the model that is most likely to be the true one, based on the given data, and is more concerned with identifying the correct model. Penalty grows with the sample size using BIC, which heavily penalizes models with more parameters, especially in larger datasets. Practically, it is crucial to find the right balance between excessively low orders, which might lead to an inadequate description of crucial oscillatory information in the vector process, and

overly high orders, which could result in overfitting, with the outcome that the model captures not only the desired information but also includes noise. Additional guidelines on model order selection include 1) utilizing multiple selection criteria and checking for consistency; if they converge on similar orders, confidence in the choice increases. 2) Fitting models at nearby orders ($\pm 1 - 2$) could help in assessing stability, since large changes in spectral measures or AR coefficients suggest sensitivity to model order. 3) Finally, visually inspecting the AR-derived PSD may be helpful, especially in the case of physiological variables whose spectral behaviour is generally well-known: overly smooth spectra may indicate too low orders, while unrealistic peaks or high-frequency noise suggest overfitting (Pernice et al., 2022a). As regards possible non-stationarity in the data, cutting large non-stationary physiological signals to identify shorter stationary segments may be a practical solution; this reduces sensitivity to model order, as the local dynamics can be captured with lower orders. As an example, for short-window heart rate variability series (on the order of ~ 300 beats), parametric modelling via AR methods has been used for complexity and spectral assessment [e.g., Porta et al. (2002b), Porta et al. (2017)]. In keeping with the precedent of moderate orders in such short windows, model orders in the range 5 – 8 are a good compromise between spectral resolution and robustness (see Section 5.1).

3.1.2 Restricted AR model

While the ARX model in Equations 16a, b provides a global representation of the overall bivariate process, to describe the linear interactions relevant to, e.g., the target process, we need to define a restricted AR model involving only Y_2 . To implement this concept, the present state of the target, $Y_{2,n}$, is described first from the past of Y_2 only through the *restricted AR model*

$$Y_{2,n} = \sum_{k=1}^{\infty} b_{Y_2 Y_2, k} Y_{2,n-k} + W_{Y_2, n} \quad (17)$$

where $b_{Y_2 Y_2, k}$ are AR coefficients and W_{Y_2} is a white noise process with variance $\lambda_{W_{Y_2}}^2$.

An issue with great practical relevance is that the order of the restricted model in Equation 17 is typically infinite and thus very difficult to identify from finite-length time series. The approach followed to face this issue in the context of causality analysis is essentially based on truncating the order of the restricted model to p , and estimating its parameters from the relevant subset of the original data. Though simple, this approach exposes to a trade-off between bias and variance of the estimates that prevents reliable model identification in most cases (Stokes and Purdon, 2017; Faes et al., 2017b). To solve this issue, methods which essentially extract the parameters of the restricted model from those of the full model have been proposed, i.e., methods based on *state-space* (SS) models (Faes et al., 2017b; Barnett et al., 2018) and on the resolution of the *Yule-Walker* (YW) equations (Barnett and Seth, 2014; Faes et al., 2015b; 2016b; Sparacino et al., 2024a). In the following, the two methods will be thoroughly described.

3.2 Identification of restricted models

3.2.1 State-space models

The method based on SS models can be applied to the bivariate AR model in Equations 16a,b to derive the parameters of the two corresponding restricted AR models of Y_1, Y_2 in the form of Equation 17, i.e., $\{b_{Y_1 Y_1}, \lambda_{W_{Y_1}}^2\}$ and $\{b_{Y_2 Y_2}, \lambda_{W_{Y_2}}^2\}$, respectively (Barnett and Seth, 2015). This class of models is the most appropriate to use because it is closed under the formation of restricted models: in fact, any restricted process obtained from a vector AR process is actually an AR process with a moving average component, or equivalently a finite-order SS process (Barnett et al., 2018). Therefore, using SS models allows to identify restricted models from the parameters of the original vector AR model estimated with a single regression, thus guaranteeing high computational reliability. We exploit SS modelling to describe the original process $\mathbf{Y}$ obeying the bivariate AR representation in Equations 16a,b using the SS model.

$$\mathbf{S}_{n+1} = \mathbf{A}\mathbf{S}_n + \mathbf{K}\mathbf{U}_{Y,n}, \quad (18a)$$

$$\mathbf{Y}_n = \mathbf{C}\mathbf{S}_n + \mathbf{U}_{Y,n}, \quad (18b)$$

where $\mathbf{S}_n = [\mathbf{Y}_{n-1}^\top, \dots, \mathbf{Y}_{n-p}^\top]^\top$ is the $2p$ -dimensional state process and the SS parameters $(\mathbf{A}, \mathbf{C}, \mathbf{K}, \mathbf{V})$ are given by the matrices $\mathbf{C} = [\mathbf{A}_{Y,1}, \dots, \mathbf{A}_{Y,p}]$, $\mathbf{K} = [\mathbf{I}_{2 \times 2(p-1)}]^\top$, $\mathbf{A} = [\mathbf{C}; \mathbf{I}_{2(p-1)} \mathbf{0}_{2(p-1) \times 2}]$, and $\mathbf{V} = \mathbb{E}[\mathbf{U}_{Y,n} \mathbf{U}_{Y,n}^\top] = \Sigma_{U_Y}$ ($\mathbf{I}$ and $\mathbf{0}$ are the identity and null matrices, respectively). Then, to represent the scalar process Y_2 , we replace Equation 17 with a restricted SS model with state equation as in Equation 18a and observation equation $Y_{2,n} = \mathbf{C}^{(2,:)} \mathbf{S}_n + W_{Y_2,n}$. The parameters of the model are $(\mathbf{A}, \mathbf{C}^{(2,:)}, \mathbf{K}\mathbf{V}\mathbf{K}^\top, \mathbf{V}^{(2,2)}, \mathbf{K}\mathbf{V}^{(:,2)})$, where the superscripts denote selection of the rows and/or columns with indices 2 in a matrix. To exploit the restricted SS model for the linear causality analysis of Y_2 it is necessary to lead its form back to that of Equations 18a,b, which reads (Barnett and Seth, 2015)

$$\mathbf{S}_{n+1} = \tilde{\mathbf{A}}\mathbf{S}_n + \tilde{\mathbf{K}}W_{Y_2,n}, \quad (19a)$$

$$Y_{2,n} = \tilde{\mathbf{C}}\mathbf{S}_n + W_{Y_2,n}. \quad (19b)$$

The parameters of the restricted model in Equations 19a,b are $(\tilde{\mathbf{A}}, \tilde{\mathbf{C}}, \tilde{\mathbf{K}}, \tilde{\mathbf{V}})$, of dimension $2p \times 2p, 1 \times 2p, 2p \times 1, 1 \times 1$, and can be derived directly from the parameters $\mathbf{A}_{Y,k}$ and Σ_{U_Y} of the original full ARX model in Equations 16a,b (Barnett and Seth, 2015): while the state and observation matrices are easily determined as $\tilde{\mathbf{A}} = \mathbf{A}$ and $\tilde{\mathbf{C}} = \mathbf{C}^{(2,:)}$, the gain $\tilde{\mathbf{K}}$ and the restricted innovation covariance $\tilde{\mathbf{V}} = \lambda_{W_{Y_2}}^2$ must be obtained by solving a discrete algebraic Riccati equation (DARE) (see Barnett and Seth (2015); Faes et al. (2017a) for detailed derivations). After identification, the model in Equations 19a,b can be analyzed in the frequency domain to study spectral interactions relevant to the process Y_2 .

3.2.2 Resolution of the Yule-Walker equations

The issue related to the formation of AR restricted models from the ARX model in Equations 16a,b can be overcome also by solving the YW equations. The restricted model coefficients, $b_{Y_2 Y_2, k}$, and the variance of the residuals, $\lambda_{W_{Y_2}}^2$, appearing in Equation 17, can be identified starting from the covariance and cross-covariance matrices between the present and past variables of the two scalar processes Y_1 and Y_2 . Using these matrices allows to identify

restricted models from the parameters of the original ARX model estimated with a single regression up to an arbitrarily large order q , thus guaranteeing computational reliability (Barnett and Seth, 2014; Faes et al., 2015b; 2016b; Mijatovic et al., 2025). For jointly Gaussian processes, these matrices contain as scalar elements the covariance between two time-lagged variables taken from the processes Y_1 and Y_2 , which in turn appear as elements of the 2×2 autocovariance of the whole observed 2-dimensional process $\mathbf{Y}_n = [Y_{1,n} Y_{2,n}]^\top$, defined at each lag $k \geq 0$ as $\Gamma_k = \mathbb{E}[\mathbf{Y}_n \mathbf{Y}_{n-k}^\top]$. The procedure exploits the possibility to compute Γ_k from the parameters of the ARX formulation of the process $\mathbf{Y}_n$ via the well-known YW equations:

$$\Gamma_k = \sum_{l=1}^p \mathbf{A}_{Y,l} \Gamma_{k-l} + \delta_{k0} \Sigma_{U_Y}, \quad (20)$$

where δ_{k0} is the Kronecker delta function. In order to solve Equation 20 for Γ_k , with $k = 0, \dots, p-1$, we first express the ARX model in Equations 16a,b in compact form as $\psi_n = \mathbf{A}\psi_{n-1} + \mathbf{E}_n$, where:

$$\begin{aligned} \psi_n &= [\mathbf{Y}_n^\top \mathbf{Y}_{n-1}^\top, \dots, \mathbf{Y}_{n-p+1}^\top]^\top; \\ \mathbf{A} &= \begin{bmatrix} \mathbf{A}_{Y,1} & \dots & \mathbf{A}_{Y,p-1} & \mathbf{A}_{Y,p} \\ \mathbf{I} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \dots & \mathbf{I} & \mathbf{0} \end{bmatrix}; \\ \mathbf{E}_n &= [\mathbf{U}_{Y,n}^\top \mathbf{0}_{1 \times 2(p-1)}]^\top. \end{aligned} \quad (21)$$

From Equation 21, the $2p \times 2p$ covariance matrix of ψ_n , which is defined as $\Psi = \mathbb{E}[\psi_n \psi_n^\top]$ and has the form

$$\Psi = \begin{bmatrix} \Gamma_0 & \Gamma_1 & \dots & \Gamma_{p-1} \\ \Gamma_1^\top & \Gamma_0 & \dots & \Gamma_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{p-1}^\top & \Gamma_{p-2}^\top & \dots & \Gamma_0 \end{bmatrix}, \quad (22)$$

can be expressed as $\Psi = \mathbf{A}\Psi\mathbf{A}^\top + \Xi$ where $\Xi = \mathbb{E}[\mathbf{E}_n \mathbf{E}_n^\top]$ is the $2p \times 2p$ covariance of $\mathbf{E}_n$. This last equation is a discrete-time Lyapunov equation, which can be solved for Ψ appearing in Equation 22 to yield the autocovariance matrices $\Gamma_0, \dots, \Gamma_{p-1}$ (Faes et al., 2015b; Mijatovic et al., 2025). Note that $\Gamma_0 \equiv \Sigma_Y$. Finally, the autocovariance can be calculated recursively for any lag $k \geq p$ by repeatedly applying YW equations in Equations 16a,b up to the desired lag q , starting from the parameters of the ARX representation in Equations 16a,b of the observed Gaussian vector process $\mathbf{Y}$.

Then, once the covariance matrices $\Gamma_0, \dots, \Gamma_q$ are derived, they need to be pruned to retain only those elements which relate to the desired restricted model. Here, we review the procedure which allows to get the restricted AR parameters $\{b_{Y_2 Y_2, k}, \lambda_{W_{Y_2}}^2\}$. The AR model in Equation 17 can be written in compact form as $Y_{2,n} = \mathbf{B}_{Y_2 Y_2} Y_{2,n}^q + W_{Y_2,n}$, where $\mathbf{B}_{Y_2 Y_2} = [b_{Y_2 Y_2, 1}, \dots, b_{Y_2 Y_2, q}]$ is the vector collecting all coefficients up to lag q . From this representation, taking the expectation $\mathbb{E}[Y_{2,n} Y_{2,n}^q]$ and solving for $\mathbf{B}_{Y_2 Y_2}$ yields:

$$\mathbf{B}_{Y_2 Y_2} = \Sigma_{Y_{2,n} Y_{2,n}^q} \cdot \Sigma_{Y_{2,n}^q}^{-1}, \quad (23)$$

where in Equation 23 the covariance $\Sigma_{Y_{2,n}^q} = \mathbb{E}[Y_{2,n}^q Y_{2,n}^{q\top}]$ is the $q \times q$ autocovariance matrix of $Y_{2,n}^q$, while $\Sigma_{Y_{2,n} Y_{2,n}^q} = \mathbb{E}[Y_{2,n} Y_{2,n}^{q\top}]$ is the

$1 \times q$ cross-covariance matrix of $Y_{2,n}$ and $Y_{2,n}^q$. The matrices $\Sigma_{Y_{2,n}^q}$ and $\Sigma_{Y_{2,n}^q, Y_{2,n}^q}$ are extracted from Γ_k . Then, the variance of the AR residuals $\lambda_{W_{Y_2}}^2$ in Equation 17 is computed as (Barnett et al., 2009):

$$\lambda_{W_{Y_2}}^2 = \sigma_{Y_2}^2 - \Sigma_{Y_{2,n}, Y_{2,n}^q} \cdot \Sigma_{Y_{2,n}^q}^{-1} \cdot \Sigma_{Y_{2,n}, Y_{2,n}^q}^T. \quad (24)$$

To summarize, the above-described procedure is based first on computing the autocovariance sequence of the bivariate process $\mathbf{Y}$ from its parameters ($\mathbf{A}_{Y,l}$, with $l = 1, \dots, p$, and Σ_{U_Y}), which are previously identified through the vector OLS approach, and then on rearranging the elements of the autocovariance matrices for building the auto- and cross-covariances to be used in the computation, according to Equation 23, 24, of the AR parameters $\{b_{Y_2 Y_2, k}, \lambda_{W_{Y_2}}^2\}$ appearing in Equation 17. The identification of the restricted model in Equation 17 can be represented in the frequency domain to study spectral patterns of causality. The same procedure applies to the AR parameters $\{b_{Y_1 Y_1, k}, \lambda_{W_{Y_1}}^2\}$ if Y_1 is assumed as the target process.

Contrary to the closed-form SS modeling approach presented in the previous subsection, the procedure described here is approximate because it retains the AR structure which has infinite order for the restricted model. The parameter determining the accuracy of the procedure is the number of lags used to truncate the past history of the process: considering the past up to lag q corresponds to calculating the autocovariance of the process in Equations 16a,b up to the matrix Γ_q . Given that the autocovariance of a stable vector AR process decays exponentially with the lag, with a rate of decay depending on the modulus of the largest eigenvalue of $\mathbf{A}$, $\rho(\mathbf{A})$, it has been suggested to compute the autocovariance up to a lag q such that $\rho(\mathbf{A})$ is smaller than a predefined numerical tolerance (Barnett and Seth, 2014). It has been found that computation of very long autocovariance sequences is not necessary for the purpose of evaluating information dynamics, because all measures stabilize to constant values already for small lags (typically $q = 10$) even for reasonably high values of $\rho(\mathbf{A})$ (Faes et al., 2013b; 2015a; b, 2016b; Mijatovic et al., 2025). In fact, the procedure described above yields results similar to the method of SS models with q sufficiently large (Antonucci et al., 2021a).

3.3 Time domain measures for linear processes

The parametric representation of the analysed bivariate process allows to derive measures of coupling and causality which are widely used for the description of symmetric and directed interactions in the time domain (Geweke, 1982; Bressler and Seth, 2011; Porta and Faes, 2015). These measures are obtained from the variances of the two analysed processes and of the prediction errors of the full and restricted models of Equations 16a,b, Equations 17. Specifically, a time-domain measure of the TD between the two random processes Y_1 and Y_2 is obtained comparing the generalized covariance of the vector noise $\mathbf{U}_{Y,n}$ of the full bivariate model of Equations 16a,b with the individual noise variances W_{Y_1} and W_{Y_2} of two separate restricted models in the form of Equation 17 as follows (Geweke, 1982; Pernice et al., 2022b)

$$F_{Y_1; Y_2} = \log \frac{\lambda_{W_{Y_1}}^2 \lambda_{W_{Y_2}}^2}{|\Sigma_{U_Y}|}. \quad (25)$$

The TD measure (Equation 25) is zero in the absence of any interaction between the two processes, resulting from a full equivalence between the bivariate AR description of $\{Y_1, Y_2\}$ and the two individual AR descriptions of Y_1 and Y_2 ; if this occurs, the two restricted models have the same error variance of the full model ($\lambda_{W_{Y_1}}^2 = \sigma_{U_{Y_1}}^2$, $\lambda_{W_{Y_2}}^2 = \sigma_{U_{Y_2}}^2$) and the two errors of the full model are uncorrelated ($\sigma_{U_{Y_1} Y_2}^2 = 0$). On the contrary, $F_{Y_1; Y_2}$ grows with the strength of the causal influence from Y_1 to Y_2 (yielding $\sigma_{U_{Y_2}}^2 < \lambda_{W_{Y_2}}^2$), with the strength of the causal influence from Y_2 to Y_1 (yielding $\sigma_{U_{Y_1}}^2 < \lambda_{W_{Y_1}}^2$), or with the strength of the instantaneous effects (yielding $\sigma_{U_{Y_1} Y_2}^2 = \sigma_{U_{Y_2} Y_1}^2 > 0$). These three components of the coupling can be isolated by decomposing the TD measure as (Geweke, 1982):

$$F_{Y_1; Y_2} = F_{Y_1 \rightarrow Y_2} + F_{Y_2 \rightarrow Y_1} + F_{Y_1 \cdot Y_2}. \quad (26)$$

Equation 26 evidences the two measures of Granger causality (GC) from Y_1 to Y_2 and from Y_2 to Y_1 , $F_{Y_1 \rightarrow Y_2}$ and $F_{Y_2 \rightarrow Y_1}$, and the measure of instantaneous causality (IC), $F_{Y_1 \cdot Y_2}$. The GC measures quantify the directed interactions between the processes according to the principle of Granger causality (Granger, 1969), whereby the improvement in predictability of the present state of a process brought by the knowledge of the past states of the other process can be quantified comparing the prediction error variances of the full and the restricted AR models of Equations 16a,b, Equations 17; specifically, GC from Y_1 to Y_2 is quantified as:

$$F_{Y_1 \rightarrow Y_2} = \log \frac{\lambda_{W_{Y_2}}^2}{\sigma_{U_{Y_2}}^2}, \quad (27)$$

and GC from Y_2 to Y_1 is computed analogously exchanging the role of the two processes. Finally, the IC measure quantifies the zero-lag correlation between the two processes, which is reflected by the off-diagonal elements of the covariance $\mathbf{U}_{Y,n}$ of the errors of the bivariate AR model in Equation 16:

$$F_{Y_1 \cdot Y_2} = \log \frac{\sigma_{U_{Y_1}}^2 \sigma_{U_{Y_2}}^2}{|\Sigma_{U_Y}|}, \quad (28)$$

in the case of strict causality ($\sigma_{U_{Y_1} Y_2}^2 = \sigma_{U_{Y_2} Y_1}^2 = 0$) the IC measure of Equation 28 vanishes, denoting the absence of instantaneous effects.

3.4 Spectral measures for linear processes

In this section we present the tools whereby the linear parametric description of time series is widely exploited to describe coupling and causality in the frequency domain in a range of applicative fields, primarily including network neuroscience and network physiology (Schulz et al., 2013; Porta and Faes, 2015; Santiago-Fuentes et al., 2022; Sorelli et al., 2022; Jahani et al., 2025). To achieve a parametric estimation of the PSD matrix of the process (Equation 3), the ARX model in Equation 16 can be represented in the Z-domain through its Z-transforms yielding $\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{U}_Y(z)$, where $\mathbf{H}(z) = [\mathbf{I} - \sum_{k=1}^p \mathbf{A}_{Y,k} z^{-k}]^{-1} = \bar{\mathbf{A}}_Y(z)^{-1}$ is the 2×2 transfer matrix (TF) matrix modelling the relationships between the input $\mathbf{U}_Y(z)$ and the output $\mathbf{Y}(z)$. Computing $\mathbf{H}(z)$ on the unit circle in the complex plane

($z = e^{i2\pi f}$), it is possible to derive the PSD of the analysed stationary random process $\mathbf{Y}$ exploiting spectral factorization:

$$\mathbf{P}_Y(\bar{f}) = \mathbf{H}(\bar{f})\mathbf{\Sigma}_{U_Y}\mathbf{H}^*(\bar{f}). \quad (29)$$

Importantly, spectral factorization is the core of the causal analysis of dynamic processes studied in the frequency domain and, while it is ubiquitously performed using AR modeling, could actually be obtained regardless of it (Baccalá and Sameshima, 2022). As specified in Section 1.2, the PSD matrix $\mathbf{P}_Y(\bar{f})$ contains information related to the spectral properties of the two processes, i.e., to their own dynamics, through the 2 diagonal elements $P_{Y_i}(\bar{f})$, $i = 1, 2$, and to the causal interactions between Y_1 and Y_2 , through the off-diagonal elements $P_{Y_iY_j}(\bar{f})$, $i, j \in \{1, 2\}, i \neq j$.

The bivariate interactions between the processes Y_1, Y_2 can be assessed by measures of spectral coupling and causality in the frequency domain, which can be directly derived from different combinations of the elements of the 2×2 PSD matrix $\mathbf{P}_Y(\bar{f})$, of the 2×2 TF matrix $\mathbf{H}(\bar{f})$ and of the innovations of the full ARX and the restricted AR models. Under the assumption that the input white noises are uncorrelated at lag zero leading to diagonality of $\mathbf{\Sigma}_{U_Y}$ (i.e., the model is strictly causal) (Ding et al., 2006; Faes et al., 2012a), from Equation 29 the ij^{th} elements of $\mathbf{P}_Y(\bar{f})$ can be factorized into:

$$P_{Y_iY_j}(\bar{f}) = \sum_{q=1}^2 \sigma_{U_{Y_q}}^2 H_{Y_iY_q}(\bar{f}) H_{Y_jY_q}^*(\bar{f}), \quad (30)$$

where $\sigma_{U_{Y_q}}^2$, $q = 1, 2$, are the diagonal elements of $\mathbf{\Sigma}_{U_Y}$. This factorization allows to decompose the frequency domain measures of coupling and causality into terms eliciting the directional information from one process to another. Indeed, exploiting Equation 30 the Coh between the two processes Y_1 and Y_2 Equation 5 is decomposed as:

$$\Gamma_{Y_1Y_2}(\bar{f}) = \sum_{q=1}^2 \gamma_{Y_1Y_q}(\bar{f}) \gamma_{Y_2Y_q}^*(\bar{f}), \quad (31)$$

where $\gamma_{Y_iY_j}(\bar{f})$ is the so-called directed coherence (DC) (Saito and Harashima, 1981; Baccalá et al., 1998) quantifying frequency-specific causal influences from Y_j to Y_i ($i, j = 1, 2$; the measure quantifies internal influences within a process when $i = j$). In particular, the squared modulus of the DC from the driver process Y_1 to the target process Y_2 appearing in Equation 31 denoted also causal coherence (Porta et al., 2002a) and defined as:

$$|\gamma_{Y_2Y_1}(\bar{f})|^2 = \frac{\sigma_{U_{Y_1}}^2 |H_{Y_2Y_1}(\bar{f})|^2}{\sigma_{U_{Y_1}}^2 |H_{Y_2Y_1}(\bar{f})|^2 + \sigma_{U_{Y_2}}^2 |H_{Y_2Y_2}(\bar{f})|^2}, \quad (32)$$

can be interpreted as a measure of the influence of Y_1 onto Y_2 , being 0 in the absence of directional coupling from Y_1 to Y_2 at the frequency f , and 1 in the presence of full coupling. Importantly, if the bivariate process is strictly causal the squared DC $|\gamma_{Y_2Y_1}(\bar{f})|^2$ reflects the coupling strength from Y_1 to Y_2 intended as the normalized proportion of $P_{Y_2}(\bar{f})$ which is causally due to Y_1 . This interpretation arises from the observation that, for strictly causal processes, Equation 30 written for the target process becomes $P_{Y_2}(\bar{f}) = P_{Y_2|Y_1}(\bar{f}) + P_{Y_2|Y_2}(\bar{f})$, where $P_{Y_2|Y_1}(\bar{f}) = \sigma_{U_{Y_1}}^2 |H_{Y_2Y_1}(\bar{f})|^2$ is the numerator of the DC (Equation 32), the

denominator being the whole spectrum $P_{Y_2}(\bar{f})$. Therefore $P_{Y_2|Y_1}(\bar{f}) = |\gamma_{Y_2Y_1}(\bar{f})|^2 P_{Y_2}(\bar{f})$ is the part of $P_{Y_2}(\bar{f})$ due to Y_1 , which is usually referred to as the *causal* part of $P_{Y_2}(\bar{f})$; $P_{Y_2|Y_2}(\bar{f}) = |\gamma_{Y_2Y_2}(\bar{f})|^2 P_{Y_2}(\bar{f})$ measures the part of $P_{Y_2}(\bar{f})$ due to none of the other processes, but to the process Y_2 itself, which has been referred to as the *isolated* part of the target spectrum (Sparacino et al., 2024a). It is worth stressing that the DC (Equation 32) is a non-negative measure of causal coupling, but its meaningful physical interpretation as the amount of signal power transferred from the driver to the target process holds only under strict causality, because the denominator of (Equation 32) is not the PSD of the target process if instantaneous effects are present. To overcome this interpretational gap, extended versions of the causal coherence have been proposed in the literature (Faes and Nollo, 2010; 2013) which make use of AR models with instantaneous effects incorporated into the regression parameters.

Another line of research introduced independently frequency-domain measures of symmetric and directed interaction between two stationary jointly Gaussian processes based on the spectral expansion of the time-domain measures reviewed in Section 2.3) (Geweke, 1982; Ding et al., 2006; Chicharro, 2011). Given the spectral density matrix of the bivariate process, the frequency-specific measure of TD between Y_1 and Y_2 , $f_{Y_1Y_2}(\bar{f})$, was introduced as in Equation 6; the measure is null at the frequencies for which the two processes are uncoupled, and does not have an upper bound; it is indeed related to the magnitude squared coherence through Equation 7. Moreover, the measure satisfies the so-called *spectral integration property*, which relates the spectral measure (Equation 6) with the time-domain coupling measure (Equation 25) as follows (Gelfand and Yaglom, 1959):

$$F_{Y_1Y_2} = 2 \int_0^{\frac{1}{2}} f_{Y_1Y_2}(\bar{f}) d\bar{f}. \quad (33)$$

Moreover, exploiting the bivariate AR model and spectral factorization, Geweke formulated the so-called *linear feedback* measure defined as (Geweke, 1982)

$$f_{Y_1 \rightarrow Y_2}(\bar{f}) = \log \frac{P_{Y_2}(\bar{f})}{\sigma_{U_{Y_2}}^2 |H_{Y_2Y_2}(\bar{f})|^2}, \quad (34)$$

which is non-negative and linked to the corresponding time domain GC measure (Equation 27) by the spectral integration property according to a relation similar to that of Equation 33:

$$F_{Y_1 \rightarrow Y_2} = 2 \int_0^{\frac{1}{2}} f_{Y_1 \rightarrow Y_2}(\bar{f}) d\bar{f}. \quad (35)$$

The Geweke measure of spectral causality is an upper unbounded measure of GC which, if the bivariate process is strictly causal, can be related to the normalized measure of causal coherence. In fact, combining Equations 32, 34 one can easily show that the DC and the spectral GC are linked by the relation $f_{Y_1 \rightarrow Y_2}(\bar{f}) = -\log(1 - |\gamma_{Y_2Y_1}(\bar{f})|^2)$ (Geweke, 1982; Chicharro, 2011; Faes et al., 2012a; Pernice et al., 2022b; Sparacino et al., 2024a). We note that, while the TD measure (Equation 6) is always non-negative, the two causal measures $f_{Y_i \rightarrow Y_j}(\bar{f})$, $i, j = 1, 2$, can take negative values at some frequencies if the

process $\mathbf{Y}$ is not strictly causal (i.e., if the innovation covariance Σ_{U_Y} is not diagonal) (Pernice et al., 2022b).

Finally, the spectral measure of *instantaneous causality* (IC) was chosen *ad hoc* as (Geweke, 1982)

$$f_{Y_1:Y_2}(\bar{f}) = \log \frac{\sigma_{U_{Y_1}}^2 |H_{Y_1 Y_1}(\bar{f})|^2 \sigma_{U_{Y_2}}^2 |H_{Y_2 Y_2}(\bar{f})|^2}{|\mathbf{P}_Y(\bar{f})|} \quad (36)$$

to satisfy the *Geweke decomposition of the total dependence* in the frequency domain, i.e.,

$$f_{Y_1:Y_2}(\bar{f}) = f_{Y_1 \rightarrow Y_2}(\bar{f}) + f_{Y_2 \rightarrow Y_1}(\bar{f}) + f_{Y_1, Y_2}(\bar{f}), \quad (37)$$

in such a way to be linked to the corresponding time domain IC measure (28) by the spectral integration property similarly to the relations established in Equations 33, 35 i.e.,

$$F_{Y_1:Y_2} = 2 \int_0^{\frac{1}{2}} f_{Y_1:Y_2}(\bar{f}) \, d\bar{f}. \quad (38)$$

The spectral measure defined in Equation 36 to satisfy Equation 37 and to be related to the time-domain measure of instantaneous causality by Equation 38 does not fulfil all the requirements of Geweke, different from what occurs in the time domain. Indeed, it may be negative for some frequencies and has no clear physical meaning (Geweke, 1982). In the absence of instantaneous causality, $F_{Y_1:Y_2} = 0$ because the prediction error covariance Σ_{U_Y} is diagonal; however, $f_{Y_1:Y_2}(\bar{f})$ is generally found to be non-zero for some frequencies. Since the integral of $f_{Y_1:Y_2}(\bar{f})$ has to be null when $F_{Y_1:Y_2} = 0$, not being zero for all frequencies, this implies the violation of non-negativity (Pernice et al., 2022b).

The issue of instantaneous causality in the computation of frequency domain measures of GC is a relevant one, and several efforts have been made to interpret instantaneous GC and to provide corrected measures (Faes and Nollo, 2010; Faes et al., 2013a; Nuzzi et al., 2021; Baccalá and Sameshima, 2021a). For instance, methods that identify a preferred direction for the instantaneous effects and then incorporate them together with lagged effects into directional measures of extended causality, were proposed by Faes and Nollo (2010) and Faes et al. (2013a): the methods exploit the Cholesky decomposition of the AR parameters and set the direction of zero-lag effects based on *a-priori* assumptions subjectively relying on physical knowledge (Faes and Nollo, 2010) or objectively relying on non-gaussianity of the AR residuals (Faes et al., 2013a). These methods were successfully applied to electroencephalographic rhythms (Faes et al., 2013a), as well as to cardiovascular and cerebrovascular oscillations (Pernice et al., 2022a). Measures which include instantaneous causality in the modelling approach are generally preferred because they enforce that zero-lag effects are ascribed to one of the two causal directions and therefore become zero both in time and frequency domain (Pernice et al., 2022a), essentially solving the issue about interpretability. Nevertheless, the assumptions about prior physiological knowledge or non-gaussianity of the residuals are not always satisfied, and therefore several alternatives based on keeping instantaneous effects as undirected but including them in extended spectral causality measures have been proposed to face the problem. As an example, Baccalá and Sameshima (2021a) discussed the theoretical interpretation of instantaneous GC within a spectral

framework, and decomposed frequency-domain measures of causality, namely, directed transfer function and partial directed coherence, into lagged and instantaneous connectivity terms without the need of including the zero lag in AR models. On the other hand, Nuzzi et al. (2021) introduced an alternative frequency-domain decomposition of GC by obtaining a novel index of undirected instantaneous causality and a novel measure of GC including instantaneous effects, with the purpose to mitigate the confounding effect of zero-lag coupling. The issue of how it is best to treat instantaneous effects in the analysis of physiological interactions, e.g., cardiovascular interactions, where zero-lag interdependencies are expected to contribute significantly to the baroreflex mechanism (see, e.g., Faes et al. (2013a)), and of cardiorespiratory interactions, where the information transfer from respiration to heart rate variability is expected to be within-beat (Faes et al., 2012b), still remains an active area of research.

3.5 Connection between information-theoretic and spectral formulations

When formulated for jointly stationary Gaussian processes, the Geweke spectral and time domain measures of coupling and causality reviewed in the previous subsections have a straightforward information-theoretic interpretation (Geweke, 1982; Pernice et al., 2022b; Sparacino et al., 2024a). Indeed, the spectral measures of the bivariate interactions between two processes, defined by the spectral TD (Equation 6), the spectral GC (Equation 34) and the spectral IC (Equation 36), are closely related to the information-theoretic measures defined in Equations 11, 13, 15 by means of the spectral integration property (Geweke, 1982; Chicharro, 2011):

$$I_{Y_1:Y_2} = \frac{F_{Y_1:Y_2}}{2} = \int_0^{\frac{1}{2}} f_{Y_1:Y_2}(\bar{f}) \, d\bar{f}, \quad (39a)$$

$$T_{Y_i \rightarrow Y_j} = \frac{F_{Y_i \rightarrow Y_j}}{2} = \int_0^{\frac{1}{2}} f_{Y_i \rightarrow Y_j}(\bar{f}) \, d\bar{f}; i, j = 1, 2, i \neq j \quad (39b)$$

$$I_{Y_1, Y_2} = \frac{F_{Y_1, Y_2}}{2} = \int_0^{\frac{1}{2}} f_{Y_1, Y_2}(\bar{f}) \, d\bar{f}. \quad (39c)$$

The spectral integration property is very important not only to connect the information-theoretic and spectral formulations of the interaction measures, but also to allow quantification of these measures with reference to specific oscillatory components contained within spectral bands of interest. Examples of band-specific integration of the spectral interaction measures to obtain values related to peculiar oscillations of a group of random processes are reported in the next sections for real-world systems.

4 Implementation through model-free approaches

The practical implementation of the information-theoretic measures of symmetric and directional coupling through model-free approaches assumes that the measures are estimated directly from data without assuming a parametric model for the underlying probability distribution. These approaches are especially useful in

high-dimensional, non-Gaussian, or complex distributions where classical parametric methods fail. This section presents three widely used model-free approaches for estimating entropy-based measures of coupling and causality, i.e., the k -nearest neighbour (Section 3.1), the binning (Section 3.2) and the permutation (Section 3.3) approaches. Other techniques, including kernel and slope estimators, are also discussed in the literature; however, they are not considered here for brevity reasons, and the reader is referred to Barà et al. (2024) for further details.

4.1 Nearest-neighbour estimator

The k -nearest neighbour (KNN) method is one of the most widely used entropy estimators (Ince et al., 2013; Wibral et al., 2013; Porta et al., 2016; Trujillo, 2019; Xiong et al., 2017; Azami et al., 2023), due to its ability to dynamically adjust resolution by adapting the distance scale to the underlying probability distribution (Victor, 2002), and its potential for bias correction through distance projection (Kraskov et al., 2004). This method builds upon the findings of Kozachenko and Leonenko (1987), who demonstrated that the average Shannon information of a d -dimensional random variable $\mathbf{W}$ can be approximated using the nearest neighbour distances among N observations $\{w_1, w_2, \dots, w_N\}$. Specifically, the expectation of the log-probability of a sample point w_n is estimated as:

$$-\mathbb{E}[\log p(w_n)] = \psi(N) - \psi(k) + d \langle \log \varepsilon_n \rangle, \quad (40)$$

where $\psi(\cdot)$ denotes the digamma function, and ε_n is twice the distance (measured under the maximum norm) between w_n and its k^{th} nearest neighbour. The notation $\langle \cdot \rangle$ indicates an average over the entire set of N samples. Then, for instance, based on Equation 40 the KNN estimation of the entropy of the present state of Y_1 can be expressed as:

$$H(Y_{1,n}) = \psi(N) - \psi(k) + \langle \log \varepsilon_n \rangle. \quad (41)$$

Besides the entropy of a scalar variable formulated as in Equation 41, the nearest neighbour estimator can be used to compute all of the entropy terms that compose a given coupling measure. However, as shown in Equation 13 for the transfer entropy, the coupling and causality measures are expressed as combinations of entropy terms computed in spaces of different dimensions. Using the same nearest-neighbour search across these spaces leads to inconsistent distance scales, introducing estimation biases not canceled by entropy differences. To mitigate computational bias, the process begins by identifying the nearest neighbours within the full high-dimensional space, and then examining how these neighbours distribute across various lower-dimensional projections (Kraskov et al., 2004). Following this approach, given the bivariate system $\mathbf{Y} = \{Y_1, Y_2\}$, the joint entropy over the combined space $[Y_{1,n}, Y_{2,n}, Y_{1,n}^q, Y_{2,n}^q]$, where q is the number of lags used to truncate the past of the processes, is first estimated as:

$$H(Y_{1,n}, Y_{2,n}, Y_{1,n}^q, Y_{2,n}^q) = \psi(N) - \psi(k) + 2(q+1) \langle \log \varepsilon_n \rangle, \quad (42)$$

where ε_n is twice the distance from $[y_{1,n}, y_{2,n}, y_{1,n}^q, y_{2,n}^q]$ to its k^{th} nearest neighbour. Then, the entropies in lower-dimensional

subspaces are computed using the same radius ε_n via range search. For the process Y_1 we obtain:

$$H(Y_{1,n}, Y_{1,n}^q) = \psi(N) - \langle \psi(N_{Y_{1,n}, Y_{1,n}^q} + 1) \rangle + (q+1) \langle \log \varepsilon_n \rangle, \quad (43a)$$

$$H(Y_{1,n}^q) = \psi(N) - \langle \psi(N_{Y_{1,n}^q} + 1) \rangle + q \langle \log \varepsilon_n \rangle, \quad (43b)$$

where $N_{Y_{1,n}, Y_{1,n}^q}$ and $N_{Y_{1,n}^q}$ count neighbours within a distance $\varepsilon_n/2$ from $[y_{1,n}, y_{1,n}^q]$ and $y_{1,n}^q$, respectively. Analogously, for the process Y_2 we obtain:

$$H(Y_{2,n}, Y_{2,n}^q) = \psi(N) - \langle \psi(N_{Y_{2,n}, Y_{2,n}^q} + 1) \rangle + (q+1) \langle \log \varepsilon_n \rangle, \quad (44a)$$

$$H(Y_{2,n}^q) = \psi(N) - \langle \psi(N_{Y_{2,n}^q} + 1) \rangle + q \langle \log \varepsilon_n \rangle, \quad (44b)$$

with $N_{Y_{2,n}, Y_{2,n}^q}$ and $N_{Y_{2,n}^q}$ defined in the same way. The joint entropy of the past states of both processes is estimated as:

$$H(Y_{1,n}, Y_{2,n}^q) = \psi(N) - \langle \psi(N_{Y_{1,n}, Y_{2,n}^q} + 1) \rangle + 2q \langle \log \varepsilon_n \rangle, \quad (45)$$

where $N_{Y_{1,n}, Y_{2,n}^q}$ counts neighbours within $\varepsilon_n/2$ from $[y_{1,n}^q, y_{2,n}^q]$. Finally, using the entropy expressions derived in Equations 42–45 and substituting them into Equation 13, the following estimator of TE is derived (Faes et al., 2015a; Pinto et al., 2025a):

$$T_{Y_1 \rightarrow Y_2} = \langle \psi(N_{Y_{2,n}, Y_{1,n}^q, Y_{2,n}^q} + 1) - \psi(N_{Y_{1,n}^q, Y_{2,n}^q} + 1) - \psi(N_{Y_{2,n}, Y_{2,n}^q} + 1) + \psi(N_{Y_{1,n}^q} + 1) \rangle. \quad (46)$$

From Equation 46, the IC measure and the MIR can be computed as:

$$I_{Y_1, Y_2} = \psi(k) + \langle \psi(N_{Y_{1,n}, Y_{2,n}^q} + 1) - \psi(N_{Y_{1,n}, Y_{1,n}^q, Y_{2,n}^q} + 1) - \psi(N_{Y_{2,n}, Y_{1,n}^q, Y_{2,n}^q} + 1) \rangle, \quad (47a)$$

$$I_{Y_1, Y_2} = \psi(k) + \langle \psi(N_{Y_{1,n}^q} + 1) + \psi(N_{Y_{2,n}^q} + 1) - \psi(N_{Y_{1,n}, Y_{1,n}^q} + 1) - \psi(N_{Y_{2,n}, Y_{2,n}^q} + 1) - \psi(N_{Y_{1,n}^q, Y_{2,n}^q} + 1) \rangle. \quad (47b)$$

The accuracy of entropy estimators can vary depending on both the data size and the chosen number of nearest neighbours (k). Specifically, smaller values of k yield more local estimates that capture subtle data structures but are more sensitive to noise. Conversely, larger values produce smoother, more stable estimates that may overlook fine-scale dynamics (Lombardi and Pant, 2016). Thus, selecting k requires balancing the trade-off between bias and variance, as thoroughly described in many previous works on the topic [see, e.g., Kugiumtzis (2013); Barà et al. (2023); Pinto et al. (2024); Barà et al. (2024); Pinto et al. (2025b)].

4.1.1 Embedding procedures

Finding embedding vectors that adequately approximate the infinite-dimensional past states of the processes is a critical step in estimating information-theoretical measures using model-free approaches. When working with time series of finite length, e.g., the 300 samples typically used for the analysis of short-term physiological time series (Shaffer and Ginsberg, 2017), the employment of high-dimensional vectors to provide a more complete description of past processes leads to the curse of dimensionality and unreliable estimates of entropy quantities (Faes and Porta, 2014). A selection technique widely used in this frame is the *uniform embedding* approach, which simply uses a fixed number of equally spaced variables. Nevertheless, this method may

overlook the most informative lags, potentially limiting the effectiveness of information-theoretic analyses in capturing relevant temporal dynamics (Kugiumtzis, 2013). An alternative approach was introduced to limit the size of the descriptive patterns and maximize their informational content about process dynamics, i.e., the *non-uniform embedding* approach introduced in (Faes et al., 2011; 2015a). In brief, a set of candidates including all possible states of the processes up to a maximum lag q is first considered. In the case of the TE estimation, with, e.g., Y_2 as the target process, a sequential approach is applied to fill progressively the embedding vector with the components taken from $[Y_{1,n}^q, Y_{2,n}^q]$ which maximize the information shared with the present state of the target process $Y_{2,n}$. Starting with an empty embedding vector, the additional information brought by each candidate above and beyond that provided by the previously selected variables is evaluated at each step, and the candidate bringing the maximum of such information is selected; the metric used to perform this maximization is the conditional MI. then, the selected candidate is retained and added to the embedding vector if the contribution that it brings to the target is statistically significant; significance is typically assessed over randomized datasets obtained shuffling randomly the samples of the target series, and setting a threshold on the conditional MI based on percentiles. This step is repeated until no remaining candidate contributes a statistically significant amount of information (Faes et al., 2011; 2015a).

4.2 Binning estimator

The binning estimation approach is based on performing uniform or non-uniform quantization of a continuous random variable and then estimating the entropy of the variable approximating probabilities with the frequency of visitation of the quantized states, or bins. Specifically, uniform quantization simplifies implementation by dividing the range of values into equal intervals (Faes and Porta, 2014), whereas non-uniform quantization better preserves the dynamic structure of data by adapting to its distribution (Darbellay and Vajda, 1999).

Let W be a generic continuous random variable defined over the interval $\mathcal{D}_W = [w_{\min}, w_{\max}]$. Quantization transforms W into a discrete random variable B_W that assumes values from a finite alphabet $\mathcal{A}_{B_W} = \{1, 2, \dots, b\}$, where b denotes the total number of quantization bins. In the case of uniform quantization, an observation $v \in \mathcal{D}_W$ is assigned to bin $b_w = i$ if it satisfies the condition $w_{\min} + (i-1)r \leq v < w_{\min} + ir$, where $r = (w_{\max} - w_{\min})/b$ represents the uniform bin width (Azami et al., 2023; Barà et al., 2023). Following the quantization process, the probability associated with each symbol in the discrete alphabet $\mathcal{A}_{B_W}$ is naturally approximated by its empirical frequency across a large number of observations, that is $\hat{p}(i) = \Pr\{b_w = i\} = N_i/N$, where N_i is the number of observations of the variable that fall into the bin. The entropy of the original variable W can then be estimated by computing the entropy of the quantized variable B_W by

$$H(W) = - \sum_{b_w \in \mathcal{A}_{B_W}} \hat{p}(b_w) \log \hat{p}(b_w). \quad (48)$$

The concepts outlined above extend naturally to multivariate cases, where quantization is performed independently on each

scalar element of the analysed vector variable. Specifically, if we consider a d -dimensional continuous random vector $\mathbf{W} = [W_1, \dots, W_d]$, each component can be discretized using b uniform bins; the resulting discrete vector variable $\mathcal{B}_W = [\mathcal{B}_{W_1}, \dots, \mathcal{B}_{W_d}]$ assumes values from a finite alphabet containing b^d unique combinations of bin indices.

In the context of dynamic processes, the two stochastic processes Y_1 and Y_2 are represented by temporally-correlated vector-valued random variables. These vectors are formed by concatenating the current values with previous samples from both processes, extending up to q time steps into the past (Barà et al., 2023). The computation of TE and IC, as defined in Equations 13, 15, involves four entropy terms; therefore, Equation 48 can be used to estimate these measures. For TE, the required entropies correspond to the following vector configurations: (i) the past of the target process alone (b^q); (ii) the past together with its current state (b^{q+1}); (iii) the joint past of both processes (b^{2q}); and (iv) the joint past combined with the present state of the target process (b^{2q+1}). For IC, the entropy terms involve: (i) the joint past of both processes (b^{2q}); (ii) the joint past with the present of one process (b^{2q+1}); and (iii) the joint past with the present of both processes (b^{2q+2}). Finally, the estimation of the MIR relies on the TE and IT estimators, according to the decomposition given in Equation 14.

A key issue in implementing discretization methods is the selection of the free parameters of the estimator, i.e., the memory length q used to capture the process history and, for the binning estimator, the number of quantization levels b used for coarse graining. These choices relate to estimating entropy in high-dimensional variables from limited data, tied to the curse of dimensionality (Runge et al., 2012). Empirically, parameter optimization aims to keep the alphabet size comparable to the series length N (Faes and Porta, 2014). Specifically, as regards the impact of the memory length, too small q implies that important dynamics or dependencies in the past are ignored, leading to underestimation of TE or MIR. On the other hand, too large q determines an increase of embedding dimension, such that the number of cells in the q -dimensional space grows exponentially ($\sim b^q$); in the case of small sample sizes this implies that the observations spread in the high-dimensional space, and most of the cells remain empty or sparsely populated, finally resulting in high variance of the estimates (the estimator becomes noisy and unreliable). The binning approach also requires a suitable choice of the number of quantization levels. Too small b leads to coarse quantization such that TE and MIR may be underestimated; conversely, too large b determines fine quantization, such that many bins have few or zero samples again resulting in high variance and possible overestimation due to noise [see, e.g., Barà et al. (2023), Barà et al. (2024)].

4.3 Permutation estimator

Permutation-based methods carry out symbolization by operating directly on discrete vector variables derived emphasizing the relative ordering of neighbouring sample amplitudes within each realization, rather than their specific absolute amplitude values (Bandt and Pompe, 2002). Given a general d -dimensional continuous random vector

$\mathbf{W} = [W_1, \dots, W_d]$, a discrete variable R_W is derived by applying a rank-ordering transformation. For any realization $\mathbf{w} = [w_1, \dots, w_d]$ of $\mathbf{W}$, the associated realization of R_W is denoted by $r_W = [r_{w_1}, \dots, r_{w_d}] \in \mathcal{A}_{R_W}$, where each $r_{w_i} \in \{1, \dots, d\}$ represents the rank position of w_i when the elements of $\mathbf{w}$ are arranged in ascending order. For instance, $r_{w_i} = 1$ if $w_i = \min(\mathbf{w})$, and $r_{w_i} = d$ if $w_i = \max(\mathbf{w})$. In the case of equal values, the element appearing later in the original sequence is assigned the smaller rank. After the discretization step, the probability associated with each symbol in the alphabet $\mathcal{A}_{R_W}$ is estimated in a straightforward manner by computing its relative frequency across a large number of observations. Consequently, the entropy of the original continuous variable $\mathbf{W}$ is approximated by computing the entropy of its discretized counterpart R_W , using a formulation similar to Equation 48 (Barà et al., 2023).

It is worth noting that the permutation strategy is favoured when compared with the binning approach for the estimation of entropy from a limited number of observations of the variable under analysis since for the first the continuous d -dimensional variable $\mathbf{W}$ takes values inside an alphabet with cardinality $|\mathcal{A}_{R_W}| = d!$, which is usually smaller than the cardinality of the alphabet obtained quantizing the variable with b bins, $|\mathcal{A}_{B_W}| = b^d$ (Azami et al., 2023; Sarà et al., 2023).

Analogously to the binning approach, estimation of the TE, presented in (Equation 13), requires that the relevant variables include either the past of the target process alone or combined with its present state, resulting in alphabet sizes of $q!$ and $(q+1)!$, respectively. Additionally, the joint past of both processes is considered, with or without the past of the target, leading to alphabet sizes of $(q!)^2$ or $q! \cdot (q+1)!$. IC, formulated in Equation 15, is estimated using combinations of the past and present values of Y_1 and Y_2 . The corresponding alphabet sizes are $(q!)^2$ for past-only combinations, $q! \cdot (q+1)!$ when one present state is included, and $((q+1)!)^2$ when both are included. Finally, the MIR is estimated based on its decomposition into TE and IC, as given in Equation 14.

For the permutation approach, the memory length q controls the temporal resolution and sensitivity to dependencies. With increasing q the number of possible ordinal patterns grows factorially ($\sim q!$); with limited data, many patterns may never occur thus leading to high variance and noisy TE/MIR estimates. Again, a bias-variance trade-off is required to get reliable estimates [see, e.g., Sarà et al. (2023), Sarà et al. (2024)].

5 Practical computation of bivariate interaction measures

The practical computation of the information-theoretic and spectral measures of coupling and causality from two time series of L samples measured from a physical system, $\mathbf{y}_i = \{y_i(1), \dots, y_i(L)\}$, where $i = 1, 2$ and L is the length of the time series, is based on considering the series as a finite length realization of the vector process $\mathbf{Y} = \{Y_1, Y_2\}$ that describes the evolution of the system over time.

The software and the codes relevant to this work are collected in the BIM Matlab toolbox and available for free download from <https://github.com/LauraSparacino>.

5.1 Non-parametric estimation of the PSD

Spectral measures of coupling can be computed directly from the PSD terms of the set $\mathbf{y} = \{\mathbf{y}_1, \mathbf{y}_2\}$ as evidenced in Section 1.2; time domain and information-theoretic measures of dynamic coupling can then be obtained exploiting the spectral integration property applied to the spectral estimates under the hypothesis of linearity (see Section 2.4 and Section 2.5).

A widely used *non-parametric* estimator of the PSD is the weighted covariance (WC) method (function `bim_WCspectra.m`). This estimator leverages the FT of the sample CCF of the observed data $\hat{R}_{\mathbf{y}_1\mathbf{y}_2}(k)$ (Blackman and Tukey, 1958) and is expressed as:

$$\hat{P}_{\mathbf{y}_1\mathbf{y}_2}(\bar{f}) = \sum_{k=-\tau}^{\tau} w(k) \hat{R}_{\mathbf{y}_1\mathbf{y}_2}(k) e^{-j2\pi\bar{f}k}, \quad (49)$$

where $\tau \leq L-1$ is the maximum lag considered for estimating the CCF. The window function $w(k)$ has width 2τ , satisfies $w(k) = 0$ for $|k| > \tau$, is normalized such that $0 \leq w(k) \leq w(0) = 1$, and is symmetric, i.e., $w(-k) = w(k)$ (Kay, 1988). To ensure non-negative spectral estimates as a result of Equation 49, it is common to use biased estimators for the CCF producing semi-definite sequences. A biased estimator is given by:

$$\hat{R}_{\mathbf{y}_1\mathbf{y}_2}(k) = \frac{1}{L} \sum_{n=0}^{L-1-k} \mathbf{y}_{1,n}^* \mathbf{y}_{2,n-k}, \quad (50)$$

where the asterisk denotes the complex conjugate. The definition in Equation 50 applies for $k = 0, \dots, L-1$, while for negative lags $k = -(L-1), \dots, -1$, the cross-correlation is defined by symmetry as $\hat{R}_{\mathbf{y}_1\mathbf{y}_2}(k) = \hat{R}_{\mathbf{y}_1\mathbf{y}_2}^*(-k)$. The same reasoning applies to the computation of the autocorrelation function $\hat{R}_{\mathbf{y}_i}(k)$ to get the autospectrum $\hat{P}_{\mathbf{y}_i}(k)$, with $i = 1, 2$. Window selection is usually performed by providing mathematical formulations for the window function which allow to control the spectral leakage introduced by the profile of the window itself (Pinna et al., 1996). Following this rationale, the Parzen window can be suitably selected, since it shows a significantly lower side-lobe level compared to Hanning and Hamming windows; furthermore, it is non-negative for all frequencies, and produces non-negative spectral estimates (Priestley, 1981). For the Parzen window, the relationship between the bandwidth (B_w) of the spectral window and the lag τ at which correlation estimates are truncated is $B_w = 1.273 f_s / \tau$ (Pinna et al., 1996). To resolve the corresponding peaks in the spectrum, the window bandwidth can be set equal to 25 Hz, which brings to $\tau \approx 0.05$ using $f_s = 1$ Hz.

5.2 Linear autoregressive models

The linear ARX equation in Equations 16a,b is seen as a model of how the observed data have been generated, and an identification procedure (function `bim_idARX.m`) is applied after model order selection (function `bim_mos_idARX.m`) to provide estimates of the coefficients and innovation variances to be used for computing the coupling and causality measures in the information-theoretic and spectral domains. Then, computation of the information-theoretic measures of MIR, TE and IT amounts to identify

restricted linear models through methods which extract the parameters of the restricted model from those of the full model (Section 2.2), i.e., based on (i) the resolution of the YW equations (function `bim_MIRdec_lin_YW.m`) or, equivalently, on (i) SS models (function `bim_MIRdec_lin_SS.m`). Side functions of (i) are `bim_LinReg.m`, which performs a linear regression of the present state of input target processes from the past states of input driver processes, and `bim_Yule.m`, which provides solution of the YW equations for a ARX process using discrete time Lyapunov equation; side functions of (ii) are `bim_SSmodel.m`, which computes the parameters of a SS model from those of a ARX model, and `bim_submodel.m`, which derives a submodel of a SS model. On the other hand, information-theoretic measures of coupling and causality can be obtained directly exploiting the spectral integration property (Section 2.4 and Section 2.5). If this is the case, *spectral* measures of TD, GC and IC are computed by first estimating the parametric PSD matrix of the set $\{y_1, y_2\}$ (function `bim_VARSpectra.m`). Then, the function `bim_fGC_lin.m` computes the spectral measures, accepting as inputs the PSD of the process, the transfer function of the full model and the innovation variances. The information-theoretic counterparts are finally obtained as the integrals of the corresponding spectral functions up to a factor 2.

5.3 Model-free approaches

Model-free estimation of the MIR, TE and IT measures can be performed by exploiting different approaches for the computation of entropy measures (Xiong et al., 2017; Azami et al., 2023). Regarding the approaches reviewed in this work, the KNN, binning and permutation estimators are implemented through the functions `bim_MIRdec_KNN.m`, `bim_MIRdec_bin.m` and `bim_MIRdec_perm.m`, respectively, taking as inputs the set $\{y_1^T, y_2^T\}$ in the form of a $L \times 2$ matrix, the number of past samples in the embedding vectors, and the vector of embedding delays τ . Further, the number of neighbours and quantization bins must be specified for the KNN and binning estimators, respectively. Side functions are `bim_SetLag.m`, which returns the list of candidates taking as input the vectors of embedding dimensions and delays (uniform embedding), `bim_quantization.m`, which quantizes the input series with a given number of quantization levels (for binning estimator), and `bim_H.m`, which computes entropy of a discrete multidimensional variable. The example `bim_thsim_modelfree.m` replicates physiological interactions and shows the behaviour of the three estimators at varying hyperparameters such as memory length, quantization levels, or the number of nearest neighbors.

5.4 Assessment of statistical significance

This section presents the use of *surrogate data analysis* to statistically validate the proposed measures of coupling and causality in the information-theoretic and spectral domains. Validation is performed at the level of individual realizations of the observed random processes $\{Y_1, Y_2\}$, obtained in the form of the set of time series y_i , $i = 1, 2$.

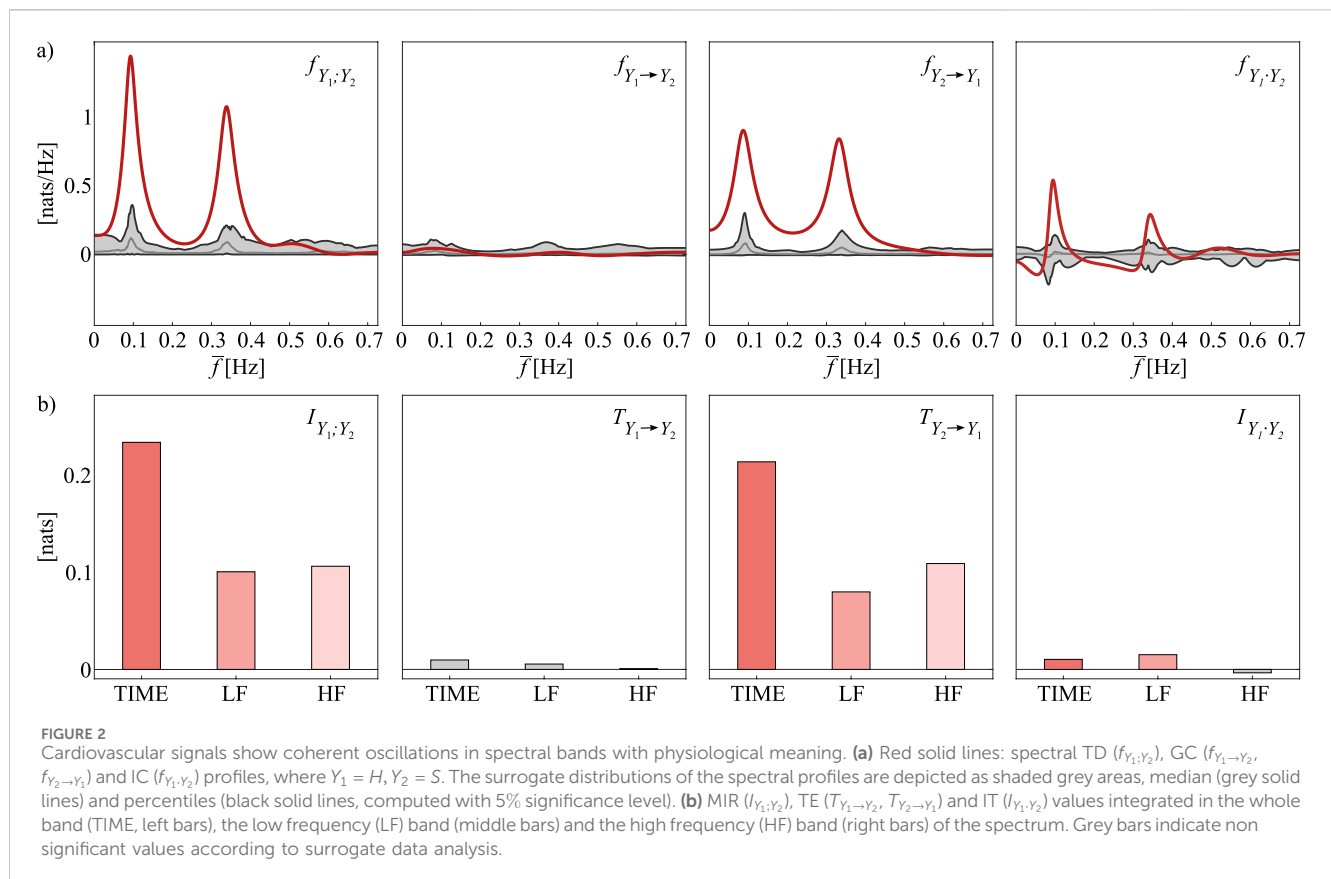
The method of surrogate data (Theiler et al., 1992; Zhang, 2023) is employed to set a significance level for the coupling and causality measures. Specifically, to assess the significance of conditional mutual information measures (i.e., $T_{Y_i \rightarrow Y_j}$ and I_{Y_i, Y_j} , $i, j = 1, 2, i \neq j$ and/or their Gaussian logarithmic counterparts - see Section 2.5), as well as of the MIR/TD, it is sufficient to destroy the coupling of the two series, while it is preferable to maintain the statistical properties of the individual series. Therefore, *random time shift bivariate surrogates* are generated according to the null hypothesis of independent random processes, shifting the samples of the time series Y_1 over time (while wrapping the extra values around the beginning of the series) and leaving the other series unchanged; the shift is chosen randomly, imposing a minimum shift of τ_{min} lags (function `bim_surrtimeshift.m`). For each pair of original time series, n_s pairs of surrogate time series are generated to obtain the set y_i^s ($i = 1, 2; s = 1, \dots, n_s$). The considered measure is then computed both on the original series and on each surrogate pair $\{y_1^s, y_2^s\}$, yielding the surrogate distribution from which the significance threshold is derived taking the $100(1 - \alpha)^{th}$ percentile. Then, each conditional mutual information measure is deemed as statistically significant if its value computed on the original series is higher than the significance threshold. The same procedure applies to both information-theoretic and spectral measures, the latter obtained by integrating the spectral profiles within specific frequency bands. The reader is referred to (Pinto et al., 2024) for methodologies and codes relevant to the assessment of coupling and nonlinearity in bivariate time series.

6 Exemplary applications to real-world time series

6.1 Cardiovascular interactions

In this section, we analyse physiological time series collected to study the effect of postural stress on cardiovascular variability (Faes et al., 2013c; Bari et al., 2016). One representative subject was selected for the following analyses, chosen from a dataset comprising healthy controls enrolled at the Neurology Division of Sacro Cuore Hospital, Negrar, Italy. Electrocardiogram (ECG) was acquired synchronously with arterial pressure (AP) measured at the level of middle finger through a photoplethysmographic device (Finapres Medical Systems, Ohmenda, Netherlands) at a sampling rate of 1 kHz. From the raw signals, stationary time series of heart period (H) and systolic AP (S) were measured as detailed in (Faes et al., 2013c; Bari et al., 2016) during the supine resting state condition, and regarded as realizations of the H and S discrete-time processes, in turn assumed as uniformly sampled with a sampling frequency equal to the inverse of the mean heart period. The series, each of length equal to 251 beats, were preprocessed reducing the slow trends with an AR high-pass filter (zero phase; cut-off frequency 0.0156 Hz), removing the mean value and normalizing to unit variance.

Information and spectral measures of coupling (MIR/TD) and causality (TE/GC, IT/IC) were computed from the parameters of an ARX model (least squares estimation, model order set according to the AIC - maximum scanned model order: 8; selected model order:



6), restricted through resolution of the YW equations and represented in the frequency domain to get a parametric estimation of the PSD matrix. The mathematical derivations are presented in Section 2, while the practical computation is detailed in Section 4. The spectral profiles were integrated within two frequency bands of physiological interest, i.e., the low frequency (LF, $f \in [0.04 - 0.15]$ Hz) and the high frequency (HF, $f \in [0.15 - 0.4]$ Hz) band, as well as over the whole frequency range ($f \in [0 - f_s/2]$ Hz) to get the corresponding time domain values (see Section 2.4 and Section 2.5). In heart rate variability analysis, the LF band reflects both sympathetic and parasympathetic activity, while the HF band indicates parasympathetic (vagal) influence, often linked to respiration (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). Surrogate data analysis was applied as in Section 4.1 to assess the statistical significance of the computed measures, with a minimum shift of $\tau_{min} = 20$ lags, $N_s = 100$ iterations and $\alpha = 0.05$ significance level.

Figure 2a displays the spectral profiles of TD ($f_{Y_1;Y_2}$, with $Y_1 = H$, $Y_2 = S$), GC ($f_{Y_1 \rightarrow Y_2}$ and $f_{Y_2 \rightarrow Y_1}$), and IC ($f_{Y_1;Y_2}$) - red line - along with their surrogate distributions - gray lines and shades. Figure 2b depicts the corresponding information-theoretic domain measures, obtained by integrating the spectral measures in a) over the TIME ($0 - f_s/2$), LF ($0.04 - 0.15$ Hz) and HF ($0.15 - 0.4$ Hz) bands as in Equations 39a–39c. As regards the total coupling (panel a, left), two peaks are observed at ~ 0.1 Hz and ~ 0.35 Hz, corresponding to significant MIR values in

the LF and HF bands, respectively (panel b, left). These findings are consistent with previous studies confirming the presence of LF oscillatory rhythms in heart rate and blood pressure variability, as well as the dominant role of HF spectral components of the respiratory signal during supine rest [see, e.g., Pagani et al. (1986); Montano et al. (1994); Cooley et al. (1998); Stefanovska (2002)]. The observed coupling primarily occurs in the direction $Y_2 \rightarrow Y_1$, that is, from S to H, rather than in the reverse direction (panels a,b, middle). This is supported by the low and statistically non-significant values of both the frequency domain measure $f_{Y_1 \rightarrow Y_2}$ and the time domain measure $T_{Y_1 \rightarrow Y_2}$. In contrast, $f_{Y_2 \rightarrow Y_1}$ exhibits two distinct peaks in the LF and HF bands, corresponding to sympathetic and vagal activity, respectively. The instantaneous interactions are predominantly concentrated in the LF band, as indicated by significance of the IT measure $I_{Y_1;Y_2}$ (panel b, right). Although the spectral profile $f_{Y_1;Y_2}$ (panel a, right) also shows a peak in the HF band, the integrated contribution over this band is not statistically significant. This behaviour is expected, as systolic arterial pressure and heart period are known to interact at zero lag due to shared physiological mechanisms such as the baroreflex and central autonomic regulation. This phenomenon has been documented in several studies (e.g., Porta et al., 2000; Nollo et al., 2005), supporting the interpretation that instantaneous causality between these cardiovascular parameters reflects genuine physiological coupling rather than non-physiological factors (e.g., unobserved confounders) (Faes et al., 2013a; Nuzzi et al., 2021).

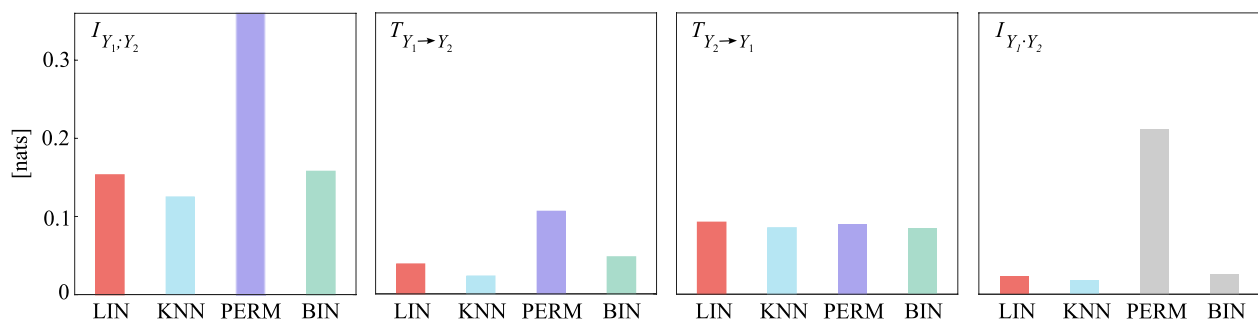


FIGURE 3

Different estimation approaches of coupling and causality measures suggest a bidirectional transfer of information in representative climate time series. The time domain MIR ($I_{Y_1:Y_2}$), TE ($T_{Y_1 \rightarrow Y_2}$, $T_{Y_2 \rightarrow Y_1}$) and IT ($I_{Y_1:Y_2}$) measures are estimated exploiting the linear parametric (LIN), k -nearest neighbours (KNN), permutation-based (PERM) and binning (BIN) approaches, with $Y_1 = \text{SOI}$, $Y_2 = \text{NINO34}$. Grey bars indicate non significant values according to surrogate data analysis.

6.2 Case study in climate science

In this section, to showcase the use of the tools presented in this paper also outside the field of Network Physiology, we consider an exemplary case study in climate science, i.e., we investigate the interactions among the most representative indices descriptive of El Niño and the Southern Oscillation (ENSO). ENSO is a periodic fluctuation in the sea surface temperature and air pressure of the atmosphere overlying the equatorial Pacific Ocean, which is considered as the most prominent interannual climate variability on Earth (McPhaden et al., 2006). Since the exact initiating causes of an ENSO warm or cool events are not fully understood, it is important to analyze the statistical relationship between its two main components, i.e., atmospheric pressure and sea surface temperature. Such components are measured respectively by NINO34 (the East Central Tropical Pacific sea surface temperature anomaly, also called El Niño) and SOI (Southern Oscillation Index, the standardized difference in surface air pressure between Tahiti and Darwin), and are dynamically related to several other indexes that represent large scale climate patterns (Chang et al., 2003; Silini et al., 2023; Pinto et al., 2025c). The analyzed climate indices are taken from a public database (Silini et al., 2023), of which we consider the series SOI and NINO34 measured with a monthly sampling rate during the period 1950–2016 (792 data points) for which all time series values are available. The series were first detrended and deseasonalized.

Model-based and model-free information-theoretic measures of MIR, TE and IT were computed exploiting the linear parametric, KNN, binning and permutation approaches. Specifically, linear parametric measures were computed from the parameters of an ARX model (least squares estimation, model order set according to the AIC - maximum scanned model order: 12 (Faes et al., 2012a); selected model order: 10); restricted models were obtained via resolution of the YW equations (Section 2). The KNN estimator was implemented through the uniform embedding procedure, by fixing a maximum number of past samples in the embedding vectors of $q = 3$ samples, while a number of neighbours $k = 10$ was set to estimate the information measures of coupling and causality (Section 3.1). As regards the binning approach, we set $b = 3$ and $q = 2$, so as to deal with a number of quantization levels equal to $3^{2 \times 2} = 81$ and adhere with the empirical rule stated in Section 3.2. On

the other hand, we used the pattern length typically adopted in permutation entropy analyses when the permutation approach was implemented, which is the minimum advised to guarantee variability in the discrete patterns, i.e., $q = 3$ (Section 3.3). For all the non-linear estimators, the embedding delay was set equal to $\tau = 3$. All the procedures are detailed in Section 4. Surrogate data analysis was applied as in Section 4.1 to assess the statistical significance of the computed measures, with a minimum shift of $\tau_{\min} = 20$ lags, $N_s = 100$ iterations and $\alpha = 0.05$ significance level.

Figure 3 summarizes the coupling and causality measures of MIR ($I_{Y_1:Y_2}$, with $Y_1 = \text{SOI}$, $Y_2 = \text{NINO34}$), TE ($T_{Y_1 \rightarrow Y_2}$ and $T_{Y_2 \rightarrow Y_1}$) and IT ($I_{Y_1:Y_2}$) computed for a representative pair of climate time series estimated using the linear parametric (LIN) and three different model-free approaches, i.e., the KNN, permutation (PERM) and binning (BIN) approaches. Gray bars indicate non significant values according to surrogate data analysis. Significant dynamic coupling is detected between the two processes. The same behaviour is observed for the TE in both directions (middle), that is $T_{Y_1 \rightarrow Y_2}$ and $T_{Y_2 \rightarrow Y_1}$. Finally, as regards the instantaneous transfer, only the LIN and KNN methods yield statistically significant results, although the values are very low and close to zero ($I_{\text{LIN}} = 0.022773$ nats; $I_{\text{KNN}} = 0.017997$ nats), probably suggesting that the NINO34 and SOI time series do not exhibit meaningful instantaneous information exchange. Rather, their interaction is likely dominated by delayed, feedback-driven dynamics typical of coupled ocean-atmosphere processes such as ENSO. This aligns with previous studies highlighting the temporal structure and lagged dependencies in ENSO-related indices, where causality is more evident over seasonal timescales than at zero lag (Faes et al., 2025b).

Overall, although the LIN, KNN, PERM and BIN estimators aim to quantify the same underlying information flow, their varying theoretical assumptions (continuous- or discrete-valued random variables), data manipulation steps (no transformation, discretization based on quantization or based on ordinal patterns), estimation approach (model-based vs. model-free), as well as their different sensitivity to data length, noise, and embedding parameters, explain why distinct numerical values may arise, even substantial as depicted in Figure 3. As a matter of fact, estimating information-theoretic measures such as transfer

entropy or mutual information rate yields results that depend strongly on how probability distributions are inferred or system dynamics are modelled. Therefore, comparison between measures derived from different approaches should be avoided. Moreover, given an estimation approach, the parameter setting should be as much as possible uniform when comparing the same measure across different experimental conditions. In general, the selection of the appropriate estimator should be guided by the nature of the data, including its linearity, stationarity, noise characteristics, and available sample size.

7 Conclusion

This work provides a comprehensive review, theoretical description and practical implementation of the most popular time-domain, spectral and information-theoretic approaches for the investigation of both symmetrical and directional interactions in bivariate time series. Coupling and causality measures are described in their formulation, evaluated critically highlighting advantages and limitations, connected identifying their reciprocal relations, and showcased in exemplary applications in Network Physiology and Climate Science. Thanks to the freely available toolbox that practically implements the measures using model-based and model-free estimators, our groundwork provides researchers with a robust foundation for quantifying and interpreting several bivariate interdependencies across a wide range of applications.

The practical implementation of the coupling and causality measures through both model-based and model-free approaches allows a complete characterization of the complex interplay occurring in a variety of real-world scenarios. We show how parametric modelling offers interpretability and computational efficiency under the assumption of joint Gaussianity of the bivariate process analysed, and that fully model-free estimation techniques, including binning, permutation and nearest-neighbours estimators, achieve a non-linear description of the complex interdependencies among the data. Although inherently model-free, information-theoretic measures are herein contextualized through linear model-based interpretations, which enable frequency-specific insights into oscillatory dynamics.

The present work can thus serve as a driving force for future endeavours in the development and critical assessment of functional connectivity measure in bivariate systems, as well as help researchers to test such measures in a variety of applicative scenarios where the activity of dynamic systems is measured in terms of time series. Moreover, the systematic description and categorization of bivariate measures pursued in this work can pose solid basis to extend them to multivariate time series data, in contexts where more than two dynamic processes are simultaneously monitored. This approach is very popular in the field of Network Physiology, historically regarding the reconstruction of causal networks (Günther et al., 2022) where a range of extensions of the methods reviewed here has been proposed in the multivariate setting (Rosenberg et al., 1989; Baccalá and Sameshima, 2001; Faes et al., 2012b; Barnett and Seth, 2014; Montalto et al., 2014; Baccalá and Sameshima, 2021b; Baccalá and Sameshima, 2022), and more recently regarding the study of high-order interactions (Scaglierini et al., 2023) whereby collective interactions among three or more processes which cannot be reduced to pairwise dependencies are increasingly investigated

extending tools for bivariate analysis (Faes et al., 2022; Scaglierini et al., 2023; Faes et al., 2025b; Sparacino et al., 2025b; Mijatovic et al., 2025; Faes et al., 2025a). These causal and high-order analyses of multivariate time series are closely connected (see, e.g., Stramaglia et al. (2024)), and their practical implementation implies to face similar issues to those tackled by the bivariate methods presented here, even though exacerbated by the need to work in higher-dimensional settings. In fact, as dimensionality increases, the joint probability space expands rapidly, raising theoretical constraints, computational demands and the risk of biased estimation, particularly with limited data. Up to now, while multivariate linear methods applied appear feasible across time, frequency and information-theoretic domains, truly multivariate non-parametric approaches to the study of causality networks and high-order interactions are scarce. The extension of the bivariate model-free methods reviewed here will have to face the theoretical and practical challenges posed by the curse of dimensionality, and is an open question for the ongoing research in the field.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://github.com/LauraSparacino>.

Ethics statement

The studies involving humans were approved by Ethics Committee, Sacro Cuore Hospital, Negrar, Italy. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

LS: Writing – original draft, Software, Investigation, Data curation, Visualization, Methodology, Validation, Formal Analysis. HP: Data curation, Formal Analysis, Visualization, Methodology, Software, Investigation, Writing – original draft. CB: Writing – review and editing, Methodology, Formal Analysis, Validation, Data curation, Software. YA: Validation, Formal Analysis, Methodology, Writing – review and editing, Investigation, Supervision. RP: Resources, Validation, Writing – review and editing, Supervision. AR: Supervision, Investigation, Writing – review and editing. LF: Methodology, Conceptualization, Investigation, Project administration, Supervision, Funding acquisition, Writing – review and editing.

Funding

The authors declare that financial support was received for the research and/or publication of this article. 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 by European Union-NextGenerationEU - funds from Italian MUR, D.M. 737/2021 - EUROSTART 2021 research project “Novel Computational tools for Patient Stratification in cardiovascular diseases and brain disorders”. YA, RP, and LF were supported by SiciliAn MicronanOTech Research And Innovation CEntEr “SAMOTHRACE” (MUR, PNRR-M4C2, ECS_00000022), spoke 3–Università degli Studi di Palermo S2-COMMs–Micro and Nanotechnologies for Smart and Sustainable Communities. HP and AR were partially supported by CMUP, member of LASI, which is financed by national funds through Fundação para a Ciência e Tecnologia (FCT), under the project UID/00144/2025 (<https://doi.org/10.54499/UID/00144/2025>). HP thanks FCT, Portugal for the Ph.D. Grant 2022.11423.BD (<https://doi.org/10.54499/2022.11423.BD>).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors RP, APR, and LF declared that they were editorial board members of Frontiers

at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions Automatic Control* 19, 716–723. doi:10.1109/TAC.1974.1100705
- Antonacci, Y., Astolfi, L., Nollo, G., and Faes, L. (2020). Information transfer in linear multivariate processes assessed through penalized regression techniques: validation and application to physiological networks. *Entropy* 22, 732. doi:10.3390/e22070732
- Antonacci, Y., Astolfi, L., and Faes, L. (2021a). “Testing different methodologies for granger causality estimation: a simulation study,” in *2020 28th European signal processing conference (EUSIPCO)*, 940–944. doi:10.23919/Eusipco47968.2020.9287405
- Antonacci, Y., Minati, L., Nuzzi, D., Mijatovic, G., Pernice, R., Marinazzo, D., et al. (2021b). Measuring high-order interactions in rhythmic processes through multivariate spectral information decomposition. *IEEE Access* 9, 149486–149505. doi:10.1109/ACCESS.2021.3124601
- Antonacci, Y., Barà, C., Sparacino, L., Pirovano, I., Mastropietro, A., Rizzo, G., et al. (2025). Spectral information dynamics of cortical signals uncover the hierarchical organization of the human brain's motor network. *IEEE Trans. Biomed. Eng.* 72, 1655–1664. doi:10.1109/TBME.2024.3516943
- Ash, R. B. (2012). Information theory. *Cour. Corp.*
- Azami, H., Faes, L., Escudero, J., Humeau-Heurtier, A., and Silva, L. E. (2023). “Entropy analysis of univariate biomedical signals: review and comparison of methods,” in *Frontiers in entropy across disciplines: a panorama of entropy theory, computation, and applications*. 233–286. doi:10.1142/9789811259401_0009
- Bacalá, L. A., and Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biol. Cybern.* 84, 463–474. doi:10.1007/PL00007990
- Bacalá, L. A., and Sameshima, K. (2021a). Partial directed coherence: twenty years on some history and an appraisal. *Biol. Cybern.* 115, 195–204. doi:10.1007/s00422-021-00880-y
- Bacalá, L. A., and Sameshima, K. (2021b). Frequency domain repercussions of instantaneous granger causality. *Entropy* 23, 1037. doi:10.3390/e23081037
- Bacalá, L. A., and Sameshima, K. (2022). Partial directed coherence and the vector autoregressive modelling myth and a caveat. *Front. Netw. Physiology* 2, 845327. doi:10.3389/fnetp.2022.845327
- Bacalá, L. A., Sameshima, K., Ballester, G., Valle, A., Yoshimoto, C. E., and Timolaria, C. (1998). Studying the interaction between brain via direct coherence and granger causality. *Appl. Signal Process* 5, 40–48.
- Bandt, C., and Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Lett.* 88, 174102. doi:10.1103/PhysRevLett.88.174102
- Baptista, M. S., and Kurths, J. (2008). Transmission of information in active networks. *Phys. Rev. E* 77, 26205. doi:10.1103/PhysRevE.77.026205
- Barà, C., Sparacino, L., Pernice, R., Antonacci, Y., Porta, A., Kugiumtzis, D., et al. (2023). Comparison of discretization strategies for the model-free information-theoretic assessment of short-term physiological interactions. *Chaos Interdiscip. J. Nonlinear Sci.* 33, 033127. doi:10.1063/5.0140641
- Barà, C., Pernice, R., Catania, C. A., Hilal, M., Porta, A., Humeau-Heurtier, A., et al. (2024). Comparison of entropy rate measures for the evaluation of time series complexity: simulations and application to heart rate and respiratory variability. *Biocybern. Biomed. Eng.* 44, 380–392. doi:10.1016/j.bbe.2024.04.004
- Barabási, A.-L. (2013). Network science. *Philosophical Trans. R. Soc. A Math. Phys. Eng. Sci.* 371, 20120375. doi:10.1098/rsta.2012.0375
- Bari, V., Marchi, A., Maria, B. D., Rossato, G., Nollo, G., Faes, L., et al. (2016). Nonlinear effects of respiration on the crosstalk between cardiovascular and cerebrovascular control systems. *Phil. Trans. R. Soc.* 374, 20150179. doi:10.1098/rsta.2015.0179
- Barnett, L., and Seth, A. K. (2014). The mvgc multivariate granger causality toolbox: a new approach to granger-causal inference. *J. Neuroscience Methods* 223, 50–68. doi:10.1016/j.jneumeth.2013.10.018
- Barnett, L., and Seth, A. K. (2015). Granger causality for state-space models. *Phys. Rev. E* 91, 040101. doi:10.1103/PhysRevE.91.040101
- Barnett, L., and Seth, A. K. (2023). Dynamical Independence: discovering emergent macroscopic processes in complex dynamical systems. *Phys. Rev. E* 108, 014304. doi:10.1103/PhysRevE.108.014304
- Barnett, L., Barrett, A. B., and Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for gaussian variables. *Phys. Review Letters* 103, 238701. doi:10.1103/PhysRevLett.103.238701
- Barnett, L., Barrett, A. B., and Seth, A. K. (2018). Solved problems for granger causality in neuroscience: a response to stokes and purdon. *NeuroImage* 178, 744–748. doi:10.1016/j.neuroimage.2018.05.067
- Barrett, A. B., Barnett, L., and Seth, A. K. (2010). Multivariate granger causality and generalized variance. *Phys. Rev. E* 81, 041907. doi:10.1103/PhysRevE.81.041907
- Barrett, A. B., Murphy, M., Bruno, M.-A., Noirhomme, Q., Boly, M., Laureys, S., et al. (2012). Granger causality analysis of steady-state electroencephalographic signals during propofol-induced anaesthesia. *PloS One* 7, e29072. doi:10.1371/journal.pone.0029072
- Baselli, G., Porta, A., Rimoldi, O., Pagani, M., and Cerutti, S. (1997). Spectral decomposition in multichannel recordings based on multivariate parametric identification. *IEEE Transactions Biomedical Engineering* 44, 1092–1101. doi:10.1109/10.641336

- Bashan, A., Bartsch, R. P., Kantelhardt, J. W., Havlin, S., and Ivanov, P. C. (2012). Network physiology reveals relations between network topology and physiological function. *Nat. Communications* 3, 702. doi:10.1038/ncomms1705
- Bassett, D. S., and Sporns, O. (2017). Network neuroscience. *Nat. Neuroscience* 20, 353–364. doi:10.1038/nn.4502
- Benesty, J., Chen, J., Huang, Y., and Cohen, I. (2009). *Pearson correlation coefficient*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-00296-0_5
- Blackman, R. B., and Tukey, J. W. (1958). The measurement of power spectra from the point of view of communications Engineering—Part I. *Bell Syst. Tech. J.* 37, 185–282. doi:10.1002/j.1538-7305.1958.tb03874.x
- Bressler, S. L., and Seth, A. K. (2011). Wiener–granger causality: a well established methodology. *Neuroimage* 58, 323–329. doi:10.1016/j.neuroimage.2010.02.059
- Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R., and Bressler, S. L. (2004). Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by granger causality. *Proc. Natl. Acad. Sci.* 101, 9849–9854. doi:10.1073/pnas.0308538101
- Buckner, R. L., Snyder, A. Z., Shannon, B. J., LaRossa, G., Sachs, R., Fotenos, A. F., et al. (2005). Molecular, structural, and functional characterization of alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *J. Neurosci.* 25, 7709–7717. doi:10.1523/JNEUROSCI.2177-05.2005
- Cairo, B., Gelpi, F., Bari, V., Anguissola, M., Singh, P., Maria, B. D., et al. (2025). A model-based spectral directional approach reveals the long-term impact of covid-19 on cardiorespiratory control and baroreflex. *Biomed. Eng. Online* 24, 8. doi:10.1186/s12938-024-01327-8
- Cellucci, C. J., Albano, A. M., and Rapp, P. E. (2005). Statistical validation of mutual information calculations: comparison of alternative numerical algorithms. *Phys. Rev. E* 71, 066208. doi:10.1103/PhysRevE.71.066208
- Chang, P., Saravanan, R., and Ji, L. (2003). Tropical Atlantic seasonal predictability: the roles of El Niño remote influence and thermodynamic air-sea feedback. *Geophys. Res. Lett.* 30. doi:10.1029/2002GL016119
- Charleston-Villalobos, S., Javorka, M., Faes, L., and Voss, A. (2023). Granger causality and information transfer in physiological systems: basic research and applications
- Chicharro, D. (2011). On the spectral formulation of granger causality. *Biol. Cybernetics* 105, 331–347. doi:10.1007/s00422-011-0469-z
- Clemson, P. T., Hoag, J. B., Cooke, W. H., Eckberg, D. L., and Stefanovska, A. (2022). Beyond the baroreflex: a new measure of autonomic regulation based on the time-frequency assessment of variability, phase coherence and couplings. *Front. Netw. Physiology* 2, 891604. doi:10.3389/fnetp.2022.891604
- Cliff, O. M., Novelli, L., Fulcher, B. D., Shine, J. M., and Lizier, J. T. (2021). Assessing the significance of directed and multivariate measures of linear dependence between time series. *Phys. Rev. Res.* 3, 013145. doi:10.1103/PhysRevResearch.3.013145
- Coben, R., Clarke, A. R., Hudspeth, W., and Barry, R. J. (2008). Eeg power and coherence in autistic spectrum disorder. *Clin. Neurophysiology* 119, 1002–1009. doi:10.1016/j.clinph.2008.01.013
- Cohen, M. A., and Taylor, J. A. (2002). Short-term cardiovascular oscillations in man: measuring and modelling the physiologies. *J. Physiology* 542, 669–683. doi:10.1113/jphysiol.2002.017483
- Cooley, R. L., Montano, N., Cogliati, C., Van de Borne, P., Richenbacher, W., Oren, R., et al. (1998). Evidence for a central origin of the low-frequency oscillation in rr-interval variability. *Circulation* 98, 556–561. doi:10.1161/01.CIR.98.6.556
- Cover, T. M. (1999). *Elements of information theory*. John Wiley and Sons. doi:10.1002/047174882X
- Daoud, M., Ravier, P., and Butelli, O. (2018). Use of cardiorespiratory coherence to separate spectral bands of the heart rate variability. *Biomed. Signal Process. Control* 46, 260–267. doi:10.1016/j.bspc.2018.08.003
- Darbellay, G., and Vajda, I. (1999). Estimation of the information by an adaptive partitioning of the observation space. *IEEE Trans. Inf. Theory* 45, 1315–1321. doi:10.1109/18.761290
- Difrancesco, S., van Baardewijk, J., Cornelissen, A., Varon, C., Hendriks, R., and Brouwer, A. (2023). Exploring the use of granger causality for the identification of chemical exposure based on physiological data. *Front. Netw. Physiology* 3, 1106650. doi:10.3389/fnetp.2023.1106650
- Ding, M., Chen, Y., and Bressler, S. L. (2006). *Granger causality: basic theory and application to neuroscience*. John Wiley and Sons, Ltd, 437–460. chap. 17. doi:10.1002/9783527609970.ch17
- Duncan, T. E. (1970). On the calculation of mutual information. *SIAM J. Appl. Math.* 19, 215–220. doi:10.1137/0119020
- Edinburgh, T., Eglen, S. J., and Ercole, A. (2021). Causality indices for bivariate time series data: a comparative review of performance. *Chaos Interdiscip. J. Nonlinear Sci.* 31, 083111. doi:10.1063/5.0053519
- Faes, L., and Nollo, G. (2010). Extended causal modeling to assess partial directed coherence in multiple time series with significant instantaneous interactions. *Biol. Cybernetics* 103, 387–400. doi:10.1007/s00422-010-0406-6
- Faes, L., and Nollo, G. (2013). Measuring frequency domain granger causality for multiple blocks of interacting time series. *Biol. Cybernetics* 107, 217–232. doi:10.1007/s00422-013-0547-5
- Faes, L., and Porta, A. (2014). “Conditional entropy-based evaluation of information dynamics in physiological systems,” in *Directed information measures in neuroscience* (Springer), 61–86. doi:10.1007/978-3-642-54474-3_3
- Faes, L., Widesott, L., Del Greco, M., Antolini, R., and Nollo, G. (2005). Causal cross-spectral analysis of heart rate and blood pressure variability for describing the impairment of the cardiovascular control in neurally mediated syncope. *IEEE Transactions Biomedical Engineering* 53, 65–73. doi:10.1109/TBME.2005.859788
- Faes, L., Nollo, G., and Porta, A. (2011). Information-based detection of nonlinear granger causality in multivariate processes via a nonuniform embedding technique. *Phys. Rev. E—Statistical, Nonlinear, Soft Matter Phys.* 83, 051112. doi:10.1103/PhysRevE.83.051112
- Faes, L., Erla, S., and Nollo, G. (2012a). Measuring connectivity in linear multivariate processes: definitions, interpretation, and practical analysis. *Comput. Mathematical Methods Medicine* 2012, 140513. doi:10.1155/2012/140513
- Faes, L., Nollo, G., and Porta, A. (2012b). Non-uniform multivariate embedding to assess the information transfer in cardiovascular and cardiorespiratory variability series. *Comput. Biology Medicine* 42, 290–297. doi:10.1016/j.combiomed.2011.02.007
- Faes, L., Erla, S., Porta, A., and Nollo, G. (2013a). A framework for assessing frequency domain causality in physiological time series with instantaneous effects. *Philosophical Trans. R. Soc. A Math. Phys. Eng. Sci.* 371, 20110618. doi:10.1098/rsta.2011.0618
- Faes, L., Montalto, A., Nollo, G., and Marinazzo, D. (2013b). “Information decomposition of short-term cardiovascular and cardiorespiratory variability,” in *Computing in cardiology* 2013, 113–116.
- Faes, L., Porta, A., Rossato, G., Adami, A., Tonon, D., Corica, A., et al. (2013c). Investigating the mechanisms of cardiovascular and cerebrovascular regulation in orthostatic syncope through an information decomposition strategy. *Aut. Neurosci.* 178, 76–82. doi:10.1016/j.autneu.2013.02.013
- Faes, L., Kugiumtzis, D., Nollo, G., Jurysta, F., and Marinazzo, D. (2015a). Estimating the decomposition of predictive information in multivariate systems. *Phys. Rev. E* 91, 032904. doi:10.1103/PhysRevE.91.032904
- Faes, L., Porta, A., and Nollo, G. (2015b). Information decomposition in bivariate systems: theory and application to cardiorespiratory dynamics. *Entropy* 17, 277–303. doi:10.3390/e17010277
- Faes, L., Marinazzo, D., Nollo, G., and Porta, A. (2016a). An information-theoretic framework to map the spatiotemporal dynamics of the scalp electroencephalogram. *IEEE Trans. Biomed. Eng.* 63, 2488–2496. doi:10.1109/TBME.2016.2569823
- Faes, L., Porta, A., Nollo, G., and Javorka, M. (2016b). Information decomposition in multivariate systems: definitions, implementation and application to cardiovascular networks. *Entropy* 19 (5), 5. doi:10.3390/e19010005
- Faes, L., Marinazzo, D., and Stramaglia, S. (2017a). Multiscale information decomposition: exact computation for multivariate gaussian processes. *Entropy* 19, 408. doi:10.3390/e19080408
- Faes, L., Stramaglia, S., and Marinazzo, D. (2017b). On the interpretability and computational reliability of frequency-domain granger causality. *F1000Research* 6, 1710. doi:10.12688/f1000research.12694.1
- Faes, L., Pernice, R., Mijatovic, G., Antonacci, Y., Krohova, J. C., Javorka, M., et al. (2021). Information decomposition in the frequency domain: a new framework to study cardiovascular and cardiorespiratory oscillations. *Philosophical Trans. R. Soc. A* 379, 20200250. doi:10.1098/rsta.2020.0250
- Faes, L., Mijatovic, G., Antonacci, Y., Pernice, R., Barà, C., Sparacino, L., et al. (2022). A new framework for the time-and frequency-domain assessment of high-order interactions in networks of random processes. *IEEE Trans. Signal Process.* 70, 5766–5777. doi:10.1109/TSP.2022.3221892
- Faes, L., Mijatovic, G., Sparacino, L., and Porta, A. (2025a). Predictive information decomposition as a tool to quantify emergent dynamical behaviors in physiological networks. *IEEE Trans. Biomed. Eng.* 72, 3515–3524. doi:10.1109/TBME.2025.3570937
- Faes, L., Sparacino, L., Mijatovic, G., Antonacci, Y., Ricci, L., Marinazzo, D., et al. (2025b). Partial information rate decomposition. *Phys. Rev. Lett.* 135, 187401. doi:10.1103/nrwj-n8lj
- Freeman, J. R. (1983). Granger causality and the times series analysis of political relationships. *Am. J. Political Sci.* 27, 327–358. doi:10.2307/2111021
- Gelfand, I. M., and Yaglom, A. M. (1959). “Calculation of the amount of information about a random function contained,” in *Another such function* (Providence, RI: American Mathematical Society).
- Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *J. Am. Statistical Association* 77, 304–313. doi:10.1080/01621459.1982.10477803
- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424–438. doi:10.2307/1912791
- Günther, M., Kantelhardt, J. W., and Bartsch, R. P. (2022). The reconstruction of causal networks in physiology. *Front. Netw. Physiology* 2, 893743. doi:10.3389/fnetp.2022.893743

- Hannan, E. J. (1979). "The statistical theory of linear systems," *Dev. Statistics*, 2, 83–121. doi:10.1016/b978-0-12-426602-5.50008-6
- Ince, R. A. A., Panzeri, S., and Kayser, C. (2013). Neural codes formed by small and temporally precise populations in auditory cortex. *J. Neurosci.* 33, 18277–18287. doi:10.1523/JNEUROSCI.2631-13.2013
- Jahani, A., Begin, C., Toffa, D. H., Obaid, S., Nguyen, D. K., and Bou Assi, E. (2025). Preictal connectivity dynamics: exploring inflow and outflow in iieeg networks. *Front. Netw. Physiology* 5, 1539682. doi:10.3389/fnetp.2025.1539682
- Jwo, D.-J., Cho, T.-S., and Biswal, A. (2023). Geometric insights into the multivariate gaussian distribution and its entropy and mutual information. *Entropy* 25, 1177. doi:10.3390/e25081177
- Karimi, M. (2011). Order selection criteria for vector autoregressive models. *Signal Process.* 91, 955–969. doi:10.1016/j.sigpro.2010.09.021
- Kay, S. M. (1988). *Modern spectral estimation: theory and application*. Englewood Cliffs, NJ: Prentice Hall.
- Korhonen, I., Mainardi, L., Loula, P., Carraut, G., Baselli, G., and Bianchi, A. (1996). Linear multivariate models for physiological signal analysis: theory. *Comput. Methods Programs Biomed.* 51, 85–94. doi:10.1016/0169-2607(96)01764-6
- Koutlis, C., Kimiskidis, V. K., and Kugiumtzis, D. (2021). Comparison of causality network estimation in the sensor and source space: simulation and application on eeg. *Front. Netw. Physiology* 1, 706487. doi:10.3389/fnetp.2021.706487
- Kozachenko, L. F., and Leonenko, N. N. (1987). Sample estimate of the entropy of a random vector. *Probl. Peredachi Inf.* 23, 9–16.
- Kraskov, A., Stögbauer, H., and Grassberger, P. (2004). Estimating mutual information. *Phys. Rev. E* 69, 066138. doi:10.1103/PhysRevE.69.066138
- Kugiumtzis, D. (2013). Direct-coupling information measure from nonuniform embedding. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* 87, 062918. doi:10.1103/PhysRevE.87.062918
- Lombardi, D., and Pant, S. (2016). Nonparametric k -nearest-neighbor entropy estimator. *Phys. Rev. E* 93, 013310. doi:10.1103/PhysRevE.93.013310
- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science and Business Media.
- Martins, A., Pernice, R., Amado, C., Rocha, A. P., Silva, M. E., Javorka, M., et al. (2020). Multivariate and multiscale complexity of long-range correlated cardiovascular and respiratory variability series. *Entropy* 22, 315. doi:10.3390/e22030315
- McGraw, M. C., and Barnes, E. A. (2018). Memory matters: a case for granger causality in climate variability studies. *J. Climate* 31, 3289–3300. doi:10.1175/JCLI-D-17-0334.1
- McPhaden, M. J., Zebiak, S. E., and Glantz, M. H. (2006). Enso as an integrating concept in Earth science. *Science* 314, 1740–1745. doi:10.1126/science.1132588
- Mijatovic, G., Antonacci, Y., Javorka, M., Marinazzo, D., Stramaglia, S., and Faes, L. (2025). Network representation of higher-order interactions based on information dynamics. *IEEE Trans. Netw. Sci. Eng.* 12, 1872–1884. doi:10.1109/tNSE.2025.3540982
- Milde, T., Leistritz, L., Astolfi, L., Miltner, W. H., Weiss, T., Babiloni, F., et al. (2010). A new kalman filter approach for the estimation of high-dimensional time-variant multivariate ar models and its application in analysis of laser-evoked brain potentials. *Neuroimage* 50, 960–969. doi:10.1016/j.neuroimage.2009.12.110
- Möller, E., Schack, B., Arnold, M., and Witte, H. (2001). Instantaneous multivariate eeg coherence analysis by means of adaptive high-dimensional autoregressive models. *J. Neuroscience Methods* 105, 143–158. doi:10.1016/S0165-0270(00)00350-2
- Montalto, A., Faes, L., and Marinazzo, D. (2014). Mute: a matlab toolbox to compare established and novel estimators of the multivariate transfer entropy. *PLoS One* 9, e109462. doi:10.1371/journal.pone.0109462
- Montano, N., Ruscone, T. G., Porta, A., Lombardi, F., Pagani, M., and Malliani, A. (1994). Power spectrum analysis of heart rate variability to assess the changes in sympathovagal balance during graded orthostatic tilt. *Circulation* 90, 1826–1831. doi:10.1161/01.CIR.90.4.1826
- Nollo, G., Faes, L., Porta, A., Antolini, R., and Ravelli, F. (2005). Exploring directionality in spontaneous heart period and systolic pressure variability interactions in humans: implications in the evaluation of baroreflex gain. *Am. J. Physiology-Heart Circulatory Physiology* 288, H1777–H1785. doi:10.1152/ajpheart.00594.2004
- Nuzzi, D., Faes, L., Javorka, M., Marinazzo, D., and Stramaglia, S. (2021). Inclusion of instantaneous influences in the spectral decomposition of causality: application to the control mechanisms of heart rate variability. *Eur. Signal Process. Conf. 2021-January*, 930–934. doi:10.23919/Eusipco47968.2020.9287642
- Orini, M., Bailón, R., Mainardi, L. T., Laguna, P., and Flandrin, P. (2011). Characterization of dynamic interactions between cardiovascular signals by time-frequency coherence. *IEEE Transactions Biomedical Engineering* 59, 663–673. doi:10.1109/TBME.2011.2171959
- Pagani, M., Lombardi, F., Guzzetti, S., Rimoldi, O., Furlan, R., Pizzinelli, P., et al. (1986). Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog. *Circulation Research* 59, 178–193. doi:10.1161/01.RES.59.2.178
- Papoulis, A., and Pillai, S. U. (2002). *Probability, random variables, and stochastic processes*. 4th edn. New York: McGraw-Hill.
- Pereda, E., Quiroga, R. Q., and Bhattacharya, J. (2005). Nonlinear multivariate analysis of neurophysiological signals. *Prog. Neurobiology* 77, 1–37. doi:10.1016/j.pneurobio.2005.10.003
- Pernice, R., Faes, L., Feucht, M., Benninger, F., Mangione, S., and Schiecke, K. (2022a). Pairwise and higher-order measures of brain-heart interactions in children with temporal lobe epilepsy. *J. Neural Eng.* 19, 045002. doi:10.1088/1741-2552/ac7fba
- Pernice, R., Sparacino, L., Bari, V., Gelpi, F., Cairo, B., Mijatovic, G., et al. (2022b). Spectral decomposition of cerebrovascular and cardiovascular interactions in patients prone to postural syncope and healthy controls. *Aut. Neurosci.* 103021doi, 103021. doi:10.1016/j.autneu.2022.103021
- Pichot, V., Corbier, C., and Chouchou, F. (2024). The contribution of granger causality analysis to our understanding of cardiovascular homeostasis: from cardiovascular and respiratory interactions to central autonomic network control. *Front. Netw. Physiology* 4, 1315316. doi:10.3389/fnetp.2024.1315316
- Pinna, G., Maestri, R., and Di Cesare, A. (1996). Application of time series spectral analysis theory: analysis of cardiovascular variability signals. *Med. Biol. Eng. Comput.* 34, 142–148. doi:10.1007/BF02520019
- Pinto, H., Lazic, I., Antonacci, Y., Pernice, R., Gu, D., Barà, C., et al. (2024). Testing dynamic correlations and nonlinearity in bivariate time series through information measures and surrogate data analysis. *Front. Netw. Physiology* 4, 1385421. doi:10.3389/fnetp.2024.1385421
- Pinto, H., Antonacci, Y., Barà, C., Pernice, R., Lazic, I., Faes, L., et al. (2025a). Estimating the mutual information rate of short time series from coupled dynamic systems. SSRN. doi:10.2139/ssrn.5118280
- Pinto, H., Antonacci, Y., Barà, C., Pernice, R., Lazic, I., Faes, L., et al. (2025b). Estimating the mutual information rate of short time series from coupled dynamic systems. *Commun. Nonlinear Sci. Numer. Simul.* 152, 109383. doi:10.1016/j.cnsns.2025.109383
- Pinto, H., Antonacci, Y., Mijatovic, G., Sparacino, L., Stramaglia, S., Faes, L., et al. (2025c). Information-theoretic sequential framework to elicit dynamic high-order interactions in high-dimensional network processes. *Mathematics* 13, 2081. doi:10.3390/math13132081
- Pinto, H., Dias, C., Barà, C., Antonacci, Y., Faes, L., and Rocha, A. P. (2025d). "Exploring the mutual information rate decomposition in situations of pathological stress," in *New frontiers in statistics and data science*. Editors L. Henriques-Rodrigues, R. Menezes, L. M. Machado, S. Faria, and M. de Carvalho (Cham: Springer Nature Switzerland), 243–257. doi:10.1007/978-3-031-68949-9_18
- Porta, A., and Faes, L. (2013). Assessing causality in brain dynamics and cardiovascular control. *Philosophical Trans. R. Soc. A Math. Phys. Eng. Sci.* 371, 20120517. doi:10.1098/rsta.2012.0517
- Porta, A., and Faes, L. (2015). Wiener-granger causality in network physiology with applications to cardiovascular control and neuroscience. *Proc. IEEE* 104, 282–309. doi:10.1109/JPROC.2015.2476824
- Porta, A., Baselli, G., Rimoldi, O., Malliani, A., and Pagani, M. (2000). Assessing baroreflex gain from spontaneous variability in conscious dogs: role of causality and respiration. *Am. J. Physiology-Heart Circulatory Physiology* 279, H2558–H2567. doi:10.1152/ajpheart.2000.279.5.H2558
- Porta, A., Furlan, R., Rimoldi, O., Pagani, M., Malliani, A., and Van De Borne, P. (2002a). Quantifying the strength of the linear causal coupling in closed loop interacting cardiovascular variability signals. *Biol. Cybernetics* 86, 241–251. doi:10.1007/s00422-001-0292-z
- Porta, A., Guzzetti, S., Montano, N., Furlan, R., Pagani, M., Malliani, A., et al. (2002b). Entropy, entropy rate, and pattern classification as tools to typify complexity in short heart period variability series. *IEEE Trans. Biomed. Eng.* 48, 1282–1291. doi:10.1109/10.959324
- Porta, A., De Maria, B., Bari, V., Marchi, A., Marinou, K., Sideri, R., et al. (2016). "Comparison between k-nearest-neighbor approaches for conditional entropy estimation: application to the assessment of the cardiac control in amyotrophic lateral sclerosis patients," in *2016 38th annual international conference of the IEEE engineering in medicine and biology Society (EMBC)*, 2933–2936. doi:10.1109/EMBC.2016.7591344
- Porta, A., Bari, V., Ranuzzi, G., De Maria, B., and Baselli, G. (2017). Assessing multiscale complexity of short heart rate variability series through a model-based linear approach. *Chaos Interdiscip. J. Nonlinear Sci.* 27, 093901. doi:10.1063/1.4999353
- Priestley, M. B. (1981). *Spectral analysis and time series: probability and mathematical statistics*, 04; QA280. New York: Academic Press, P7.
- Reinsel, G. C. (2003). *Elements of multivariate time series analysis*. Springer Science and Business Media. doi:10.1007/978-1-4684-0198-1
- Rodgers, J. L., and Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *Am. Statistician* 42, 59–66. doi:10.2307/2685263

- Rosenberg, J., Amjad, A., Breeze, P., Brillinger, D., and Halliday, D. (1989). The fourier approach to the identification of functional coupling between neuronal spike trains. *Prog. Biophysics Mol. Biol.* 53, 1–31. doi:10.1016/0079-6107(89)90004-7
- Rosenblum, M., and Kurths, J. (1998). “Analysing synchronization phenomena from bivariate data by means of the hilbert transform,” in *Nonlinear analysis of physiological data* (Springer), 91–99.
- Runge, J., Heitzig, J., Petoukhov, V., and Kurths, J. (2012). Escaping the curse of dimensionality in estimating multivariate transfer entropy. *Phys. Rev. Lett.* 108, 258701. doi:10.1103/PhysRevLett.108.258701
- Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. *Sci. Adv.* 5, eaau4996. doi:10.1126/sciadv.aau4996
- Saito, Y., and Harashima, H. (1981). “Tracking of information within multichannel eeg record causal analysis in eeg,” in *recent advances in [EEG] and [EMG] data processing*, Editors N yamaguchi, and K fujisawa,
- Santiago-Fuentes, L. M., Charleston-Villalobos, S., González-Camarena, R., Voss, A., Mejía-Avila, M. E., Buendía-Roldán, I., et al. (2022). Effects of supplemental oxygen on cardiovascular and respiratory interactions by extended partial directed coherence in idiopathic pulmonary fibrosis. *Front. Netw. Physiology* 2, 834056. doi:10.3389/fnetp.2022.834056
- Scagliarini, T., Sparacino, L., Faes, L., Marinazzo, D., and Stramaglia, S. (2023). Gradients of o-information highlight synergy and redundancy in physiological applications. *Front. Netw. Physiology* 3, 1335808. doi:10.3389/fnetp.2023.1335808
- Schögl, A. (2006). A comparison of multivariate autoregressive estimators. *Signal Processing* 86, 2426–2429. doi:10.1016/j.sigpro.2005.11.007
- Schreiber, T. (2000). Measuring information transfer. *Phys. Rev. Lett.* 85, 461–464. doi:10.1103/PhysRevLett.85.461
- Schulz, S., Adochiei, F.-C., Edu, I.-R., Schroeder, R., Costin, H., Bär, K.-J., et al. (2013). Cardiovascular and cardiorespiratory coupling analyses: a review. *Philosophical Trans. R. Soc. A Math. Phys. Eng. Sci.* 371, 20120191. doi:10.1098/rsta.2012.0191
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals Statistics* 6, 461–464. doi:10.1214/aos/1176344136
- Seth, A. K., Barrett, A. B., and Barnett, L. (2015). Granger causality analysis in neuroscience and neuroimaging. *J. Neurosci.* 35, 3293–3297. doi:10.1523/JNEUROSCI.4399-14.2015
- Shaffer, F., and Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Front. Public Health* 5, 258. doi:10.3389/fpubh.2017.00258
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27, 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x
- Shao, K., Logothetis, N. K., and Besserve, M. (2023). Information theoretic measures of causal influences during transient neural events. *Front. Netw. Physiology* 3, 1085347. doi:10.3389/fnetp.2023.1085347
- Silini, R., Tirabassi, G., Barreiro, M., Ferranti, L., and Masoller, C. (2023). Assessing causal dependencies in climatic indices. *Clim. Dyn.* 61, 79–89. doi:10.1007/s00382-022-06562-0
- Smirnov, D. A., and Mokhov, I. I. (2009). From granger causality to long-term causality: application to climatic data. *Phys. Rev. E—Statistical, Nonlinear, Soft Matter Phys.* 80, 016208. doi:10.1103/PhysRevE.80.016208
- Sorelli, M., Hutson, T. N., Iasemidis, L., and Bocchi, L. (2022). Linear and nonlinear directed connectivity analysis of the cardio-respiratory system in type 1 diabetes. *Front. Netw. Physiology* 2, 840829. doi:10.3389/fnetp.2022.840829
- Sparacino, L., Antonacci, Y., Barà, C., Valenti, A., Porta, A., and Faes, L. (2024a). A method to assess granger causality, isolation and autonomy in the time and frequency domains: theory and application to cerebrovascular variability. *IEEE Trans. Biomed. Eng.* 71, 1454–1465. doi:10.1109/TBME.2023.3340011
- Sparacino, L., Antonacci, Y., Barà, C., Švec, D., Javorka, M., and Faes, L. (2024b). A method to assess linear self-predictability of physiologic processes in the frequency domain: application to beat-to-beat variability of arterial compliance. *Front. Netw. Physiology* 4, 1346424. doi:10.3389/fnetp.2024.1346424
- Sparacino, L., Antonacci, Y., Mijatovic, G., and Faes, L. (2025a). Measuring hierarchically-organized interactions in dynamic networks through spectral entropy rates: theory, estimation, and illustrative application to physiological networks. *Neurocomputing* 630, 129675. doi:10.1016/j.neucom.2025.129675
- Sparacino, L., Mijatovic, G., Antonacci, Y., Ricci, L., Marinazzo, D., Stramaglia, S., et al. (2025b). Decomposing multivariate information rates in networks of random processes. *Phys. Rev. E* 112, 044313. doi:10.1103/mn8p-kf6t
- Stefanovska, A. (2002). Cardiorespiratory interactions. *Nonlinear Phenom. Complex Systems* 5, 462–469.
- Stokes, P. A., and Purdon, P. L. (2017). A study of problems encountered in granger causality analysis from a neuroscience perspective. *Proc. National Academy Sciences* 114, E7063–E7072. doi:10.1073/pnas.1704663114
- Stramaglia, S., Faes, L., Cortes, J. M., and Marinazzo, D. (2024). Disentangling high-order effects in the transfer entropy. *Phys. Rev. Res.* 6, L032007. doi:10.1103/PhysRevResearch.6.L032007
- Sugihara, G., May, R., Ye, H., hao Hsieh, C., Deyle, E., Fogarty, M., et al. (2012). Detecting causality in complex ecosystems. *Science* 338, 496–500. doi:10.1126/science.1227079
- Sytkulski, A. M., Olhede, S. C., Lilly, J. M., and Early, J. J. (2017). Frequency-domain stochastic modeling of stationary bivariate or complex-valued signals. *IEEE Trans. Signal Process.* 65, 3136–3151. doi:10.1109/tsp.2017.2686334
- Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation* 93, 1043–1065. doi:10.1161/01.CIR.93.5.1043
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Phys. D. Nonlinear Phenom.* 58, 77–94. doi:10.1016/0167-2789(92)90102-S
- Trujillo, L. T. (2019). K-th nearest neighbor (Knn) entropy estimates of complexity and integration from ongoing and stimulus-evoked electroencephalographic (eeg) recordings of the human brain. *Entropy* 21, 61. doi:10.3390/e21010061
- Vakithilil, N., Sainbhi, A. S., Islam, A., Gomez, A., Stein, K. Y., Froese, L., et al. (2025). Multivariate linear time-series modeling and prediction of cerebral physiologic signals: review of statistical models and implications for human signal analytics. *Front. Netw. Physiology* 5, 1551043. doi:10.3389/fnetp.2025.1551043
- Victor, J. D. (2002). Binless strategies for estimation of information from neural data. *Phys. Rev. E* 66, 51903. doi:10.1103/PhysRevE.66.051903
- Wang, R., Wang, J., Yu, H., Wei, X., Yang, C., and Deng, B. (2015). Power spectral density and coherence analysis of alzheimer’s eeg. *Cogn. Neurodynamics* 9, 291–304. doi:10.1007/s11571-014-9325-x
- Wibral, M., Pampu, N., Priesemann, V., Siebenhühner, F., Seiwert, H., Lindner, M., et al. (2013). Measuring information-transfer delays. *PLOS ONE* 8, 1–19. doi:10.1371/journal.pone.0055809
- Wibral, M., Vicente, R., and Lizier, J. T. (2014). *Directed information measures in neuroscience*, 724. Springer. doi:10.1007/978-3-642-54474-3
- Wiener, N. (1956). “The theory of prediction,” in *Modern mathematics for the engineer*. Editors E. F. Beckenbach (New York, NY: McGraw-Hill) 1, 125–139.
- Wold, H. (1938). *A study in the analysis of stationary time series*. Almqvist and Wiksell. Doctoral thesis.
- Xiong, W., Faes, L., and Ivanov, P. C. (2017). Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: effects of artifacts, nonstationarity, and long-range correlations. *Phys. Rev. E* 95, 1–37. doi:10.1103/PhysRevE.95.062114
- Zhang, H. (2023). *How can I conduct surrogate analyses, and how should I shuffle?* Cham: Springer International Publishing, 567–577. chap. 35. doi:10.1007/978-3-031-20910-9_35
- Zhao, H., and Gui, L. (2019). “Nonparametric and parametric methods of spectral analysis,” in *Matec web conf.*, vol. 283, 07002, doi:10.1051/mateconf/201928307002