frontiers | Research Topics

Granger causality and information transfer in physiological systems: Basic research and applications

Edited by
Sonia Charleston-Villalobos, Michal Javorka,
Luca Faes and Andreas Voss

Published in
Frontiers in Network Physiology

ISSN 1664-8714
ISBN 978-2-8325-3806-7
DOI 10.3389/978-2-8325-3806-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Granger causality and information transfer in physiological systems: Basic research and applications

Topic editors

Sonia Charleston-Villalobos — Metropolitan Autonomous University, Mexico
Michal Javorka — Comenius University, Slovakia
Luca Faes — University of Palermo, Italy
Andreas Voss — Ilmenau University of Technology, Germany

Citation

Charleston-Villalobos, S., Javorka, M., Faes, L., Voss, A., eds. (2023). *Granger causality and information transfer in physiological systems: Basic research and applications*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-3806-7

Table of contents

frontiers | Frontiers in Network Physiology

TYPE Editorial
PUBLISHED 13 October 2023
DOI 10.3389/fnetp.2023.1284256

OPEN ACCESS

EDITED AND REVIEWED BY
Sebastiano Stramaglia,
University of Bari Aldo Moro, Italy

*CORRESPONDENCE
Sonia Charleston-Villalobos,
schv@xanum.uam.mx

RECEIVED 28 August 2023
ACCEPTED 09 October 2023
PUBLISHED 13 October 2023

CITATION
Charleston-Villalobos S, Javorka M, Faes L and Voss A (2023), Editorial: Granger causality and information transfer in physiological systems: basic research and applications.
Front. Netw. Physiol. 3:1284256.
doi: 10.3389/fnetp.2023.1284256

Editorial: Granger causality and information transfer in physiological systems: basic research and applications

Sonia Charleston-Villalobos[1]*, Michal Javorka[2], Luca Faes[3] and Andreas Voss[4]

[1]Department of Electrical Engineering, Metropolitan Autonomous University, Mexico City, Mexico, [2]Jessenius Faculty of Medicine, Comenius University, Martin, Slovakia, [3]Department of Engineering, University of Palermo, Palermo, Italy, [4]Institute for Biomedical Engineering and Computer Science, Ilmenau University of Technology, Ilmenau, Germany

KEYWORDS

Granger causality, information transfer, physiological systems, basic research, applications, network physiology

Editorial on the Research Topic
Granger causality and information transfer in physiological systems: basic research and applications

The concept of causality provides a theoretical framework to gain insights into the mechanisms underlying driver-response relationships in coupled systems by estimating the involved subsystems' directed interaction. In recent years, this approach has become fundamental to investigating coupled dynamical systems in several fields, including physiology, physics, and economics, among others (Bressler and Seth, 2011). In this context, a class of information-theoretic methods has been proposed for causality estimation of time series, starting from the fundamental definition of Granger causality (GC) based on multivariate linear autoregressive models (Barrett et al., 2010) and extending to non-linear and non-parametric information transfer approaches (Faes et al., 2011). In the field of network physiology, causality analysis has become a reference tool (Porta and Faes, 2016; Ivanov, 2021) as it has allowed studying the dynamics of brain connectivity, interactions of the brain with other organ systems, as well as the interactions between pairs of physiological systems such as the cardiovascular, cardiorespiratory, and cerebrovascular ones (Porta and Faes, 2016; Khandoker et al., 2019).

The present Research Topic highlights advances in causality analysis applied to bi- and multivariate physiological and simulated time series. The Research Topic has attracted nine high-quality papers involving novel methodologies, revisiting established approaches, and presenting innovative applications.

The methodological paper from Shao et al. introduced a novel measure of causal influence strength in the context of recurrent transient neural events—the relative dynamic causal strength (rDCS). This measure was compared to previously used tools and verified by simulated and experimentally recorded neurophysiological data. The significance of this approach to studying brain activity mechanisms was further discussed.

A modified, corrected measure of the mutual information rate (cMIR) was developed by Mijatovic et al. for the study of bivariate point processes. The functionality of the measure

was first demonstrated in various simulation studies, showing its good performance in terms of bias reduction, and then in real point process data represented by heartbeat times and arrival times of the sphygmic wave at the body periphery. In the application, a significant coupling between the two processes could be demonstrated in healthy subjects studied across different experimental conditions. The statistically significant changes in cMIR observed during physiological stress suggested that this index may reflect neuroautonomic modulation of heartbeat and vascular dynamics.

As a decision aid for interaction analyses, Günther et al. compared strengths and weaknesses of the GC approach and the Bivariate Phase Rectified Signal Averaging (BPRSA). Among other results, they found that BPRSA requires more data or stronger interactions than GC and that the latter, unlike BPRSA, can detect direct causal relationships from indirect relationships. In contrast to this, BPRSA is suitable for the analysis of non-stationary data. The suitability of GC for the coupling analysis of causal networks in subjects with and without sleep apnea (analysis of heart rate, respiratory rate, and EEG alpha amplitude) was successfully demonstrated.

Baccalà and Sameshima delved into the methodological aspects of the analysis of GC, particularly focusing on the assumptions underlying the computation of GC measures in both time and frequency domains. Importantly, they show that using the very popular linear vector autoregressive model representation of the observed set of time series is not a mandatory requirement for computing GC but is rather a convenient representation that can be substituted by other means of spectral factorization of the spectral density matrix into minimum phase factors.

Important methodological aspects in the computation of GC are also addressed by Koutlis et al., who challenged the notion that the structure of neural connectivity underlying multichannel electroencephalographic signals can be better inferred by analyzing the cortical source time series obtained through inverse source reconstruction. Using both simulated and real signals measured from epileptic subjects, they showed that causality networks constructed at the sensor and source levels differ significantly, and the former can yield better discriminative ability of network topological indices.

Froese et al. employed GC and other statistical methods applied to autonomic response variables such as heart rate and blood pressure variability to investigate the relationship between such variables and cerebrovascular reactivity in patients with moderate to severe traumatic brain injury. The study showed that the sympathetic autonomic response, monitored via spectral indexes and baroreflex sensitivity, is closely linked to impaired cerebrovascular reactivity.

Incorporating measures of physiological interaction in classification tasks is gaining importance due to consideration of the relevant physiological information. In that sense, Difrancesco et al. set out to improve the performance of machine learning models, specifically support vector machine classification algorithms (SVM), by including cardiorespiratory interactions to identify and classify acute toxicity effects of two chemicals on guinea pigs, an opioid and a nerve agent. For this purpose, by continuously monitoring the ECG and respiration signals through wearable sensors, the F-statistics of bivariate prediction models between respiration features, as the tidal volume, and ECG morphology characteristics, as the ST elevation, were fed to an SVM. Although chemicals affected cardiorespiratory interactions differently, they did not improve SVM performance since respiratory features were the most important for the classification task.

In the search for new ways to measure cerebral blood flow (CBF) in humans non-invasively and continuously, González et al. proposed the rheoencephalograph (REG) signal as a CBF surrogate to provide patient monitoring in surgeries. Furthermore, the authors addressed the estimation of bivariate GC to understand better the anesthetics' effects on brain hemodynamics. GC was estimated between REG's features, linear and non-linear, and global hemodynamics by HR or MAP, as well as EEG-based parameters related to the depth of anesthesia. Results pointed out that REG effectively contains information on CBF, and the most frequent causal interaction occurred from CBF_{REG} to the different EEG spectral density bands but not in the opposite direction.

In a clinicall-oriented research study, Santiago-Fuentes et al. applied univariate and causal bivariate analyses on cardiorespiratory signals obtained from patients with idiopathic pulmonary fibrosis (IPF) before and during oxygen supplementation to characterize cardiovascular control and cardiorespiratory interactions. In IPF patients, a shift of the sympathovagal balance towards sympathetic dominance was observed, accompanied by a decreased cardiorespiratory interaction, increased blood pressure variability, and decreased baroreflex function during oxygen supplementation. These results point towards a persistence of autonomic control impairment despite oxygen administration.

The papers of this Research Topic consider fundamental aspects of causality estimation and include novel proposals. Also, most of the contributions showed the power of the causality framework for dissecting intricate directed interactions from physiological networks studied in diverse physiological and pathological scenarios. These efforts are hoped to increase the impact and importance of causality estimation in clinical settings, among others.

Author contributions

SC-V: Writing–original draft, Writing–review and editing. MJ: Writing–original draft, Writing–review and editing. LF: Writing–original draft, Writing–review and editing. AV: Writing–original draft, Writing–review and editing.

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 SC-V, MJ, LF, and AV 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.

Publisher's note

References

Barrett, A. B., Barnett, L., and Seth, A. K. (2010). Multivariate Granger causality and generalized variance. *Phys. Rev. E.* 81 (4), 041907. doi:10.1103/PhysRevE.81.041907

Bressler, S. L., and Seth, A. K. (2011). Wiener–Granger causality: a well established methodology. *Neuroimage* 58 (2), 323–329. doi:10.1016/j.neuroimage.2010.02.059

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.* 83 (5), 051112. doi:10.1103/PhysRevE.83.051112

Ivanov, P. C. (2021). The new field of network physiology: building the human physiolome. *Front. Netw. Physiol.* 1, 711778. doi:10.3389/fnetp.2021.711778

Khandoker, A. H., Schulz, S., Al-Angari, H. M., Voss, A., and Kimura, Y. (2019). Alterations in maternal–fetal heart rate coupling strength and directions in abnormal fetuses. *Front. Physiol.* 10, 482. doi:10.3389/fphys.2019.00482

Porta, A., and Faes, L. (2016). Wiener–granger causality in network physiology with applications to cardiovascular control and neuroscience. *Proc. IEEE* 104 (2), 282–309. doi:10.1109/JPROC.2015.2476824

frontiers in Network Physiology

ORIGINAL RESEARCH
published: 29 September 2021
doi: 10.3389/fnetp.2021.706487

Comparison of Causality Network Estimation in the Sensor and Source Space: Simulation and Application on EEG

Christos Koutlis[1], *Vasilios K. Kimiskidis*[2] *and Dimitris Kugiumtzis*[3]*

[1]Information Technologies Institute, Centre of Research and Technology Hellas, Thessaloniki, Greece, [2]1st Department of Neurology, Medical School, Aristotle University of Thessaloniki, Thessaloniki, Greece, [3]Division of Electronics and Computing, Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece

The usage of methods for the estimation of the true underlying connectivity among the observed variables of a system is increasing, especially in the domain of neuroscience. Granger causality and similar concepts are employed for the estimation of the brain network from electroencephalogram (EEG) data. Also source localization techniques, such as the standardized low resolution electromagnetic tomography (sLORETA), are widely used for obtaining more reliable data in the source space. In this work, connectivity structures are estimated in the sensor and in the source space making use of the sLORETA transformation for simulated and for EEG data with episodes of spontaneous epileptiform discharges (ED). From the comparative simulation study on high-dimensional coupled stochastic and deterministic systems originating in the sensor space, we conclude that the structure of the estimated causality networks differs in the sensor space and in the source space. Moreover, different network types, such as random, small-world and scale-free, can be better discriminated on the basis of the data in the original sensor space than on the transformed data in the source space. Similarly, in EEG epochs containing epileptiform discharges, the discriminative ability of network topological indices was significantly better in the sensor compared to the source level. In conclusion, causality networks constructed at the sensor and source level, for both simulated and empirical data, exhibit significant structural differences. These observations indicate that further studies are warranted in order to clarify the exact relationship between data registered in the sensor and source space.

Keywords: multi-channel EEG analysis, sensor space analysis, source space analysis, brain networks, Granger causality, sLORETA

OPEN ACCESS

Edited by:
Luca Faes,
University of Palermo, Italy

Reviewed by:
Guido Nolte,
University Medical Center Hamburg-Eppendorf, Germany
Sebastiano Stramaglia,
University of Bari Aldo Moro, Italy

***Correspondence:**
Dimitris Kugiumtzis
dkugiu@auth.gr

Specialty section:
This article was submitted to Information Theory, a section of the journal Frontiers in Network Physiology

Received: *07 May 2021*
Accepted: *09 September 2021*
Published: *29 September 2021*

Citation:
Koutlis C, Kimiskidis VK and Kugiumtzis D (2021) Comparison of Causality Network Estimation in the Sensor and Source Space: Simulation and Application on EEG. Front. Netw. Physiol. 1:706487. doi: 10.3389/fnetp.2021.706487

1 INTRODUCTION

There is an increasing interest in neuroscience in working on the true current distribution of the sources in the grey matter of the brain, termed source space, and not on the extracranial electroencephalograms (EEG) or magnetic encephalogram (MEG) recordings, termed sensor space (De Vico Fallani et al., 2010; Lehmann et al., 2014). Many strategies have been proposed in order to obtain the primary current distribution of the sources in the brain from EEG or MEG. These are basically head models that take into account the volume conduction of the brain and other

properties in order to estimate the activity of the initial sources that explain best the extracranial electric potential measurements. Alternatively, it has also been proposed to reconstruct the electrostatic field within a predefined cubic grid of center points, based on the assumption of an ellipsoid and electromagnetically homogeneous head model, without using dipole modeling or other priors, an approach called 3D vector field tomography (Papadaniil and Hadjileontiadis, 2015; Papadaniil et al., 2016). The methods of source reconstruction are separated in two main classes, the over-determined and the under-determined inverse models (Michel et al., 2004). The over-determined inverse models presuppose that a small number of points in the source space is capable of explaining the extracranial measurements sufficiently. In this case, a unique solution, in terms of source location and current activity, is provided when the number of parameters to be estimated is less or equal to the number of sensor space channels. On the contrary, the under-determined inverse models consider a dense three dimensional (3D) grid of points in the brain having fixed positions and being many more than the sensor space channels, which leads to infinite solutions, as stated first in (Helmholtz, 1853). The objective for these models then is to determine a unique optimal source electric activity distribution over the grid of points.

In recent years, the concepts of Granger causality (Granger, 1969) and causality networks are of increasing interest in many branches of science such as finance (Hong et al., 2009; Billio et al., 2012), socioeconomics (Bollen et al., 2011), climatology (Runge et al., 2015; Dijkstra et al., 2019) and neuroscience (Kugiumtzis and Kimiskidis, 2015; Porta and Faes, 2016; He et al., 2019; Rossini et al., 2019). Intuitively, given two variables, X_1 and X_2, the Granger causality from variable X_1 to variable X_2 exists if the past and present of X_1 provide information that improves forecasts for the future of X_2. The concept of Granger causality is implemented in a number of measures in the time, frequency and phase domain, and extended to account for the presence of other observed variables and thus estimate only direct causal effects. In several studies, linear and nonlinear as well as bivariate and multivariate causality measures are compared, e.g., see the recent comparative studies in (Bakhshayesh et al., 2019; Siggiridou et al., 2019; Sommariva et al., 2019) and references therein, and it turns out that the most appropriate measures are the direct and nonlinear measures, but these are also the harder to be estimated reliably. The dimension reduction, intrinsically taking place in the algorithm of some Granger causality measures, allows for the estimation of direct causal effects in high-dimensional time series, such as multi-channel EEG. We consider a linear and a nonlinear such measure in this study, the restricted conditional Granger causality index (RCGCI) (Siggiridou and Kugiumtzis, 2016) and the partial mutual information from mixed embedding (PMIME) (Kugiumtzis, 2013), respectively, both found to perform well in high-dimensional settings (Koutlis and Kugiumtzis, 2016; Siggiridou et al., 2019).

The direct measures are particularly relevant for the formation of causality networks from multivariate time series because only direct causal effects are estimated. The causality networks are graphs with nodes that represent the observed variables of a system and connections between nodes that are weighted by the values of the selected causality measure. These networks are usually analyzed with simple metrics of network theory in order to estimate important characteristics of their topology (Rubinov and Sporns, 2010; Fornito et al., 2016; Geier and Lehnertz, 2017). In neuroscience, the estimation of causality effects among brain areas using EEG recordings is a widely used approach for the observation of brain reactions to certain stimuli and for the discrimination of normal and abnormal states of brain function (Schelter et al., 2006; Lehnertz et al., 2014; Fornito et al., 2016).

For the computation of the connectivity in the source space, in some cases the cortical activity may be estimated for a limited number of nodes and that makes the computations cost-efficient (De Vico Fallani et al., 2010), while in other cases, the number of nodes is prohibitive for a multivariate analysis of the estimated cortical activity (Milz et al., 2014). In the latter cases, the usual strategy is to discretize the cortex in different regions of interest (ROIs) and to consider as new nodes the centers of these regions and assign to them the average activity of all nodes of this region (Hata et al., 2016; Toppi et al., 2016). The connectivity of the estimated cortical activity is estimated using the measures of Granger causality discussed above [see also (Lei et al., 2015)], as well as linear measures (Schoffelen and Gross, 2009; De Vico Fallani et al., 2010; Lehmann et al., 2014; Milz et al., 2014) and nonlinear measures (Mulert et al., 2011; Pascual-Marqui et al., 2014) of similarity and synchronization, i.e. phase synchronization or correlation in the time and frequency domain. Recent studies found that the choice of the inverse method impacts the brain connectivity analysis, e.g., see source-space coherence analysis on magnetoencephalogram (MEG) (Hincapié et al., 2017).

To the best of our knowledge there have not been any reports for systematic comparison of the brain connectivity estimated in the sensor and source space. These two approaches are tested in simulation examples and discussed in (Brunner et al., 2016; Van de Steen et al., 2019). Furthermore, the effect of noise and source locations on the estimation of connectivity in source space is studied in (Anzolin et al., 2019). It is well accepted that EEG sensors do not capture signals exclusively from areas beneath EEG electrodes and the location and orientation of the sources influence critically the signals registered at the sensor level (Van de Steen et al., 2019). For that very reason, in the present study, we examine the overall connectivity and not distinguish one-to-one relationship between connections of individual anatomic areas in the sensor and source space, as done in the two studies in (Brunner et al., 2016; Van de Steen et al., 2019). Our study differs from the two studies also in that it assumes the simulated systems in the sensor rather than the source space, where the latter is physiologically more appropriate as sensors do not interact with each other but brain sources do. However, the data at hand are the scalp EEG (sensor space) and it is thus reasonable to assume simulation systems for the observed dynamics. Moreover, the study is on the overall structure of causality networks and not causal relationships of distinct sources as in (Anzolin et al., 2019; Van de Steen et al., 2019), and for this to develop a high dimensional system in the source space under

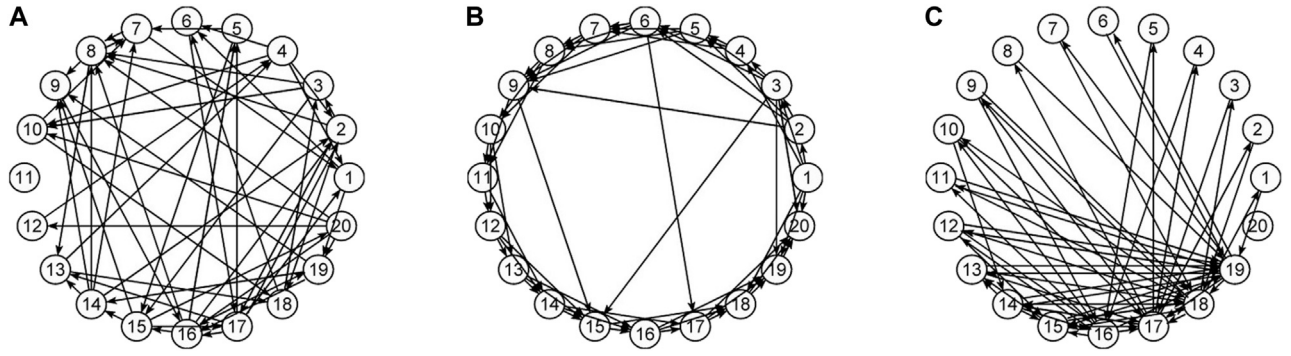


FIGURE 1 | Illustration of network types: **(A)** random, **(B)** small world, and **(C)** scale-free.

realistic assumptions is a hard task. For the comparison of causality networks in sensor and source space it was reported strong correlation of the global functional connectivity between the two domains on real scalp EEG of eyes open and closed (Lai et al., 2018). We extend this study using effective connectivity measures and a different real EEG setting of changing connectivity structure.

In this work, the objective is beyond investigating differences in the estimated connectivity networks in the sensor and source space, which is expected due to the transform and we show it here analytically. Rather, the objective is to apply the same procedures for causality network estimation in the sensor and source space and to compare the causality structures that arise on each of the two workspaces. Though it is not expected to find similar causality structures, an important question is whether the estimated causality structures hold the same discriminative information in any of their characteristics, estimated by network metrics. For this, we performed a simulation study on systems of linear and nonlinear dynamics, having a predefined connectivity structure of one of the three types of random (RAND) (Erdös and Rényi, 1959), scale-free (SCF) (Barabási and Albert, 1999) and small-world (SW) (Watts and Strogatz, 1998) networks. Examples of the network types are presented in **Figure 1**. The original space of the generated multivariate time series is assumed to be the sensor space and the data are transformed by sLORETA to source space. The objective of the simulation study is the discrimination of different coupling structures (the three network types) by a network characteristic in both spaces. As established in our previous works, the two causality measures RCGCI and PMIME have a very good discrimination capability in the sensor space (the generated time series in the simulation study) and the interest is merely to investigate the deformation of the network structure in the source space. For the neuroscience data analysis, EEG data with episodes of epileptiform discharges (ED) are used in order to compare the causality structure of the brain before, during and after an ED in the sensor and source space.

The paper is organized as follows: in **Section 2** the source localization method sLORETA, the causality measures and the network indices are presented, in **Section 3** the simulated and the EEG data are briefly discussed, in **Section 4** the results are presented and in **Section 5** conclusions are given.

2 METHODS

2.1 sLORETA

A commonly used method for the 3D localization of the sources of the electromagnetic activity of the brain is the so-called low resolution electromagnetic tomography, implemented in the LORETA software (Pascual-Marqui et al., 1994). LORETA is suggested as an improvement of the minimum norm estimate (Hämäläinen and Ilmoniemi, 1994). The minimum norm estimate has a main disadvantage in that superficial sources are preferred over deep sources due to the fact that the current density vector is not weighted for the solution of the minimization problem. LORETA overcomes this problem using as weights the norm of the columns of the lead field matrix (the matrix that transforms current density of the sources to extracranial measurements), thus giving truly 3D solutions. Additionally, it provides "smooth" solutions in terms of the minimum weighted Laplacian, which are more plausible as neighboring neurons tend to have coherent waveforms. The way to achieve a meaningful 3D unique solution is by sacrificing spatial resolution, namely the brain volume is discretized in $v = 2394$ voxels considered as separate dipoles and thus the name low resolution electromagnetic tomography. The forward model is defined as:

$$\mathbf{m} = \mathbf{G}\mathbf{d} \tag{1}$$

where $\mathbf{m} \in \mathbb{R}^c$ is the vector of measurements at a certain time point, $c \in \mathbb{N}$ is the number of channels, $\mathbf{d} \in \mathbb{R}^{3v}$ is the vector of current densities of the v voxels (for each voxel a 3D vector is considered) and $\mathbf{G} \in \mathbb{R}^{c \times 3v}$ is the lead field matrix. The discrete problem is:

$$\min_{\mathbf{d}} \|\mathbf{B}\mathbf{W}\mathbf{d}\|, \text{ under the constraint } \mathbf{m} = \mathbf{G}\mathbf{d} \tag{2}$$

where $\mathbf{W} \in \mathbb{R}^{3v \times 3v}$ is a diagonal matrix with $\mathbf{W}_{ii} = \|\mathbf{G}_i\|$, $\mathbf{G}_i$ is the ith column of $\mathbf{G}$ and $\mathbf{B}$ is the discrete Laplacian operator $3v \times 3v$

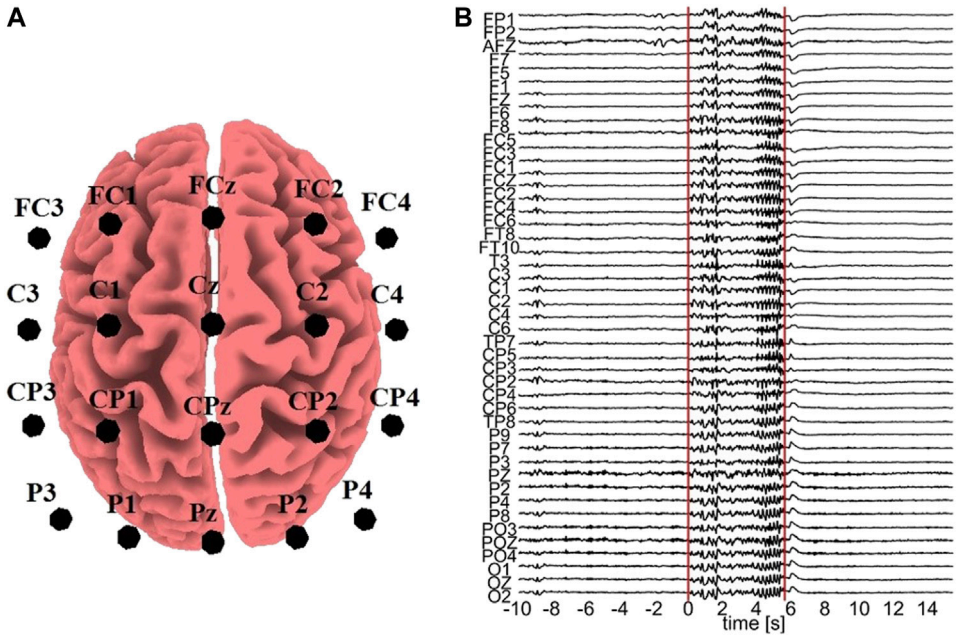


FIGURE 2 | (A) Electrode positions assumed for the sLORETA on the simulated data. **(B)** One epoch with an ED episode. The vertical lines represent the start and end points of the ED, respectively, and the time order is with respect to the ED start (0 time at start).

matrix [for more details on the computation of **B** see the Appendix of (Pascual-Marqui et al., 1994)]. The unique solution for the current densities is:

$$\mathbf{d}^* = \mathbf{Tm} \tag{3}$$

with $\mathbf{T}=(\mathbf{WB}^T\mathbf{BW})^{-1}\mathbf{G}^T\ [\mathbf{G}(\mathbf{WB}^T\mathbf{BW})^{-1}\mathbf{G}^T]^\dagger$, where the † exponent denotes the Moore-Penrose pseudo-inverse.

In (Pascual-Marqui, 2002) an improvement of LORETA is presented suggesting that it achieves zero localization error due to a standardization of the current densities. This improvement is termed standardized low resolution brain electromagnetic tomography (sLORETA) and provides a better resolution with $v = 6239$ voxels. The followed procedure is exactly the same as in LORETA but the estimation $\mathbf{d}^*$ is finally standardized by its variance. This is actually similar to the work conducted in (Dale et al., 2000) utilizing standardization as well but in a different way from that in sLORETA, which conversely achieves exact localization. More precisely, the covariance matrix $S_{\mathbf{d}^*}$ of the estimated current density $\mathbf{d}^*$ is computed and the standardized current densities are:

$$\mathbf{d}_i^* T\left\{[S_{\mathbf{d}^*}]_{ii}\right\}^{-1}\mathbf{d}_i^* \tag{4}$$

where $\mathbf{d}_i^* \in \mathbb{R}^3$ is the current density estimate at the *ith* voxel and $[S_{\mathbf{d}^*}]_{ii} \in \mathbb{R}^{3\times3}$ is the *ith* diagonal block of matrix $S_{\mathbf{d}^*}$.

2.2 Causality Network Estimation

For the estimation of the causality networks we use two Granger causality measures, one linear, and one nonlinear, and both in the time domain making use of dimension reduction and being capable of estimating direct causal effects from high-dimensional time series (Koutlis and Kugiumtzis, 2016; Siggiridou et al., 2019).

2.2.1 RCGCI

The measure restricted conditional Granger causality index (RCGCI) is the well-known linear stochastic model-based measure of conditional Granger causality index (CGCI) but based on a restricted (sparse) vector autoregressive (VAR) model (Siggiridou and Kugiumtzis, 2016). The CGCI makes use of the unrestricted VAR model (U-model) of all the lagged variables up to a maximum lag (order) p so that in total there are Kp terms, where K is the number of observed variables. Applying a selection scheme using augmented VAR models, called modified backward-in-time selection, a small subset of lagged variables are selected in the sparse VAR model to constitute the U-model for RCGCI. For high-dimensional systems with sparse coupling structure, the selected subset may have cardinality much smaller than Kp. The restricted model (R-model) is derived by excluding the lagged terms of the driving variable X, so that if the sparse VAR U-model contains no lagged terms of X then RCGCI = 0. If it contains lagged terms of X then RCGCI is computed by the log ratio of the fitted error variances of the U-model and R-model, ${s_U}^2$ and ${s_R}^2$, respectively, as for the CGCI

$$\mathrm{RCGCI}_{X\to Y} = \log\frac{{s_R}^2}{{s_U}^2}. \tag{5}$$

2.2.2 PMIME

The measure partial mutual information from mixed embedding (PMIME) is a normalized version of the partial transfer entropy

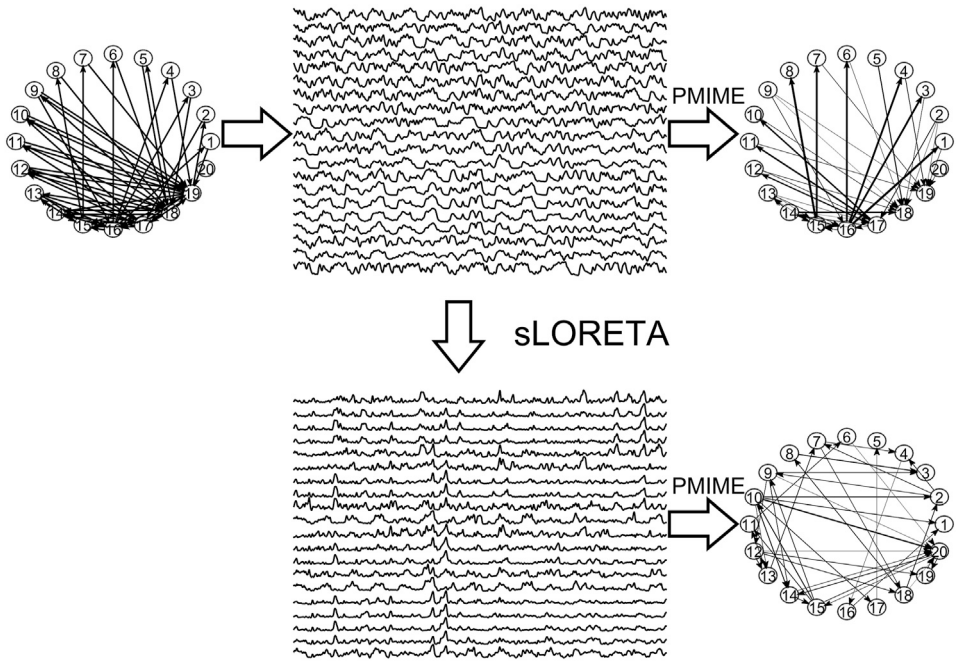


FIGURE 3 | The initial scale-free coupling structure among the variables is used to generate the $K = 20$ time series from the CMG system hypothesized to be in the sensor space. Then, sLORETA derives the respective time series in the source space and finally the PMIME is used to estimate the causal effects and give the causality networks in both the sensor and the source levels.

(Schreiber, 2000; Papana et al., 2012) restricted to the selected mixed embedding vector, formed by selecting progressively from the set of all delayed variables the most relevant ones to the future of the response variable (Kugiumtzis, 2013). The selection relies on information measures, mutual information (MI) and conditional mutual information (CMI). In particular, in the first step of the selection scheme, the lagged variable w with the highest mutual information (MI) to the response Y (one time step ahead, y_{t+1}), $I(y_{t+1}; w)$, is selected and constitutes the current subset of lagged variables $\mathbf{w}_t$. For each subsequent step, the lagged variable w with the highest CMI to the response given the current subset, $I(y_{t+1}; w|\mathbf{w}_t)$, is selected and added to $\mathbf{w}_t$. The significance of CMI (MI for the first step) is tested at each step and the selection scheme terminates when it is found statistically not significant. Similarly to RCGCI, if the derived subset does not include any lagged terms of the driving variable X then PMIME = 0. If it includes, the information of the lagged terms of X about the response is evaluated by the CMI of the lagged terms of X in $\mathbf{w}_t$, $\mathbf{w}_t^x$, and the response given all other components in $\mathbf{w}_t$, normalized by the mutual information of all the lagged terms and the response

$$\mathrm{PMIME}_{X\to Y} = \frac{I(y_{t+1}; \mathbf{w}_t^x | \mathbf{w}_t^y, \mathbf{w}_t^z)}{I(y_{t+1}; \mathbf{w}_t)}, \tag{6}$$

where $\mathbf{w}_t^y$ and $\mathbf{w}_t^z$ are the lagged terms of the response Y and the remaining variables Z in $\mathbf{w}_t$, respectively, so that $\mathbf{w}_t = [\mathbf{w}_t^x, \mathbf{w}_t^y, \mathbf{w}_t^z]$.

Both causality measures include an internal criterion to determine the presence of the driving of one variable to another and in that case a positive value of RCGCI or PMIME quantifies the strength of the driving, otherwise the measure is zero. This leads to a weighted matrix with zero and weighted connections of driving strength, thus the binarization of the causality network is performed just by accepting the positive values as existing causal effects.

2.3 Network Metrics

We select five network metrics, which capture different characteristics of the network structure (Rubinov and Sporns, 2010). These metrics are computed on the adjacency matrices formed by the positive PMIME or RCGCI values for each ordered pair of variables and regard the causality network estimated from the multivariate time series.

The five network metrics quantify different properties of the network with binary directed or non-directed connections. The centrality property of the network is quantified by the mean, deg^m, and standard deviation (SD), deg^{SD}, of the degree distribution over the nodes. The characteristic path length, λ, is used as index of functional integration. The small-worldness index, SWi, quantifies the presence of small-world structure in the network, which suggests the combination of functional segregation and functional integration in a network. As an index of resilience we consider the assortativity coefficient, r, quantifying the preference of network nodes to attach to other nodes of similar degree, typically defined as the correlation coefficient between the degrees of two nodes.

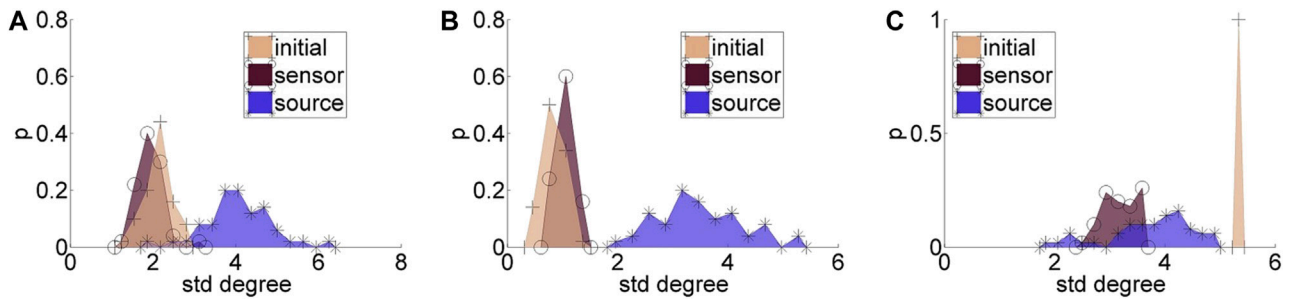


FIGURE 4 | Histograms for the standard deviation of the degree, deg^{SD}, for 50 realizations of each network type: **(A)** random (RAND), **(B)** small-world (SW), and **(C)** scale-free (SCF). At each panel, there are three histograms, for the initial network, the network estimated in the sensor space, and the network estimated in the source space, as given in the legend.

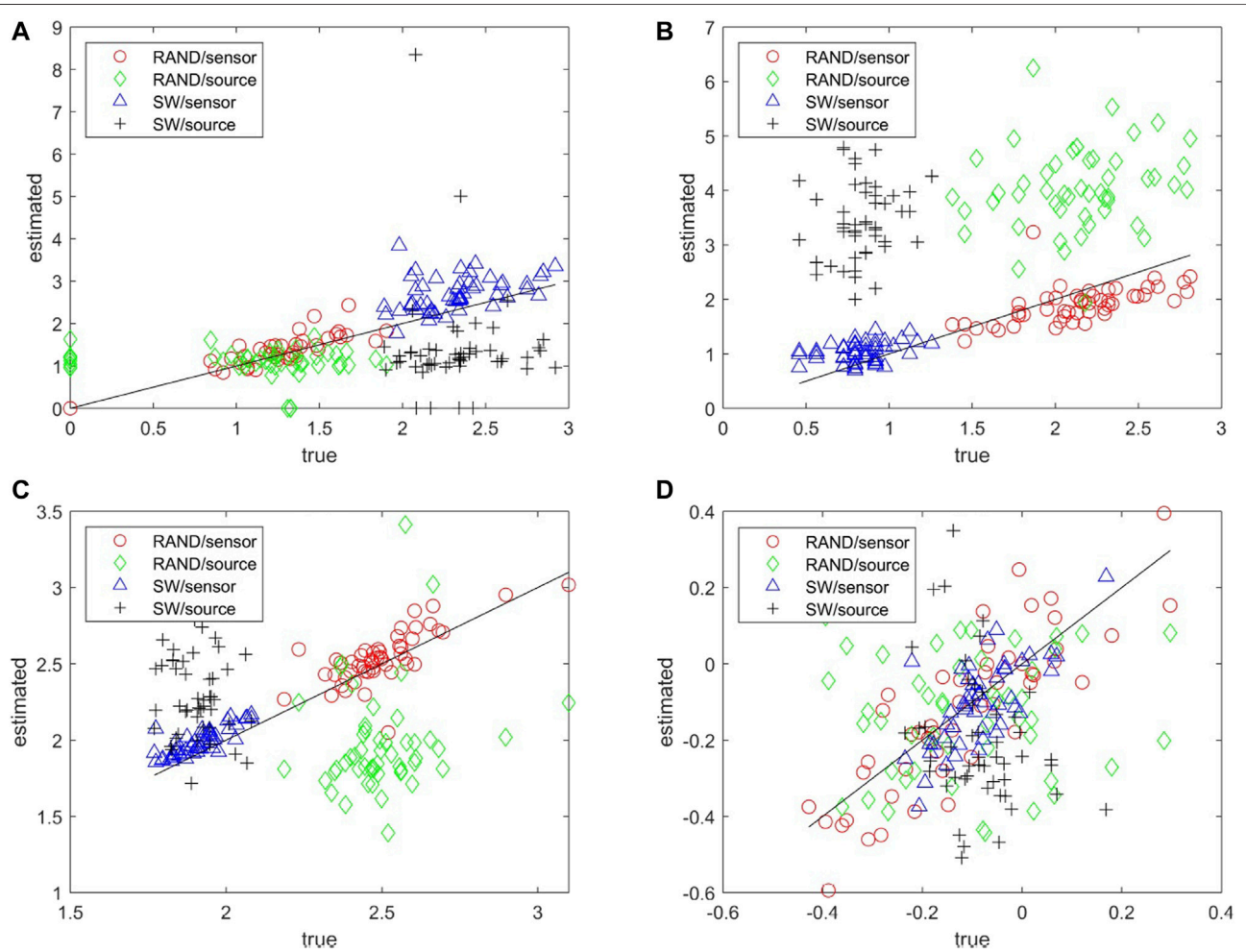


FIGURE 5 | Scatter plots of network metrics estimated on 50 realizations of random (RAND) and small-world (SW) network types, where in the *x*-axis is the metric for the true network and in the *y*-axis for the estimated network in the sensor and source space, as shown in the legend. The metrics are *SWi* in **(A)**, deg^{SD} in **(B)**, λ in **(C)** and *r* in **(D)**.

2.4 Evaluation Index

In the focus of the study is the discrimination of different connectivity structures in the sensor and source space, and we consider the three network types, i.e., random (RAND), small world (SW), and scale-free (SCF) (see **Figure 1**). We consider a number of realizations of multivariate time series from each network

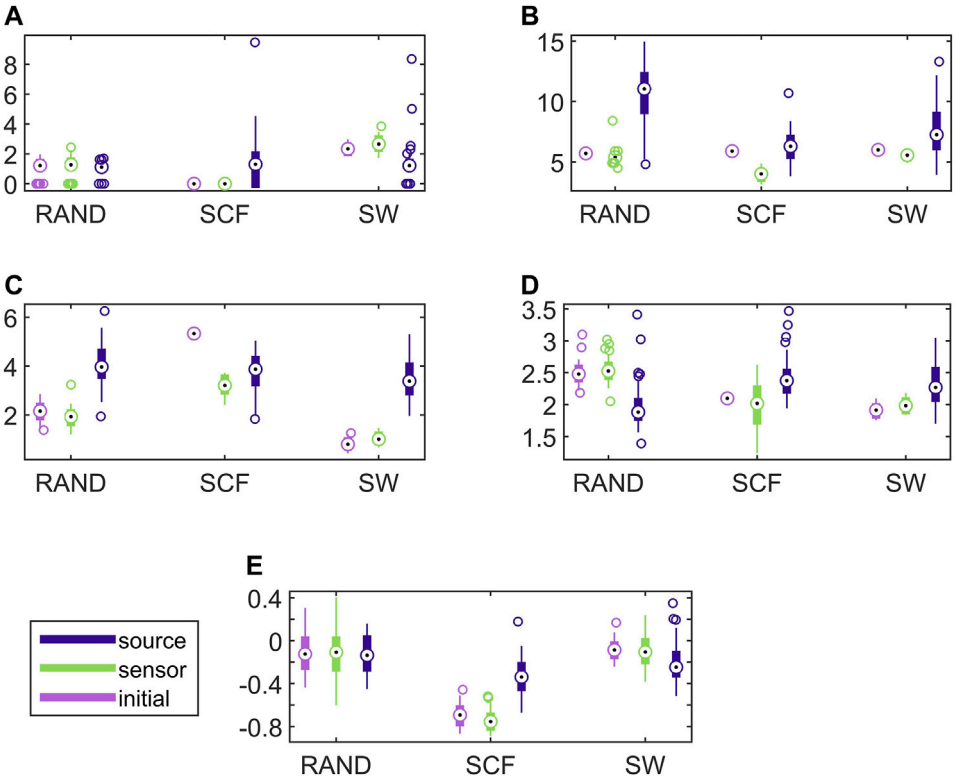


FIGURE 6 | Distribution of network metric values for three different network types (RAND, SW, and SCF) computed on the initial network and on the networks estimated by PMIME on the sensor space and on the source space, as shown in the legend. The PMIME is computed on simulated time series of the CMG system. The metrics are *SWi* in **(A)**, deg^m in **(B)**, deg^{SD} in **(C)**, λ in **(D)** and r in **(E)**.

type and compute the causality networks and subsequently the network metrics for each realization. For each pair formed from any two of RAND, SW, and SCF, the evaluation index area under receiver operating characteristic curve (AUROC) (Fawcett, 2006) is employed to quantify the overlapping of the two distributions of the network metric in the two network types. For the simulation study, the AUROC is computed for each pair of network types of the original causality network (given by the system equations) and the estimated causality networks in the sensor and source space.

3 MATERIALS

In the simulation study, the synthetic systems are considered to evolve in the sensor space. One could argue to consider the source space instead, but we do not follow this line here for two reasons. First, the assumption of the sensor space as the space of the generated time series is a natural choice since the starting data, the acquired EEG measurements, are in the sensor space. Secondly, the generation of data in the source space is a hard task and can be realized under rather unrealistic simplifications, e.g., assuming the same activity in all vertices (voxels) in a ROI on the cortical mesh (Silva Pereira et al., 2017). We consider three systems of different type. The two first systems are deterministic chaotic dynamical systems while the third is a linear multivariate stochastic process.

TABLE 1 | Average AUROC values from the three AUROC values computed in the three binary classification tasks, RAND vs. SW, RAND vs. SCF, and SW vs. SCF, for the five network metrics (labels in the first row), the three dynamical systems (labels in the first column) and for the initial networks, the networks from the sensor space and the networks from the source space.

		SWi	deg^m	deg^{SD}	λ	r
CMG	initial	0.999	1	1	0.995	0.856
	sensor	0.997	0.878	0.995	0.828	0.842
	source	0.546	0.839	0.633	0.768	0.713
CHM	initial	1	1	1	0.993	0.866
	sensor	0.99	0.957	0.988	0.924	0.713
	source	0.523	0.781	0.706	0.678	0.613
VAR	initial	0.999	1	1	0.993	0.888
	sensor	0.951	0.953	0.993	0.949	0.729
	source	0.739	0.841	0.855	0.851	0.559

The two first systems are coupled systems, i.e., are formed from a number of subsystems of the same type coupled to each other. The first system is in discrete time and the second in continuous time.

3.1 Simulated Data

The system of coupled Hénon maps (CHM) (Kugiumtzis, 2013) is a system of coupled chaotic maps defined as

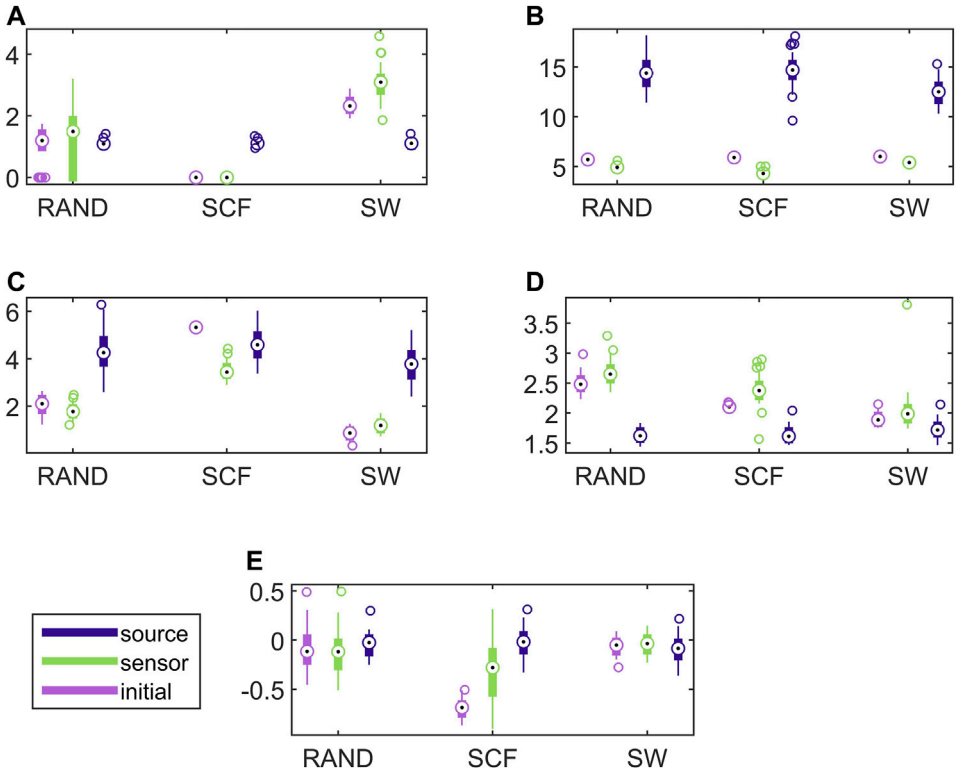


FIGURE 7 | Distribution of network metric values for three different network types (RAND, SW, and SCF) computed on the initial network and on the networks estimated by PMIME on the sensor space and on the source space, as shown in the legend. The PMIME is computed on simulated time series of the CHM system. The metrics are *SWi* in **(A)**, deg^{m} in **(B)**, deg^{SD} in **(C)**, λ in **(D)** and r in **(E)**.

$$x_{j,t} = 1.4 - \left[\frac{\sum_{i=1}^{K} C_{ij} x_{i,t-1}}{\sum_{i=1}^{K} \Theta(C_{ij})} + \left(1 - \frac{\sum_{i=1}^{K} C_{ij}}{\sum_{i=1}^{K} \Theta(C_{ij})}\right) x_{j,t-1}\right]^2 + 0.3 x_{j,t-2} \tag{7}$$

where $j = 1, 2, \ldots, K$ is the variable index, K denotes the number of variables and C_{ij} is the coupling strength (considering x_i as the driving variable and x_j as the response variable). The $\Theta(C_{ij})$ is the Heaviside function, being one if $C_{ij} > 0$ and zero if $C_{ij} = 0$. For this system, we consider $K = 20$ variables coupled with strength $C_{ij} = C = 0.2$ for the non-vanishing terms according to each of the three network types (RAND, SW, SCF). For each network type, 50 realizations are generated and the time series have length $N = 2048$.

The system of coupled Mackey-Glass (CMG) (Senthilkumar et al., 2008; Kugiumtzis, 2013) is a system of coupled identical delayed differential equations defined as

$$\dot{x}_j(t) = -0.1 x_j(t) + \sum_{i=1}^{K} \frac{C_{ij} x_i(t-\Delta)}{1 + x_i(t-\Delta)^{10}}, \tag{8}$$

where $j = 1, 2, \ldots, K$ is the variable index, and Δ is the delay parameter. For this system we consider $\Delta = 100$ time units (for the uncoupled system each equation defines a chaotic system of correlation dimension close to 7) and $C_{ij} = C = 0.1$. The solution is run in Matlab (Natick, Massachusetts: The MathWorks Inc.) using the function "dde23" and then sampled at 4 time units. Further details for the generation of the CMG time series are given in (Kugiumtzis, 2013). The setting is as for CHM, $K = 20$, 50 realizations of each of the RAND, SW, and SCF network types and $N = 2048$.

The third systems is a linear stochastic VAR process (Basu and Michailidis, 2015) on $K = 20$ variables and order $p = 1$,

$$\mathbf{x}_t = A\mathbf{x}_{t-1} + \mathbf{e}_t, \tag{9}$$

where $\mathbf{x}_t$ is the state vector of length K and $\mathbf{e}_t$ is the white noise vector of length K following Gaussian distribution with zero mean and unity covariance matrix. The components of the square coefficient matrix A are zero or positive determined by the selected network of type RAND, SW, or SCF. Initially, the non-vanishing coefficients are set to 0.9 (the rest are set to zero) and then they are iteratively reduced until the stationarity condition is fulfilled. We get 50 multivariate time series for each of the three network types of length $N = 2048$.

The PMIME is used to estimate the connectivity for the two first nonlinear systems and the RCGCI for the third linear system, as being the most appropriate measures to capture the causality effects in each system. This is so because the objective here is to have an appropriate measure that estimates best the causality network from the time series in the sensor and source space.

All the computationally generated data are treated as if they were EEG signals (in the sensor space) in order to transform them with sLORETA (to the source space) and the $K = 20$ variables are

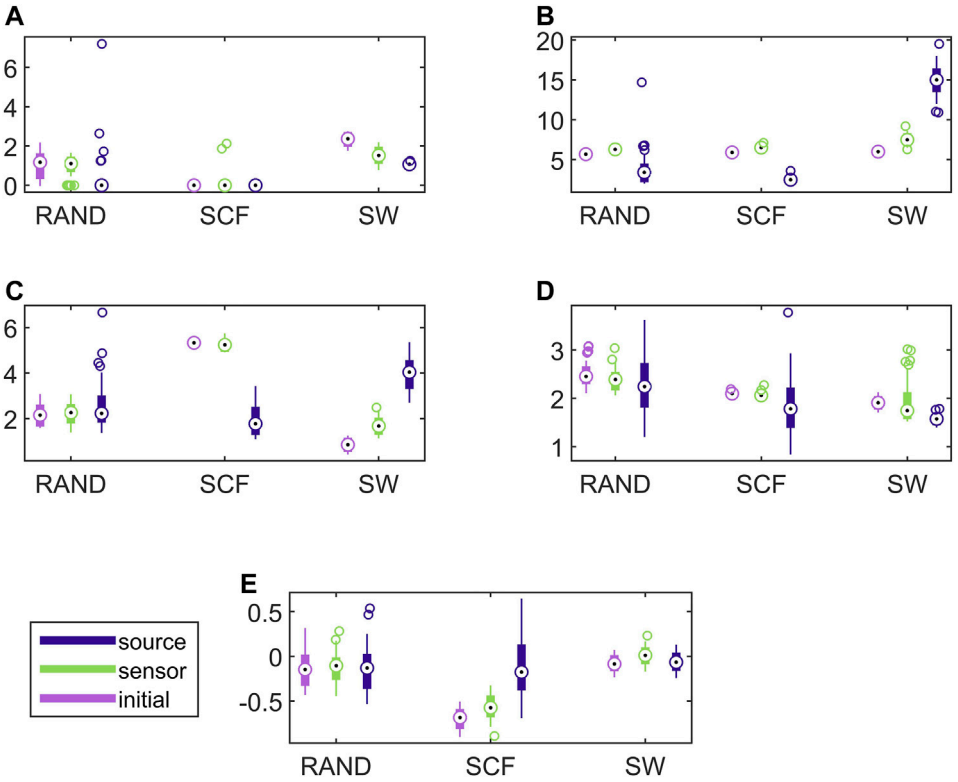


FIGURE 8 | Distribution of network metric values for three different network types (RAND, SW, and SCF) computed on the initial network and on the networks estimated by PMIME on the sensor space and on the source space, as shown in the legend. The PMIME is computed on simulated time series of the VAR system. The metrics are *SWi* in **(A)**, deg^{in} in **(B)**, deg^{SD} in **(C)**, λ in **(D)** and r in **(E)**.

considered as electrode positions on the scalp, more precisely: FC3, FC1, FCz, FC2, FC4, C3, C1, Cz, C2, C4, CP3, CP1, CPz, CP2, CP4, P3, P1, Pz, P2, P4. These positions are selected in order to cover a sufficient area of the brain rather than concentrating on a small region, as shown in **Figure 2A**.

For each of the 50 realizations, the hypothesized multi-channel ($K = 20$) time series in the sensor space is transformed with sLORETA to $v = 6239$ source space signals, which consequently are separated in ROIs and averaged, thereby resulting in $K = 20$ source space time series. The ROIs are selected with the option of "all nearest voxels" in the sLORETA software, which separates the voxels in mutually disjoint regions of the nearest to the selected electrodes voxels.

3.2 EEG Data

For the real data analysis, EEG data from three epileptic patients at rest are used. In order to avoid the confounding variables of diverse structural substrates and varying localizations of the epileptic zone, which typically characterize focal epilepsies, we used data from patients suffering from Genetic Generalized Epilepsies. From the initial recording of 60 channels, channels containing artifacts were rejected, i.e., 16 channels for subject 1, 9 channels for subject 2, and 14 channels for subject 3. The EEG data of the artifact-free channels were band-pass filtered in (0.01,70) Hz, downsampled to 200 Hz, and re-referenced to infinity (Yao, 2001), a re-referencing scheme found to be more appropriate for connectivity analysis (Qin et al., 2010). Epochs containing epileptiform discharge (ED) were selected, 7 epochs from subject 1, 3 epochs from subject 2 and 10 epochs from subject 3. An exemplary epoch from subject 1 is shown in **Figure 2B**. Each epoch contains a pre-ED period of 10 s, a spontaneous ED of duration from 2.19 to 5.71 s and a post-ED period of 10 s.

The sLORETA software is used for the transformation of the EEG signals to the source space. First, the electrode positions are selected and then the transformation matrix is generated by the software. Consequently, the signals are transformed to the source space giving a large number of $v = 6239$ time series. In accordance with the simulation study, a number of 44 ROIs corresponding to the 44 scalp electrodes for subject 1 are selected and 44 time series for the source space are obtained, and the same was done accordingly for the artifact-free channels of the other two subjects.

After the transformation of the time series, both the sensor and source time series are split in sliding overlapping windows of 2 s with a sliding step of 0.5 s to profile the brain network characteristics and all their changes during the epochs. The causality estimation is performed with PMIME and then characteristics of the causality networks are estimated by the network metrics previously presented.

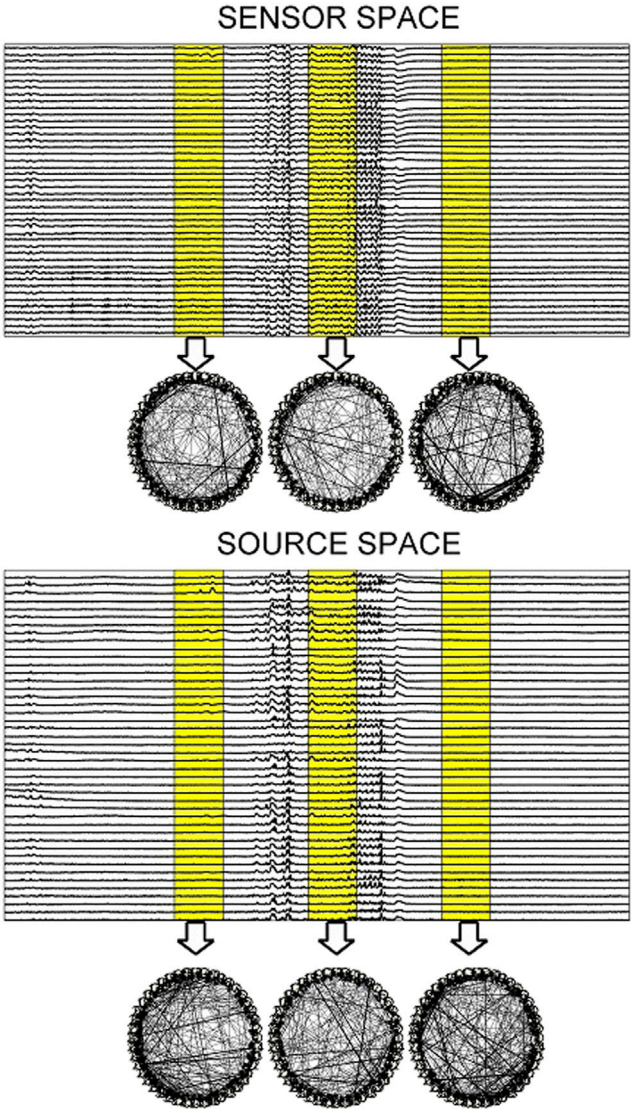


FIGURE 9 | Upper panel: One epoch of S1 comprising a 10 s pre-ED period, an ED lasting 5.71 s and a 10 s post-ED period. Three time windows are highlighted (one in the pre-ED period, one during the ED and one in the post-ED period) and the respective causality networks are presented right under the arrows. Lower panel: the same but for the sLORETA transformed time series.

4 RESULTS

Before presenting the results of the simulation study, it is elaborated that a linear transformation may change the connectivity of a set of variables in a significant degree. Although the employed methodology (sLORETA) is not a linear transformation, the following simplistic example is indicative of the changes a transformation in general can induce. Let us assume a linear VAR(1) process on K variables

$$\mathbf{x}_t = A\mathbf{x}_{t-1} + \mathbf{e}_t, \tag{10}$$

where the coefficient matrix A defines the connectivity matrix of the original system, $\mathbf{x}_t = [x_{1,t}, \ldots, x_{K,t}]$ is the vector of the state of K variables at indexed time t and $\mathbf{e}_t = [e_{1,t}, \ldots, e_{K,t}]$ is the vector of noise terms at time t having normal distribution $e_{i,t} \sim N(0, \sigma^2)$, $i = 1, \ldots, K$, and covariance matrix $\Sigma_{\mathbf{e}_t} = \sigma^2 I_K$, where I_K is the identity matrix of size $K \times K$. Consider a linear transformation of $\mathbf{x}_t$

$$\mathbf{y}_t = H\mathbf{x}_t, \tag{11}$$

where H is a matrix of size $K \times K$ with $\det(H) \neq 0$. Solving **Eq. (11)** with respect to $\mathbf{x}_t$ and substituting it in **Eq. (10)**, the VAR (1) for $\mathbf{y}_t$ is obtained as:

$$\mathbf{y}_t = HAH^{-1}\mathbf{y}_{t-1} + H\mathbf{e}_t = B\mathbf{y}_{t-1} + \mathbf{e}'_t, \tag{12}$$

where the input noise vector $\mathbf{e}'_t$ is correlated (the noise covariance matrix is not diagonal) indicating also the presence of instantaneous causality, which however is not relevant here, as we estimate only the lag causality. In analogy to the connectivity of the original VAR(1) system, the connectivity of the transformed VAR(1) system is determined by the coefficient matrix $B = HAH^{-1}$. Thus, a variety of different connectivity structures for the transformed system can be obtained, dependent on the form of H, with the only constraint being that the initial and the transformed system have connectivity matrices with the same eigenvalues. Also, it is noteworthy that the transformation matrix $H = Q_A^{-1}$, where the columns of Q_A are the eigenvectors of A, leads to a diagonal B and subsequently a linearly transformed VAR(1) system with no connections among its variables, regardless of its initial connectivity structure.

4.1 Simulation Study

In **Figure 3**, the steps of the procedure followed in the simulation study are illustrated in an example for the CMG system. Initially, a network of a predefined type (RAND, SW, or SCF) is defined, in order to form the coupling relationships in the system equations, for the generation of the 50 simulated time series of each network type in the hypothesized sensor space. Then, sLORETA is used to transform the time series to the source space. The causal effects for all pairs of observed variables are estimated by a causality measure (PMIME for CMG in **Figure 3**) in both the sensor and the source space, resulting in the respective causality networks. It is observed that the form of the time series is rather different before and after the sLORETA transformation. More precisely, the source space time series are less oscillatory and show spikes at various times contrary to the initial time series.

Regarding the causality network estimation, for the example in **Figure 3** the PMIME has a satisfactory performance on the sensor space time series capturing the original scale-free structure. The causality network of the transformed time series in the source space is qualitatively different from the original network, as expected from the previous analytical example. This is quantitatively confirmed for instance by the network metric deg^{SD} that takes the values 5.33 and 2.35, respectively, for the initial structure and the source space estimation. In **Figure 4**, aggregate results from the 50 realizations are shown for all three network structures, i.e., the initial network, the network estimated

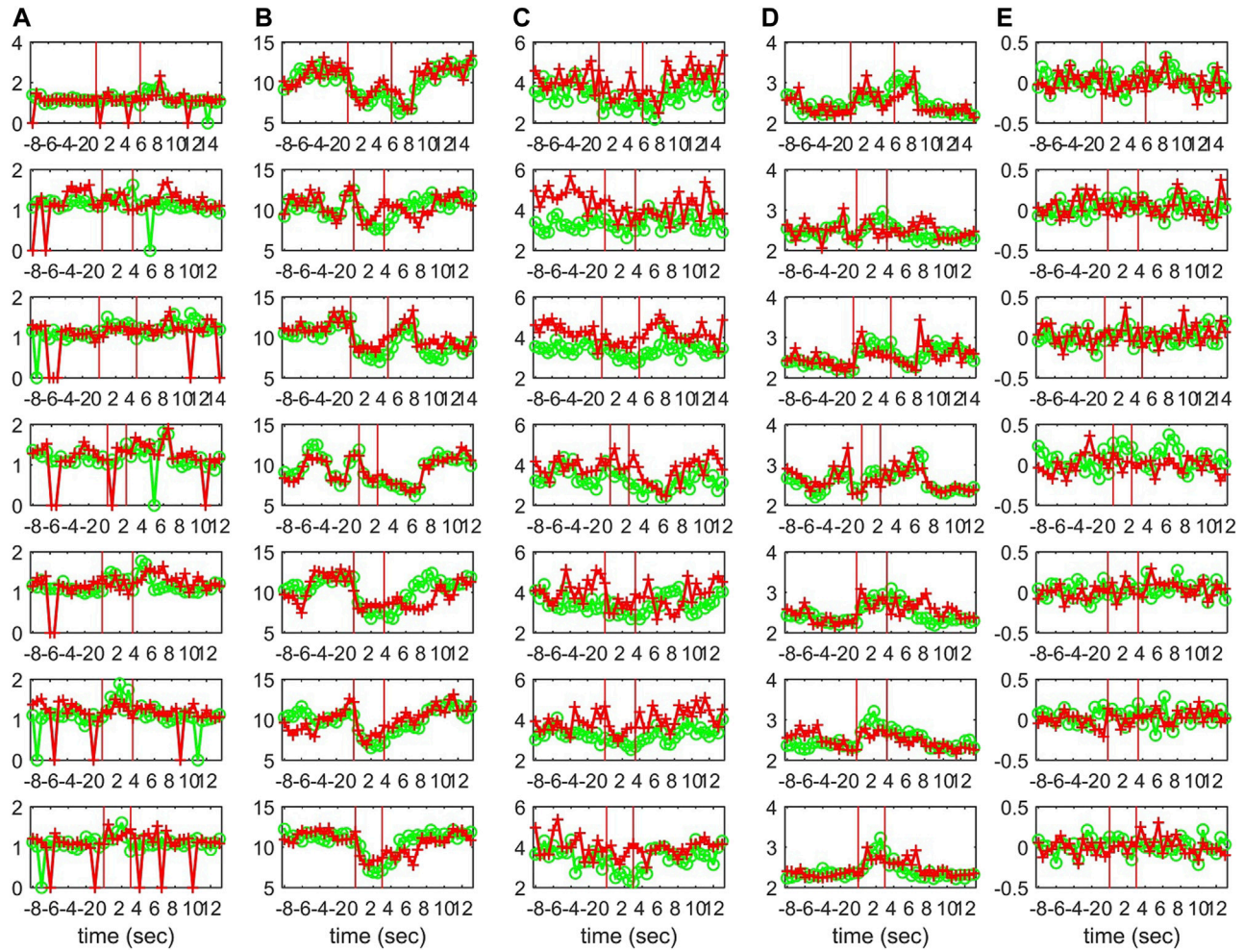


FIGURE 10 | The profile of five network indices on the sensor (green line with "o" marker) and on the source space (red line with "+" marker) for 7 epochs of S1 given in the row panels: **(A)** *SWi*, **(B)** deg^m, **(C)** deg^{SD}, **(D)** λ, **(E)** r.

in the sensor space and the network estimated in the source space. The similarity of the deg^{SD} between the original network and the estimated network in the sensor space as well as their dissimilarity to the estimated network in the source space are striking for the cases of RAND and SW networks. For SCF networks, the PMIME does not estimate with the same high accuracy the initial network (see specific results in (Koutlis and Kugiumtzis, 2016)) and therefore the histograms of deg^{SD} for the sensor and source space networks differ significantly from the respective histogram for the initial network.

The matching of the metrics of the estimated network to the true network can be seen collectively in the scatter plots for each network metric and for both RAND and SW network types in **Figure 5**. Results are not shown for the characteristic deg^m being constant for the true networks and the SCF network type as for all but the network metric r the metric of the true network is rather constant across the 50 realizations. The network metrics *SWi*, deg^{SD}, and λ computed on the sensor space tend to match well with the respective metrics of the true networks whereas for the source space the respective metrics are more spread and have the tendency to overestimate or underestimate the metrics of the true network (over or under the diagonal in **Figures 5A–C**, respectively). For the metric r in **Figure 5D** the values of the true networks are at the same range for RAND and SW network types so that at first look r values seem to spread about the same for both sensor and source space. A careful look would discern the alignment of the r values for the sensor space along the diagonal and indeed the Pearson correlation coefficient values are for RAND network type 0.84 for sensor space as opposed to 0.01 for source space and for SW network type 0.67 and −0.28, respectively.

Establishing that the connectivity changes significantly under the sLORETA transform, we investigate now whether the transform preserves the discriminative information that allows for correct classification of the different initial coupling structures. For this, we examine the discrimination of the three network types in the sensor and source space. In **Figure 6**, results for the discriminative ability of the network

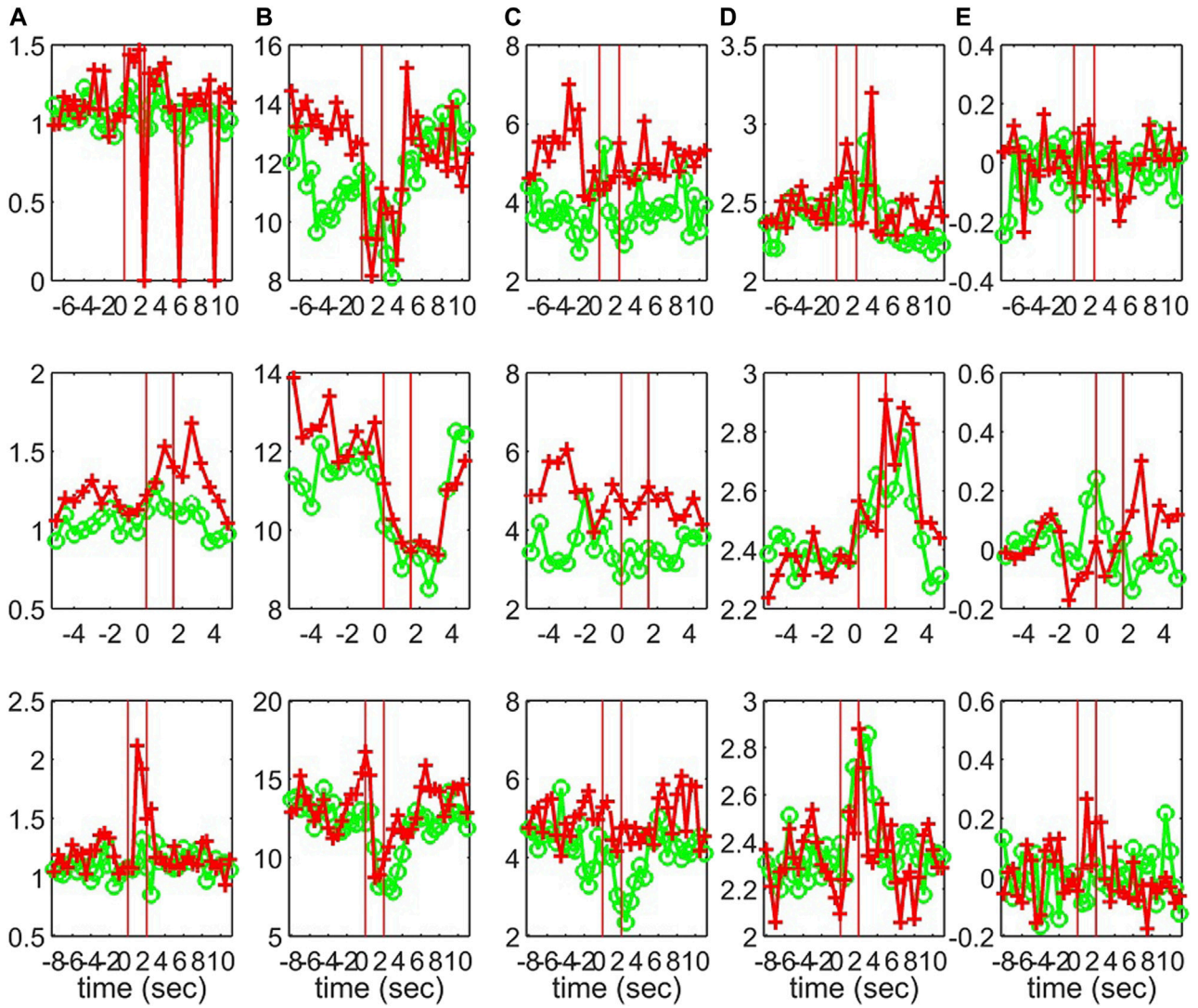


FIGURE 11 | The profile of five network indices on the sensor (green line with "o" marker) and on the source space (red line with "+" marker) for 3 epochs of S2 given in the row panels: **(A)** *SWi*, **(B)** deg^m, **(C)** deg^{SD}, **(D)** λ, **(E)** r.

metrics are presented for the CMG system, where the estimated networks are derived by the PMIME.

It is observed that the values of all network metrics for the initial network and the network estimated on the sensor space range in a similar scale. This includes also the case of deg^{SD}, where though the initial connections are not preserved well in the sensor space with the PMIME for SCF networks (see **Figure 4C**), still the SCF networks can be clearly discriminated with deg^{SD} from RAND and SW networks. The values of the network metrics for the source space are in some cases similar with those of the initial networks, e.g., *SWi* on RAND network, but in most cases they range in different scales. This result indicates that the transformation to the source space changes significantly the structure of the causality network. The discriminative information of the network metrics, regarding the three network types, remains in the sensor space estimation at a larger degree than in the source space. This result is confirmed by the average AUROC values for the differences in the three pairs of the RAND, SW, and SCF network types, presented in **Table 1**. While the AUROC values for the networks estimated in the sensor space are high and similar to these for the initial networks, the respective AUROC values for networks estimated in the source space are much lower. However, in almost all of the discrimination tasks of **Table 1** we obtain a p-value < 0.01 after ANOVA hypothesis testing for equal mean AUROC in the RAND, SW and SCF network types, with the exception of source space estimation of *SWi* on CMG, source space estimation of *SWi* on CHM, source space estimation of r on CHM and source space estimation of r on VAR.

In **Figure 7**, results for the discriminative ability of the network indices are presented for the CHM system in the same way as shown earlier for the CMG system in **Figure 6**.

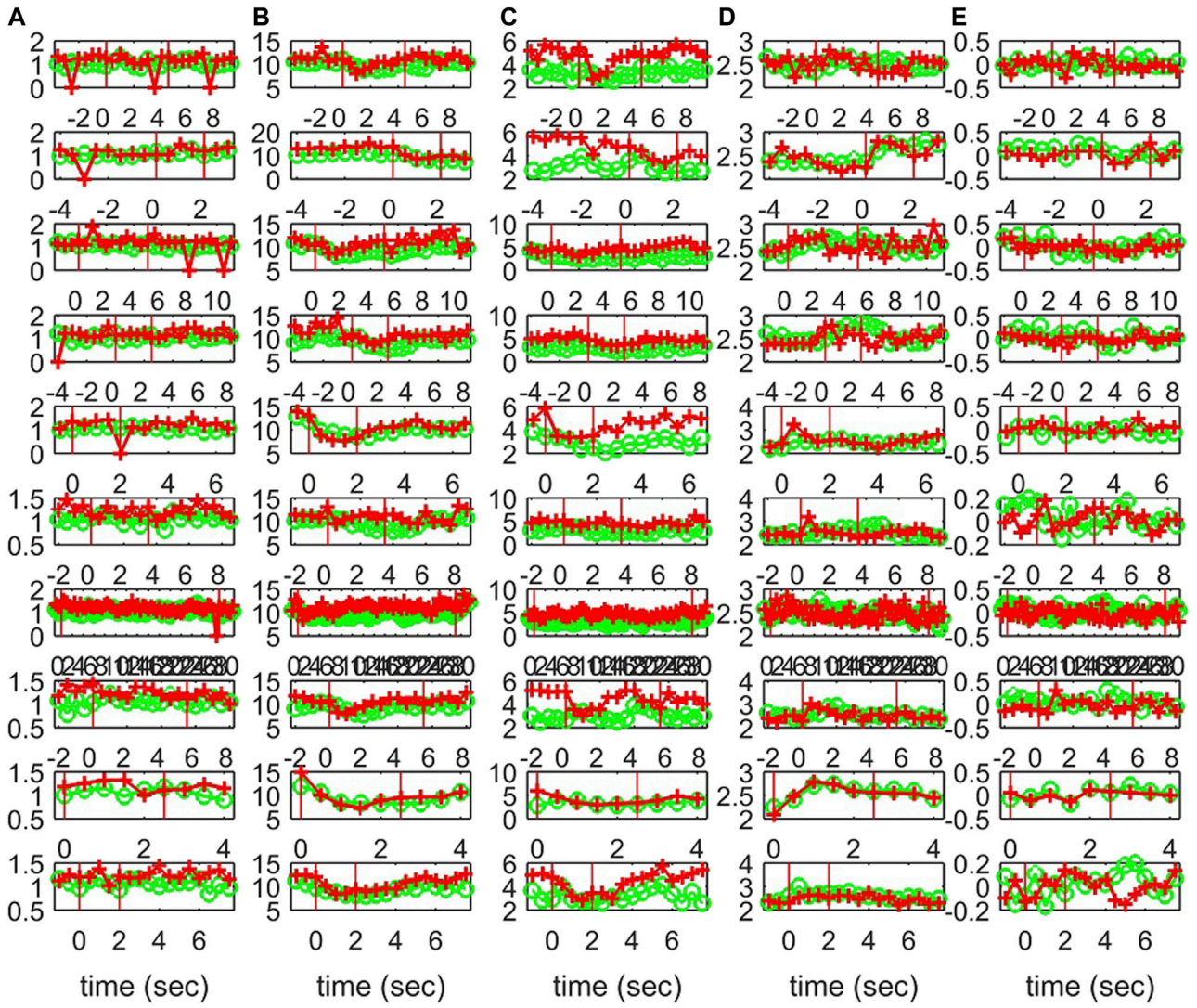


FIGURE 12 | The profile of five network indices on the sensor (green line with "o" marker) and on the source space (red line with "+" marker) for 10 epochs of S3 given in the row panels: **(A)** *SWi*, **(B)** deg^{m}, **(C)** deg^{SD}, **(D)** λ, **(E)** r.

Similarly to the CMG case, the values of all network indices for the initial network and the network estimated on the sensor space range in a similar scale. In contrast, the values of the network indices for the networks estimated on the source space range in different scales with those of the initial networks in most of the cases. Again, this is confirmed with the results of AUROC in **Table 1**.

In **Figure 8**, the respective results for the VAR system are presented. Here, the causality networks are estimated with the linear causality measure RCGCI, which is more appropriate (and simple) for the linear VAR process. Regarding the range of network index values, in many cases the initial network, the sensor space network and the source space network share the same range of values, e.g., for the random network type and all measures or the scale-free network and *SWi* and λ. Also, the discriminative information of the network indices is preserved in the sensor space in all cases, but in the source space only for *SWi* and λ, as it is shown also in **Table 1**.

4.2 EEG Data Analysis

For the EEG data analysis, the epochs from three subjects that contain epileptiform discharges (ED) are considered. The EDs of the 7 epochs of Subject 1 (S1) have duration from 2.19 to 5.71 s and the EDs of the 9 epochs of Subject 3 (S3) from 1.94 to 5.63 s, whereas the 10th epoch is exceptionally long at 29.05 s. The 3 epochs of Subject 2 (S2) have all very short EDs of duration 1.40, 1.80, and 2.09 s. Each epoch consists of a 10 s pre-ED period, an ED period and a 10 s post-ED period. The epoch is split in sliding overlapping windows of 2 s duration and 0.5 s step on which the causality estimation is performed with PMIME. As shown in **Figure 9**, for each window a causality network is estimated from the initial time series

TABLE 2 | Average AUROC values over the 7 epochs of S1 and the 10 epochs of S3 for three discrimination tasks (pre-ED vs. ED, pre-ED vs. post-ED, ED vs. post-ED), both spaces (sensor, source) and five network metrics (labels in the first row).

	SWi	deg^m	deg^{SD}	λ	r
7 epochs of S1					
sensor space					
pre-ED vs. ED	0.74	0.88	0.78	0.85	0.65
pre-ED vs. post-ED	0.65	0.61	0.58	0.62	0.57
ED vs. post-ED	0.68	0.77	0.79	0.76	0.62
source space					
pre-ED vs. ED	0.59	0.80	0.77	0.75	0.60
pre-ED vs. post-ED	0.69	0.67	0.79	0.64	0.57
ED vs. post-ED	0.60	0.70	0.69	0.67	0.58
10 epochs of S3					
sensor space					
pre-ED vs. ED	0.73	0.88	0.74	0.84	0.75
pre-ED vs. post-ED	0.71	0.78	0.73	0.77	0.80
ED vs. post-ED	0.62	0.71	0.66	0.74	0.65
source space					
pre-ED vs. ED	0.68	0.84	0.79	0.78	0.67
pre-ED vs. post-ED	0.87	0.80	0.99	0.71	0.83
ED vs. post-ED	0.62	0.74	0.72	0.64	0.60

and from the time series obtained after the sLORETA transformation.

Characteristics of the causality networks are estimated by network metrics at each time window and the profiles of the metrics are shown in **Figure 10** for S1, **Figure 11** for S2, and **Figure 12** for S3. It is observed that the network indices deg^m and λ in both the sensor and the source space have similar estimated values during the epochs, with subtle differences at various times, e.g., for S1 the coefficient of determination R^2 is 0.52 and 0.28, respectively. In contrast, SWi shows significant differences in the source space at many time points, e.g., for S1 $R^2 = 0.02$, which is in agreement with the simulation study.

The interest is mainly in the pre-ED and ED periods where the brain state changes abruptly at the ED onset, whereas from ED to post-ED the change is gradual as the brain state recovers from the ED. To quantify the differences in the three states, before the ED (pre-ED), within the ED (ED) and after the ED (post-ED), we computed the AUROC for pre-ED vs. ED, pre-ED vs. post-ED and ED vs. post-ED in both spaces (sensor and source) and all network indices. The results for S1 and S3 are shown in **Table 2**, while for S2 statistical comparison is not possible as the EDs in the three episodes are very short.

The network index deg^m discriminates well the pre-ED from ED period in all episodes of the two subjects with average AUROC value 0.88 for the sensor space for both S1 and S3 and 0.80 for S1 and 0.84 for S3 for the source space. In the post-ED period, deg^m recovers its initial level gradually.

TABLE 3 | Results of the independent sample t-test for equal means of pre-ED and ED for subjects S1 and S3 and the five network indices as given in the first row. The numbers correspond to the ED episodes of S1 and S3 for which the p-value of the test is < 0.01 (for each subject first row is for the sensor space and second row for the source space).

		SWi	deg^m	deg^{SD}	λ	r
S1	sensor	2 6 7	1 3 5 6 7	1 3 6 7	1 3 5 6 7	4
	source		1 3 5 7	1 2 3 5	3 5 7	
S2	sensor	—	1 3 7 10	3	2 3 4	—
	source		2 4	2 4	4	

For the network index λ the respective AUROC values are 0.85 and 0.84 for S1 and S3 for the sensor space while for the source space the corresponding AUROC are 0.75 and 0.78, respectively. The other two indices, SW and r, have low discrimination ability, but still the AUROC values are higher for the sensor space than for the source space for both subjects S1 and S3. These findings are also confirmed by the independent sample t-test for equal means of pre-ED and ED, as shown in **Table 3**, where for more episodes at the sensor space than at the source space the discrimination is established for the stringent significance level of 0.01.

The results show that the network characteristics do change after the transformation to the source space at a varying degree, so that the discrimination of the pre-ED vs. ED period is less significant. This result is not in agreement to a recent report of qualitatively similar results of connectivity structure in the source activity from reconstructed scalp EEG data and the connectivity from corresponding electrocorticographic sources in primates (*Macaca mulatta*) (Papadopoulou et al., 2019).

5 DISCUSSION

In this work, the level of preservation of the main network structure when estimated on the original data, also called the sensor space, and data transformed to the so-called source space has been investigated. The transformation is performed using the software sLORETA and the causality effects among the system variables are estimated with a linear method (RCGCI) and a nonlinear information based method (PMIME). After the causality network estimation, certain network characteristics are estimated by network indices in order to compare the respective network topologies in the two spaces (sensor and source).

We performed a simulation study first in order to compare the results to the predefined ground truth, defined in the sensor space. Admittedly, the true brain system is in source space and physiologically it would be more appropriate to define the simulation systems in the source space as done in (Brunner et al., 2016; Anzolin et al., 2019; Van de Steen et al., 2019). In contrast to these studies, where causal relationships of distinct sources are investigated, our study is on the overall structure of causality networks, requiring large scale systems that could only

be realized under simplifications in the source space, e.g,. in (Silva Pereira et al., 2017) this was attempted assuming the same activity in all vertices (voxels) in a ROI on the cortical mesh. On the other hand, it is common to assume the simulation systems at the domain of the acquired data (here the scalp EEG). Thus, the simulation systems refer to the observed dynamics, and the focus of the study is on the change of the overall network structure of the high-dimensional coupled system under the inverse transform defined by sLORETA.

Three different network types were considered as initial coupling structures for three high-dimensional coupled dynamical systems: $K = 20$ coupled Mackey-Glass systems (CMG), $K = 20$ Hénon coupled maps (CHM) and the linear vector autoregressive process of order one on $K = 20$ variables (VAR). The objective was first to designate the differences in the topology of the causality networks in the two spaces, and then to assess whether the information the network metrics hold regarding the discrimination of the three network types was preserved after the transformation to the source space. Then we proceeded to EEG data analysis with illustrative cases from recordings of three epileptic patients that contain epileptiform discharges (EDs). The objective was again to compare the differences in the topology of the estimated causality networks in the sensor and in the source space. In addition, we tried to clarify the degree to which the topology of the networks held discriminative information regarding the pre-ED period vs. the ED and in which space the discrimination was more clear.

In the simulation study, first we showed with an analytical example that a linear transformation of the system variables can change considerably the coupling structure of the dynamical system, but it can preserve some characteristics e.g. the eigenvalues of its adjacency matrix (similar derivation was obtained in Van de Steen et al., 2019). Then, the simulation results showed that the estimated causality networks have, as expected, considerable differences in terms of the topology characteristics. Finally, the ability of certain characteristics to discriminate the initial coupling structures is reduced at a varying degree after the transformation to the source space. We note here that causality estimation inaccuracy (false positives or negatives) exists also in the sensor space (the domain of true dynamics) due to many sources, such as the time series length, the size in conjunction with the density of the true coupling network, and the type of inherent dynamics and causal relationships (the latter may encounter the common source problem related to volume conduction for EEG). However, the utilized measures RCGCI and PMIME were found to perform well in high-dimensional coupled systems and therefore they were used here to allow for a better assessment of the differences in the estimated networks in sensor and source space.

In the EEG data analysis, we showed that the characteristics of the causality network topology were altered at varying degree. Here, the small-worldness index was found to change most under the transform to source space and had a very different profile at overlapping windows across the epochs that contain ED. The average degree of the causality networks was the most discriminative characteristic regarding the pre-ED vs. ED task. Also the characteristic path length showed good discrimination ability for the same task. For both indices the ability to discriminate the two periods was better in the sensor space than in the source space. The clinical EEG data are essentially a case study providing complementary evidence that clearly corroborates the basic conclusion of the simulation analysis. Further studies, on the basis of our preliminary findings, are warranted including a larger number of subjects.

We have tactically left out of the design of the connectivity analysis the issue of field spread and volume conduction. These factors may affect the results of connectivity analysis in the sensor space and subsequently the source space. It is not known as to what extent the volume conduction affects through the inverse transform the connectivity analysis in the source space (Schoffelen and Gross, 2009), but there is reported evidence from simulation studies that it does (Anzolin et al., 2019). Functional connectivity measures suggested in the literature to account for volume conduction, such as the imaginary coherence (Nolte et al., 2004), have been considered when comparing connectivity in sensor and source space (Schoffelen and Gross, 2009; Lai et al., 2018; Demuru et al., 2020). However, these measures are bivariate (allow for indirect connections) and thus not suitable to estimate the network structure, which is the main objective of this work. The utilized causality measures here are multivariate and when estimating a causal effect they account for the presence of all other observed significant sources, including to some extent the common source due to volume conduction. Certainly, they do not explicitly account for zero-lag effects brought about by volume conduction. One of the two measures used in the study has been recently modified to account for zero-lag effects, called PMIME0 (Koutlis et al., 2019). We focused on measures of lag-causality here and leave extension to other causality measures including zero-lag effects to future work.

We have also refrained from discussing the different solutions for the inverse transform and chose the sLORETA as one of the most popular transforms. Certainly, the estimated connectivity network in source space depends on the utilized inverse transform, and this is also an open research topic gaining so far little attention, e.g., two tested inverse transforms were found to give distinctly different connectivity network characteristics in source space (Lai et al., 2018).

In view of the increasing usage of source analysis in diverse areas of neuroscience, the above data suggest that further studies in order to clarify the relationship between sensor- and source-derived data are warranted.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Aristotle University of Thessaloniki, Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CK developed the computational procedure, made the computations and evaluation of results and wrote the manuscript. VK gave the EEG data and contributed in the discussion of the results and the editing of the manuscript. DK designed the methods and simulated systems, had the overview of the simulation study and edited the manuscript.

FUNDING

Part of the research was supported by a grant in the frame of the Greek–Israel Agreement on Bilateral Industrial R and D Cooperation from General Secretariat for Research and Technology of the Ministry of Education, Research, and Religious Affairs in Greece.

REFERENCES

Anzolin, A., Presti, P., Van de Steen, F., Astolfi, L., Haufe, S., and Marinazzo, D. (2019). Quantifying the Effect of Demixing Approaches on Directed Connectivity Estimated between Reconstructed EEG Sources. *Brain Topogr.* 32 (4), 655–674. doi:10.1007/s10548-019-00705-z

Bakhshayesh, H., Fitzgibbon, S. P., Janani, A. S., Grummett, T. S., and Pope, K. J. (2019). Detecting Connectivity in EEG: A Comparative Study of Data-Driven Effective Connectivity Measures. *Comput. Biol. Med.* 111, 103329. doi:10.1016/j.compbiomed.2019.103329

Barabási, A.-L., and Albert, R. (1999). Emergence of Scaling in Random Networks. *Science* 286 (5439), 509–512. doi:10.1126/science.286.5439.509

Basu, S., and Michailidis, G. (2015). Regularized Estimation in Sparse High-Dimensional Time Series Models. *Ann. Stat.* 43 (4), 1535–1567. doi:10.1214/15-aos1315

Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012). Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors. *J. Financ. Econ.* 104 (3), 535–559. doi:10.1016/j.jfineco.2011.12.010

Bollen, J., Mao, H., and Zeng, X. (2011). Twitter Mood Predicts the Stock Market. *J. Comput. Sci.* 2 (1), 1–8. doi:10.1016/j.jocs.2010.12.007

Brunner, C., Billinger, M., Seeber, M., Mullen, T. R., and Makeig, S. (2016). Volume Conduction Influences Scalp-Based Connectivity Estimates. *Front. Comput. Neurosci.* 10, 121. doi:10.3389/fncom.2016.00121

Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., et al. (2000). Dynamic Statistical Parametric Mapping. *Neuron* 26, 55–67. doi:10.1016/s0896-6273(00)81138-1

De Vico Fallani, F., Maglione, A., Babiloni, F., Mattia, D., Astolfi, L., Vecchiato, G., et al. (2010). Cortical Network Analysis in Patients Affected by Schizophrenia. *Brain Topogr.* 23 (2), 214–220. doi:10.1007/s10548-010-0133-2

Demuru, M., La Cava, S. M., Pani, S. M., and Fraschini, M. (2020). A Comparison between Power Spectral Density and Network Metrics: An EEG Study. *Biomed. Signal Process. Control.* 57, 101760. doi:10.1016/j.bspc.2019.101760

Dijkstra, H. A., Hernández-García, E., Masoller, C., and Barreiro, M. (2019). *Networks in Climate*. Cambridge: Cambridge University Press.

Erdös, P., and Rényi, A. (1959). On Random Graphs I. *Publicationes Math. Debrecen.* 6, 290–297.

Fawcett, T. (2006). An Introduction to ROC Analysis. *Pattern Recognit. Lett.* 27 (8), 861–874. cited By (since 1996)2503. doi:10.1016/j.patrec.2005.10.010

Fornito, A., Zalesky, A., and Bullmore, E. T. (2016). *Fundamentals of Brain Network Analysis*. 1 edition. Academic Press, Elsevier.

Geier, C., and Lehnertz, K. (2017). Which Brain Regions Are Important for Seizure Dynamics in Epileptic Networks? Influence of Link Identification and EEG Recording Montage on Node Centralities. *Int. J. Neural Syst.* 27 (1), 1650033. doi:10.1142/s0129065716500337

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* 37 (3), 424–438. doi:10.2307/1912791

Hämäläinen, M. S., and Ilmoniemi, R. J. (1994). Interpreting Magnetic fields of the Brain: Minimum Norm Estimates. *Med. Biol. Eng. Comput.* 32 (1), 35–42. doi:10.1007/bf02512476

Hata, M., Kazui, H., Tanaka, T., Ishii, R., Canuet, L., Pascual-Marqui, R. D., et al. (2016). Functional Connectivity Assessed by Resting State EEG Correlates with Cognitive Decline of Alzheimer's Disease - an eLORETA Study. *Clin. Neurophysiol.* 127 (2), 1269–1278. doi:10.1016/j.clinph.2015.10.030

He, B., Astolfi, L., Valdes-Sosa, P. A., Marinazzo, D., Palva, S. O., Benar, C.-G., et al. (2019). Electrophysiological Brain Connectivity: Theory and Implementation. *IEEE Trans. Biomed. Eng.* 66 (7), 2115–2137. doi:10.1109/tbme.2019.2913928

Helmholtz, H. (1853). Ueber einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern mit Anwendung auf die thierisch-elektrischen Versuche. *Ann. Phys. Chem.* 165 (6), 211–233. doi:10.1002/andp.18531650603

Hincapié, A.-S., Kujala, J., Mattout, J., Pascarella, A., Daligault, S., Delpuech, C., et al. (2017). The Impact of MEG Source Reconstruction Method on Source-Space Connectivity Estimation: A Comparison Between Minimum-Norm Solution and Beamforming. *Neuroimage* 156, 29–42. doi:10.1016/j.neuroimage.2017.04.038

Hong, Y., Liu, Y., and Wang, S. (2009). Granger Causality in Risk and Detection of Extreme Risk Spillover Between Financial Markets. *J. Econom.* 150, 271–287. doi:10.1016/j.jeconom.2008.12.013

Koutlis, C., and Kugiumtzis, D. (2016). Discrimination of Coupling Structures Using Causality Networks from Multivariate Time Series. *Chaos* 26 (9), 093120. doi:10.1063/1.4963175

Koutlis, C., Kimiskidis, V. K., and Kugiumtzis, D. (2019). Identification of Hidden Sources by Estimating Instantaneous Causality in High-Dimensional Biomedical Time Series. *Int. J. Neural Syst.* 29 (4), 1850051. doi:10.1142/S012906571850051X

Kugiumtzis, D., and Kimiskidis, V. K. (2015). Direct Causal Networks for the Study of Transcranial Magnetic Stimulation Effects on Focal Epileptiform Discharges. *Int. J. Neur. Syst.* 25, 1550006. doi:10.1142/s0129065715500069

Kugiumtzis, D. (2013). Direct-coupling Information Measure from Nonuniform Embedding. *Phys. Rev. E Stat. Nonlin Soft Matter Phys.* 87, 062918. doi:10.1103/PhysRevE.87.062918

Lai, M., Demuru, M., Hillebrand, A., and Fraschini, M. (2018). A Comparison between Scalp- and Source-Reconstructed EEG Networks. *Sci. Rep.* 8, 12269. doi:10.1038/s41598-018-30869-w

Lehmann, D., Faber, P. L., Pascual-Marqui, R. D., Milz, P., Herrmann, W. M., Koukkou, M., et al. (2014). Functionally Aberrant Electrophysiological Cortical Connectivities in First Episode Medication-Naive Schizophrenics from Three Psychiatry Centers. *Front. Hum. Neurosci.* 8, 635. doi:10.3389/fnhum.2014.00635

Lehnertz, K., Ansmann, G., Bialonski, S., Dickten, H., Geier, C., and Porz, S. (2014). Evolving Networks in the Human Epileptic Brain. *Physica D* 267, 7–15. doi:10.1016/j.physd.2013.06.009

Lei, X., Wu, T., and Valdes-Sosa, P. A. (2015). Incorporating Priors for EEG Source Imaging and Connectivity Analysis. *Front. Neurosci.* 9, 284. doi:10.3389/fnins.2015.00284

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., and Grave De Peralta, R. (2004). EEG Source Imaging. *Clin. Neurophysiol.* 115 (10), 2195–2222. doi:10.1016/j.clinph.2004.06.001

Milz, P., Faber, P. L., Lehmann, D., Kochi, K., and Pascual-Marqui, R. D. (2014). sLORETA Intracortical Lagged Coherence During Breath Counting in Meditation-Naïve Participants. *Front. Hum. Neurosci.* 8, 303. doi:10.3389/fnhum.2014.00303

Mulert, C., Kirsch, V., Pascual-Marqui, R., Mccarley, R. W., and Spencer, K. M. (2011). Long-range Synchrony of Gamma Oscillations and Auditory Hallucination Symptoms in Schizophrenia. *Int. J. Psychophysiol.* 79 (1), 55–63. doi:10.1016/j.ijpsycho.2010.08.004

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., and Hallett, M. (2004). Identifying True Brain Interaction from EEG Data Using the Imaginary Part of Coherency. *Clin. Neurophysiol.* 115 (10), 2292–2307. doi:10.1016/j.clinph.2004.04.029

Papadaniil, C. D., and Hadjileontiadis, L. J. (2015). Tomographic Reconstruction of 3-D Irrotational Vector fields via a Discretized ray Transform. *J. Math. Imaging Vis.* 52 (2), 285–302. doi:10.1007/s10851-015-0559-y

Papadaniil, C. D., Kosmidou, V. E., Tsolaki, A., Tsolaki, M., Kompatsiaris, I., and Hadjileontiadis, L. J. (2016). Cognitive MMN and P300 in Mild Cognitive

Impairment and Alzheimer's Disease: A High Density EEG-3D Vector Field Tomography Approach. *Brain Res.* 1648 (Pt A), 425–433. doi:10.1016/j.brainres.2016.07.043

Papadopoulou, M., Friston, K., and Marinazzo, D. (2019). Estimating Directed Connectivity from Cortical Recordings and Reconstructed Sources. *Brain Topogr.* 32 (4), 741–752. doi:10.1007/s10548-015-0450-6

Papana, A., Kugiumtzis, D., and Larsson, P. G. (2012). Detection of Direct Causal Effects and Application in the Analysis of Electroencephalograms from Patients with Epilepsy. *Int. J. Bifurcation Chaos* 22 (9), 1250222. doi:10.1142/s0218127412502227

Pascual-Marqui, R. D., Michel, C. M., and Lehmann, D. (1994). Low Resolution Electromagnetic Tomography: A New Method for Localizing Electrical Activity in the Brain. *Int. J. Psychophysiol.* 18 (1), 49–65. doi:10.1016/0167-8760(84)90014-x

Pascual-Marqui, R. D., Biscay, R. J., Bosch-Bayard, J., Lehmann, D., Kochi, K., Kinoshita, T., et al. (2014). Assessing Direct Paths of Intracortical Causal Information Flow of Oscillatory Activity with the Isolated Effective Coherence (Icoh). *Front. Hum. Neurosci.* 8, 448. doi:10.3389/fnhum.2014.00448

Pascual-Marqui, R. D. (2002). Standardized Low-Resolution Brain Electromagnetic Tomography (sLORETA): Technical Details. *Methods Find Exp. Clin. Pharmacol.* 24 Suppl D, 5–12.

Porta, A., and Faes, L. (2016). Wiener-Granger Causality in Network Physiology with Applications to Cardiovascular Control and Neuroscience. *Proc. IEEE* 104 (2), 282–309. doi:10.1109/jproc.2015.2476824

Qin, Y., Xu, P., and Yao, D. (2010). A Comparative Study of Different References for EEG Default Mode Network: The Use of the Infinity Reference. *Clin. Neurophysiol.* 121 (12), 1981–1991. doi:10.1016/j.clinph.2010.03.056

Rossini, P. M., Di Iorio, R., Bentivoglio, M., Bertini, G., Ferreri, F., Gerloff, C., et al. (2019). Methods for Analysis of Brain Connectivity: An IFCN-Sponsored Review. *Clin. Neurophysiol.* 130 (10), 1833–1858. doi:10.1016/j.clinph.2019.06.006

Rubinov, M., and Sporns, O. (2010). Complex Network Measures of Brain Connectivity: Uses and Interpretations. *Neuroimage* 52 (3), 1059–1069. doi:10.1016/j.neuroimage.2009.10.003

Runge, J., Petoukhov, V., Donges, J. F., Hlinka, J., Jajcay, N., Vejmelka, M., et al. (2015). Identifying Causal Gateways and Mediators in Complex Spatio-Temporal Systems. *Nat. Commun.* 6, 8502. doi:10.1038/ncomms9502

Schelter, B., Winterhalder, M., Maiwald, T., Brandt, A., Schad, A., Schulze-Bonhage, A., et al. (2006). Testing Statistical Significance of Multivariate Time Series Analysis Techniques for Epileptic Seizure Prediction. *Chaos* 16 (1), 013108. doi:10.1063/1.2137623

Schoffelen, J.-M., and Gross, J. (2009). Source Connectivity Analysis with MEG and EEG. *Hum. Brain Mapp.* 30 (6), 1857–1865. doi:10.1002/hbm.20745

Schreiber, T. (2000). Measuring Information Transfer. *Phys. Rev. Lett.* 85 (2), 461–464. doi:10.1103/physrevlett.85.461

Senthilkumar, D. V., Lakshmanan, M., and Kurths, J. (2008). Transition from Phase to Generalized Synchronization in Time-Delay Systems. *Chaos* 18 (2), 023118. doi:10.1063/1.2911541

Siggiridou, E., and Kugiumtzis, D. (2016). Granger Causality in Multivariate Time Series Using a Time-Ordered Restricted Vector Autoregressive Model. *IEEE Trans. Signal. Process.* 64 (7), 1759–1773. doi:10.1109/tsp.2015.2500893

Siggiridou, E., Koutlis, C., Tsimpiris, A., and Kugiumtzis, D. (2019). Evaluation of Granger Causality Measures for Constructing Networks from Multivariate Time Series. *Entropy* 21 (11), 1080. doi:10.3390/e21111080

Silva Pereira, S., Hindriks, R., Mühlberg, S., Maris, E., van Ede, F., Griffa, A., et al. (2017). Effect of Field Spread on Resting-State Magneto Encephalography Functional Network Analysis: A Computational Modeling Study. *Brain Connect.* 7 (9), 541–557. doi:10.1089/brain.2017.0525

Sommariva, S., Sorrentino, A., Piana, M., Pizzella, V., and Marzetti, L. (2019). A Comparative Study of the Robustness of Frequency-Domain Connectivity Measures to Finite Data Length. *Brain Topogr.* 32 (4), 675–695. doi:10.1007/s10548-017-0609-4

Toppi, J., Astolfi, L., Poudel, G. R., Innes, C. R. H., Babiloni, F., and Jones, R. D. (2016). Time-varying Effective Connectivity of the Cortical Neuroelectric Activity Associated with Behavioural Microsleeps. *NeuroImage* 124, 421–432. doi:10.1016/j.neuroimage.2015.08.059

Van de Steen, F., Faes, L., Karahan, E., Songsiri, J., Valdes-Sosa, P. A., and Marinazzo, D. (2019). Critical Comments on EEG Sensor Space Dynamical Connectivity Analysis. *Brain Topogr.* 32 (4), 643–654. doi:10.1007/s10548-016-0538-7

Watts, D. J., and Strogatz, S. H. (1998). Collective Dynamics of 'small-World' Networks. *Nature* 393 (6684), 440–442. doi:10.1038/30918

Yao, D. (2001). A Method to Standardize a Reference of Scalp EEG Recordings to a point at Infinity. *Physiol. Meas.* 22 (4), 693–711. doi:10.1088/0967-3334/22/4/305

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.

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.

frontiers in Network Physiology

ORIGINAL RESEARCH
published: 28 January 2022
doi: 10.3389/fnetp.2021.765332

Measuring the Rate of Information Exchange in Point-Process Data With Application to Cardiovascular Variability

*Gorana Mijatovic[1], Riccardo Pernice[2], Alessio Perinelli[3], Yuri Antonacci[4], Alessandro Busacca[2], Michal Javorka[5], Leonardo Ricci[6] and Luca Faes[2]**

[1]Faculty of Technical Science, University of Novi Sad, Novi Sad, Serbia, [2]Department of Engineering, University of Palermo, Palermo, Italy, [3]CIMeC, Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy, [4]Department of Physics and Chemistry "Emilio Segrè," University of Palermo, Palermo, Italy, [5]Department of Physiology and Biomedical Center Martin, Jessenius Faculty of Medicine, Comenius University, Martin, Slovakia, [6]Department of Physics, University of Trento, Trento, Italy

The amount of information exchanged per unit of time between two dynamic processes is an important concept for the analysis of complex systems. Theoretical formulations and data-efficient estimators have been recently introduced for this quantity, known as the mutual information rate (MIR), allowing its continuous-time computation for event-based data sets measured as realizations of coupled point processes. This work presents the implementation of MIR for point process applications in Network Physiology and cardiovascular variability, which typically feature short and noisy experimental time series. We assess the bias of MIR estimated for uncoupled point processes in the frame of surrogate data, and we compensate it by introducing a corrected MIR (cMIR) measure designed to return zero values when the two processes do not exchange information. The method is first tested extensively in synthetic point processes including a physiologically-based model of the heartbeat dynamics and the blood pressure propagation times, where we show the ability of cMIR to compensate the negative bias of MIR and return statistically significant values even for weakly coupled processes. The method is then assessed in real point-process data measured from healthy subjects during different physiological conditions, showing that cMIR between heartbeat and pressure propagation times increases significantly during postural stress, though not during mental stress. These results document that cMIR reflects physiological mechanisms of cardiovascular variability related to the joint neural autonomic modulation of heart rate and arterial compliance.

Keywords: information dynamics, point processes, mutual information rate, heart rate variability, cardiovascular time series

OPEN ACCESS

Edited by:
Dimitris Kugiumtzis, Aristotle University of Thessaloniki, Greece

Reviewed by:
Christos Frantzidis, Aristotle University of Thessaloniki, Greece
Débora Corrêa, University of Western Australia, Australia

****Correspondence:***
Luca Faes
luca.faes@unipa.it

Specialty section:
This article was submitted to Information Theory, a section of the journal Frontiers in Network Physiology

Received: *26 August 2021*
Accepted: *26 November 2021*
Published: *28 January 2022*

Citation:
Mijatovic G, Pernice R, Perinelli A, Antonacci Y, Busacca A, Javorka M, Ricci L and Faes L (2022) Measuring the Rate of Information Exchange in Point-Process Data With Application to Cardiovascular Variability. Front. Netw. Physiol. 1:765332. doi: 10.3389/fnetp.2021.765332

1 INTRODUCTION

The mutual information (MI) between two random variables is a central concept in information theory. MI is an important quantity with huge practical relevance, as it quantifies how much information is exchanged between two complex systems or is shared by two data sets. Indeed, thanks to these characteristics, MI is ubiquitously employed in diverse fields of science and engineering to

assess linear and non-linear interactions, e.g., between electronic oscillators (Minati et al., 2018), financial systems (Fiedor, 2014), climatological variables (Perinelli et al., 2021), brain units (Mijatovic et al., 2021b) or physiological systems (Valderas et al., 2019). In all these application fields, the study of dynamical systems, i.e., systems whose state evolves over time, is central to the understanding of the underlying phenomena. Therefore, dynamic formulations of MI in which the observed variables are associated with temporal information are recommended for a proper assessment of the interactions between the system units. In this study we consider the MI rate (MIR), a well-known quantity measuring the amount of information shared by two random processes per unit of time (Duncan, 1970). In particular, we focus on the computation of MIR for point processes, i.e., processes where the relevant information stands in the times of occurrence of specific events. This class of processes is widely adopted in neuroscience, for instance to study the spiking activity of neural populations acquired through multi-electrode recording techniques (Truccolo et al., 2005), and in the field of cardiovascular variability, where the point process nature of the human heartbeats has inspired the development of event-based models to describe the heart rate and its interaction with vascular, respiratory and metabolic variables (Barbieri et al., 2005; Valenza et al., 2018).

The calculation of dynamic information measures, such as the MIR or the transfer entropy rate (TER) quantifying the rate of directed (causal) information flow between stochastic processes (Schreiber, 2000; Spinney et al., 2017) is well-established for discrete-time processes, i.e., processes defined at discrete time instant, which represent the sampling rate of continuous-time signals or the rate of a physiological oscillator (e.g., the cardiac pacemaker); in this context, a number of practical approaches exist to provide data-efficient estimates (Vicente et al., 2011; Faes et al., 2015). On the other hand, the definition and practical computation of these measures for continuous-time processes defined at each time instant with arbitrarily small resolution, and more specifically for point processes, is much more cumbersome. The classical way to compute MIR and TER for point process or other event-based data typically relies on binning of the temporal axis followed by the application of discrete-time estimators (Pasquale et al., 2008), but unavoidably implies loss of information and strong dependence on the parameters related to time discretization (Mijatovic et al., 2021a; Shorten et al., 2021). Only recently, the theoretical formalism (Spinney et al., 2017; Spinney and Lizier, 2018) and the design of estimation approaches for the TER (Shorten et al., 2021) and MIR (Mijatovic et al., 2021a) has been introduced in the context of neuroscience applications. In particular, Mijatovic et al. (2021a) have shown that for point process data the MIR can be expressed as the sum of the TER computed along the two directions of interaction between the two analyzed processes, and have exploited the TER estimation methods introduced by (Shorten et al., 2021) to design a data-efficient estimator of the MIR for coupled point processes. These works are of a great practical relevance, because they open the way for a reliable non-parametric, continuous-time estimation of the information transfer for event-based processes.

In this work, we exploit the MIR estimator introduced in (Mijatovic et al., 2021a) to assess the rate of information shared between cardiovascular point processes. Specifically, we focus on cardiovascular interactions assessed between the cardiac pacemaker, studied by the heartbeat timings and measured from the ECG, and the times of arrival to the body periphery of the sphygmic wave, measured through finger photoplethysmography. The application of event-based frameworks to heartbeat and pulse arrival times entertains a different perspective on the study of cardiovascular regulation than more classical analyses performed on time series of the heart period and arterial pressure variability (Cohen and Taylor, 2002; Porta and Faes, 2015), and leads to address related but different physiological mechanisms. In particular, while classical time series analysis methods investigate cardiovascular interactions focusing on baroreflex regulation and mechanical mechanisms (Faes et al., 2013; Javorka et al., 2017), the study of coupled point processes may reveal the physiological mechanisms that modulate the arterial pressure, the contractility of the ventricles and vasomotion (Okada et al., 1996; Chan et al., 2007). Since these mechanisms typically operate on short time scales involving a few heartbeats, and due to stationarity issues, the analysis of these processes is typically restricted to short realizations (few hundred events). A practical consequence of this restriction is the difficulty of obtaining reliable estimates in the presence of short series of data. To test the applicability of the MIR estimator on short realizations of point process data, we assess the estimation bias in simulations of uncoupled point processes generated for different parametric probability distributions. When there is no coupling between the processes, a positive bias can be misinterpreted as a weak coupling, while a negative bias makes a non-negative measure like the MIR of difficult interpretation. We provide a solution to this problem, by modifying the MIR estimator and introducing a corrected MIR (cMIR) measure for which the bias is reduced; the correction employs surrogate time series, which reproduce the bias occurring for uncoupled point processes. The novel cMIR measure is tested first in simulated point process models that reproduce the coupled occurrence of the heartbeat times and of the arrival instants of the blood pressure wave at the body periphery, and then in real point process series measured from healthy subjects monitored in resting state and during postural and mental stress (Javorka et al., 2017).

2 INFORMATION-THEORETIC MEASURES TO ASSESS THE DYNAMIC INTERACTION BETWEEN STOCHASTIC PROCESSES

This section presents the mathematical background necessary to assess the information shared between continuous-time stochastic processes. Information-theoretic measures are typically employed to treat dynamic systems in discrete time, i.e. systems can be described by processes whose states are mapped by times series values. However, many theoretical and real-world systems are naturally described by processes defined in continuous time, whose available discrete-time signals represent

approximate realizations. The most accurate information-theoretic treatment of continuous-time processes is that using random functions in place of collections of random variables to quantify information dynamics (Spinney et al., 2017). In the following subsections, we show how to employ random functions to define the information dynamically shared between two continuous-time processes, how to express it in terms of the information transferred along the two directions of interaction between the processes, and how to formalize its computation and practical estimation in the particular case of point processes.

2.1 Mutual Information and Transfer Entropy Rates

Let us consider two possibly coupled dynamical systems $\mathcal{X}$ and $\mathcal{Y}$ such that their evolution over time is mapped by the continuous-time stochastic processes $X = \{X_t\}$ and $Y = \{Y_t\}$, which are defined at each continuous-time instant $t \in \mathbb{R}$. A well-known undirected measure of the dynamical interaction between X and Y is the *mutual information rate* (MIR), which quantifies the amount of information exchanged per unit of time by the two processes (Duncan, 1970). If the processes are stationary, the MIR is defined as

$$\dot{I}_{X;Y} = \lim_{\tau\to\infty} \frac{1}{\tau} I(X_{t-\tau:\,t}; Y_{t-\tau:\,t}), \tag{1}$$

where $I(\cdot;\,\cdot)$ denotes mutual information (MI) and τ is the duration of the temporal window over which the MI is computed. The notation $X_{t-\tau:t}$ denotes the random function expressing the stochastic process evaluated along the time interval of duration τ ending at the time t, also referred to as *path* (Spinney et al., 2017), i.e. $X_{t-\tau:t} = \{X_s : t - \tau \le s < t\}$ (the same holds for the process Y); note that the MI in **Eq. 1**, and consequently the MIR, are independent on t due to stationarity.

The MIR defined above, as any other information measure applied to continuous-time processes, cannot be readily formulated in terms of the probability mass functions or densities used for discrete and continuous random variables. In continuous time, a viable approach is to establish a generalized form for the information measures via measure-theoretic approaches that unify under one framework the methods specifically developed for discrete and continuous random variables (Spinney et al., 2017). In this framework, information measures can be expressed by using the Radon-Nykodim derivative between appropriate random density functions defined on paths in place of the ratio between probability distributions of random variables adopted in discrete-time (Gray, 2011). The MI measure in **Eq. 1** can be then expressed in a generalized form as (Duncan, 1970)

$$I(X_{t-\tau:\,t}; Y_{t-\tau:\,t}) = \mathbb{E}_P\left[\ln \frac{dP[x_{t-\tau:\,t}|y_{t-\tau:\,t}]}{dP[x_{t-\tau:\,t}]}\right], \tag{2}$$

where the expectation is taken over the path realizations $x_{t-\tau:t}$ and $y_{t-\tau:t}$ of the random functions $X_{t-\tau:t}$ and $X_{t-\tau:t}$, and the argument of the logarithm is the Radon-Nykodim derivative of two probability measures defined on path functions. With a similar formalism, Spinney and colleagues have formalized different measures of information dynamics for continuous-time processes (Spinney et al., 2017; Spinney and Lizier, 2018). In particular, the transfer entropy rate (TER) from the 'source' process Y to the 'target' process X is defined as (Spinney et al., 2017)

$$\dot{T}_{Y\to X}(t,\tau) = \lim_{\Delta t\to 0} \frac{1}{\Delta t} T_{Y\to X}(t,\Delta t,\tau), \tag{3}$$

where

$$T_{Y\to X}(t,\Delta t,\tau) = \mathbb{E}_P\left[\ln \frac{dP[x_{t+\Delta t}|x_{t-\tau:\,t}, y_{t-\tau:\,t}]}{dP[x_{t+\Delta t}|x_{t-\tau:\,t}]}\right] \tag{4}$$

is the transfer entropy (TE) formulated in terms of a Radon-Nykodim derivative of conditional probability measures similarly as in **Eq. 2** for the MI, and the normalization by the time interval Δt ensures convergence of the TER in the limit of small Δt (Spinney et al., 2017). For stationary processes X and Y, the TER is independent on the time t; moreover, considering realizations of infinite duration yields the constant TER measure

$$\dot{T}_{Y\to X} = \lim_{\tau\to\infty} \dot{T}_{Y\to X}(t,\tau), \tag{5}$$

which quantifies the rate of information transferred along the causal direction from Y to X. By reversing the role of the two processes, the information transferred along the opposite causal direction can be quantified by the TER measure $\dot{T}_{X\to Y}$.

The measures of the rates of information exchanged by X and Y defined in **Eqs. 1, 5** are related to each other by a decomposition that expresses the MIR between X and Y as the sum of the TER along the two directions $X \to Y$ and $Y \to X$, plus a term related to the instantaneous interaction between the two processes. Specifically, by using information-theoretic rules on **Eq. 2** and recognizing **Eq. 4** as a conditional MI, i.e., $T_{Y\to X}(t, \Delta t, \tau) = I(X_{t+\Delta t}; Y_{t-\tau:t}|X_{t-\tau:t})$, the MIR can be expanded as

$$\dot{I}_{X;Y} = \dot{T}_{X\to Y} + \dot{T}_{Y\to X} + \dot{I}^0_{X;Y}, \tag{6}$$

where the term

$$\dot{I}^0_{X;Y} = \lim_{\Delta t\to 0} \lim_{\tau\to\infty} \frac{1}{\Delta t} I(X_{t+\Delta t}; X_{t+\Delta t}|X_{t-\tau:\,t}, Y_{t-\tau:\,t}) \tag{7}$$

quantifies the rate of information instantaneously exchanged between the two processes conditioned to the knowledge of their past histories. The derivation of the important relation **Eq. 6**, where all three terms are quantified in [nats/s], is reported in the **Appendix**.

2.2 Computation for Bivariate Point Processes

In this subsection we formulate the computation of MIR for point processes. A point process is a particular class of continuous-time process that is uniquely characterized by a series of indistinguishable events described by their time of occurrence. In a bivariate context, the statistical description of two point processes is provided in terms of the instants marking the event times, i.e., by writing $X = \{x_i\}$, $i = 1, \ldots, N_X$, and $Y = \{y_j\}$, $j = 1, \ldots,$

N_Y, where x_i and y_j represent the times of the ith event in X and of the jth event in Y, respectively. For these point processes, the MIR can be computed by leveraging the decomposition provided in **Eq. 6** (Mijatovic et al., 2021a) and making the assumption that simultaneous events are not possible, i. e, $x_i \neq y_j$, $\forall i = 1, \ldots, N_X$, $j = 1, \ldots, N_Y$ (Spinney et al., 2017; Mijatovic et al., 2021a; Shorten et al., 2021). This assumption implies that the measure $\dot{I}^0_{X;Y}$ of instantaneous information exchange between X and Y defined in **Eq. 7** is null, so that the MIR between two point processes simply becomes the sum of the two TER terms

$$\dot{I}_{X;Y} = \dot{T}_{X\to Y} + \dot{T}_{Y\to X}. \tag{8}$$

Starting from **Eq. 8**, the MIR can be calculated by employing methods to define (Spinney et al., 2017) and compute (Shorten et al., 2021) the TER for point processes. Specifically, the TER from Y to X is formulated as

$$\dot{T}_{Y\to X} = \bar{\lambda}_X \mathbb{E}_{p_x}\left[\ln \frac{\lambda_{X,x_i|X^-_{x_i},Y^-_{x_i}}}{\lambda_{X,x_i|X^-_{x_i}}}\right], \tag{9}$$

where $\bar{\lambda}_X = N_X/T$ is the average event rate of X, N_X is the number of target events, and T is the duration of the target process; in **Eq. 9**, $\lambda_{X,x_i|X^-_{x_i}}$ and $\lambda_{X,x_i|X^-_{x_i},Y^-_{x_i}}$ are the instantaneous event rates of the target process X evaluated at the time of its ith event x_i, respectively conditioned on the history of X and on the histories of both X and Y. In general, the unconditioned instantaneous event rate of the process X, evaluated at the arbitrary time u, is given by $\lambda_{X,u} = \lim_{\Delta u\to 0} p_u\left(N_{X,u+\Delta u} - N_{X,u} = 1\right)/\Delta u$, where $N_X(u)$ is the counting process that returns the number of events occurred up to time u. At this point it is worth noting that, while the probability p_u is defined at any time point $u \in \mathbb{R}$, the expectation in **Eq. 9** is taken over the probability p_x of observing a quantity precisely at the time of target events x_i, $i = 1, \ldots, N_X$ (Shorten et al., 2021). This important distinction, upon expressing the conditional event rates in terms of p_u, making a Bayes inversion and noting that $\lim_{\Delta u\to 0} p_u\left(\cdot|N_{X,u+\Delta u} - N_{X,u} = 1\right) = p_x(\cdot)$, allows to reformulate the expression of the TER as (Shorten et al., 2021)

$$\dot{T}_{Y\to X} = \bar{\lambda}_X \mathbb{E}_{p_x}\left[\ln\left(\frac{p_x\left(X^-_{x_i},Y^-_{x_i}\right)}{p_u\left(X^-_{x_i},Y^-_{x_i}\right)}\cdot\frac{p_u\left(X^-_{x_i}\right)}{p_x\left(X^-_{x_i}\right)}\right)\right]. \tag{10}$$

Equation 10 shows that the TER depends on the probabilities of the process histories $X^-_{x_i}$ and $Y^-_{x_i}$, evaluated at target events and at arbitrary time points (respectively, p_x and p_u), whose statistical average is taken only at target events (i.e., over p_x). The last expression constitutes the basis for the MIR estimation strategy presented in the next subsection.

2.3 Practical Estimation

The approach for MIR estimation, devised in (Mijatovic et al., 2021a; Shorten et al., 2021) and briefly presented in the following, relies on creating history embeddings that cover the past states of the two observed point processes, implementing an operational formulation of **Eq. 10** to estimate the TER, and finally using **Eq. 8** to obtain the MIR estimate.

In the estimation of the TER from the source process Y to the target process X, the procedure for building history embeddings approximates the past history of the two processes observed either at the times of target events $x_i = 1, \ldots, N_X$, or at arbitrary time points $u_i = 1, \ldots, N_U$, sampled in continuous time. In the first case, illustrated in **Figure 1A**), the history embedding of the target X referred to the event x_i is approximated by taking l inter-event intervals, i.e., $X^-_{x_i} \approx X^l_{x_i} = \{x_{i-k+1} - x_{i-k}, k = 1, \ldots, l\}$; the history embedding of the driver Y referred to x_i is approximated as $Y^-_{x_i} \approx Y^l_{x_i} = [x_i - y_p, Y^{l-1}_{y_p}]$, where y_p is the most recent driver event preceding x_i. In the second case (**Figure 1B**), the histories of both processes as observed from u_i are approximated by taking the interval from the most recent event to u_i followed by $l-1$ inter-event intervals, i.e., $X^-_{u_i} \approx X^l_{u_i} = [u_i - x_p, X^{l-1}_{x_p}]$, $Y^-_{u_i} \approx Y^l_{u_i} = [u_i - y_p, Y^{l-1}_{y_p}]$.

The history embeddings are then used to compute the entropy terms that compose the TER computed according to **Eq. 10**. Specifically, **Eq. 10** can be expressed as

$$\hat{\dot{T}}_{Y\to X} = \bar{\lambda}_X\left[\hat{H}_{p_u}\left(X^l_{x_i},Y^l_{x_i}\right) - \hat{H}_{p_x}\left(X^l_{x_i},Y^l_{x_i}\right) + \hat{H}_{p_x}\left(X^l_{x_i}\right) - \hat{H}_{p_u}\left(X^l_{x_i}\right)\right], \tag{11}$$

where the estimates of the four entropies on the r.h.s. are obtained by approximating the past histories of infinite duration with the l – dimensional history embeddings, and computing the nearest neighbor entropy estimator (Vicente et al., 2011; Faes et al., 2015). Specifically, the terms $\hat{H}_{p_x}(\cdot)$ and $\hat{H}_{p_u}(\cdot)$ respectively refer to 'standard' differential entropy estimates where expectation is taken over the same probability distribution for which the log-likelihood is estimated, and to 'cross-entropy' estimates where the two distributions differ (a detailed procedure is given in Shorten et al. (2021); Mijatovic et al. (2021a)). The entropies are then estimated via the kNN estimator (Kozachenko and Leonenko, 1987), where the parameter k indicates the number of points used for searching the neighbors of each reference point; here, points are realizations of the history embeddings of dimension l or $2l$ specified in **Eq. 11**, and the search for neighbors is performed within the set of realizations taken at target events in the case of 'standard' entropy estimation, and within a set of realizations observed at arbitrary (randomly sampled) time points in the case of 'cross-entropy' estimation. The estimation algorithm, which is described in details in (Mijatovic et al., 2021a; Shorten et al., 2021), proceeds performing neighbor searches and range searches optimized to estimate together the four entropy terms in **Eq. 11**, in order to achieve compensation of the bias brought by the individual terms to the overall TER estimate. The TER from X to Y is estimated in the same way after reversing the role of the two point processes, and finally the MIR estimate is obtained by simply summing the two TER estimates in **Eq. 8**.

2.4 Corrected Measure of Mutual Information Rate

In this work, we face the issue of estimating the MIR from short realizations of coupled point processes. As any estimate of a measure computed on finite-length realizations of a process, the

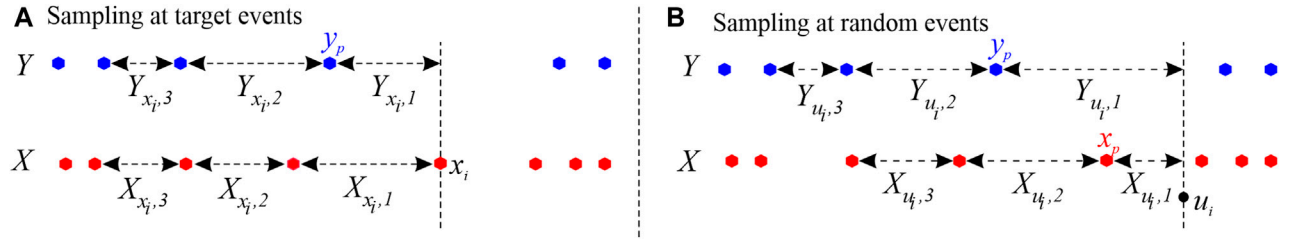


FIGURE 1 | Example of history embeddings used to approximate the past states of a target point process *X* and a source process *Y* described by the event times depicted as red and blue dots, respectively. In this example, embeddings are reconstructed with an embedding length $l = 3$.

MIR exhibits bias and variance which typically depend on the system dynamics, the analysis parameters, and the time-series length. While the variance reflects random errors which cannot be corrected, the bias of an estimator is related to systematic errors that can be compensated by knowing the true value of the measure of interest and its average value computed over several repetitions of the analyzed process. However, unfortunately the true theoretical values are generally not known for the MIR of coupled point processes, as analytical results do not exist for the sampling distribution of kNN estimates of entropy quantities. Therefore, here we resort to an empirical procedure that follows previously proposed approaches using surrogate time series to reduce the bias of information-theoretic estimates (Marschinski and Kantz, 2002; Papana et al., 2011). Specifically, first we estimate the bias of the estimator computing its average over several realizations of uncoupled surrogate event series for which the expected MIR is zero, and then we use such average value to correct the MIR estimated on the original coupled processes. While this approach can be theoretically justified as a full correction of the bias only when the true coupling between the processes is zero, it has been shown to provide a reasonable compensation of the bias of coupling and causality measures even for coupled processes (Papana et al., 2011).

The correction procedure adopted in this work is based on the generation of surrogate time series that preserve the individual dynamics of a process while destroying any correlation between pairs of processes. While surrogates are typically used to set a significance threshold in the estimate of coupling measures (Faes et al., 2004; Lancaster et al., 2018), in our approach we do not apply a formal surrogate data test but rather correct the MIR for the bias estimated in the absence of coupling. To do this, after computing the MIR estimate $\hat{I}_{X;Y}$ for a given realization of two point processes, we generate M surrogate point processes, estimate the MIR over each surrogate pair, and finally compute the corrected MIR (cMIR) as

$$\hat{I}^{(c)}_{X;Y} = \hat{I}_{X;Y} - \hat{I}^{(m)}_{X;Y}, \tag{12}$$

where $\hat{I}^{(m)}_{X;Y}$ is the median of the MIR estimated over the M surrogate pairs; we use the median instead of the mean to consider possible deviations of the MIR values from a symmetric distribution. The use of the corrected measure **Eq. 12** aims at reducing the bias of MIR in the case of absence of coupling between the two analyzed processes. To generate surrogate data, we adopted the procedure proposed by Shorten et al. (2021) in the context of TER estimation. This procedure implements a local permutation of the patterns forming the history embeddings for the two processes under the null hypothesis of independence of the present of the target and the history of the source given the history of the target. This null hypothesis is related to a more conservative test than that typically performed in TER/MIR estimation; while standard shuffling procedures destroy any relation between the current and past states of the target and the past states of the source, the local permutation test maintains the relation between the target and source histories, by decoupling only the source histories from the target events (Shorten et al., 2021). Nevertheless, to test this approach in comparison with established methods for the generation of surrogate data, we also implemented the algorithm based on random shuffling of the inter-event intervals, which preserves the probability distribution of the series of inter-event intervals; the iterative amplitude-adjusted Fourier transform (IAAFT) procedure (Schreiber and Schmitz, 1996; Perinelli et al., 2020), which preserves both distribution and power spectrum of the intervals; and the JODI algorithm (Ricci et al., 2019; Perinelli et al., 2020), which is specifically designed to preserve amplitude distribution and inter-event autocorrelation in point process data.

In all simulations and real data analyses, we implemented the nearest neighbor entropy estimator by using $k = 30$ neighbors and the maximum norm to compute distances (Faes et al., 2015), and generating a number of random time points equal to the number of target events ($N_U = N_X$) (Mijatovic et al., 2021a). Analyses were repeated varying the length of the history embedding in the range $l \in \{1, 2, 3, 4, 5\}$. In the simulation study, the dependence of MIR and cMIR on the coupling parameter, type of distribution of the inter-event intervals, and time series length was also analyzed.

3 SIMULATION STUDY

This section reports the application of the proposed method for continuous-time estimation of the MIR on point processes simulated according to three scenarios. The first is devised to assess the bias of the MIR estimate on pairs of independent point processes for different types of inter-event distribution and distribution parameters. In the second and third simulation, coupled point processes designed to mimic the conditions of the real-data application relevant to cardiovascular variability

reported in **Section 4** are considered; specifically, the dynamics of the heartbeat times and of the arrival times of the blood pressure wave in the body periphery are reproduced, and the two processes are coupled in a way such that the intensity of their interaction increases or decreases depending on different driving mechanisms modulated by the input simulation parameter.

3.1 Simulation Design

3.1.1 Simulation 1

In the first simulation, we generate pairs of uncoupled point processes according to different distributions. We consider: 1) Poisson processes, for which the inter-event intervals are i. i.d. exponential random variables with mean $1/\lambda_P$, where λ_P is the mean event-rate, here varied in the set $\lambda_P \in \{1, 2, 3, 4, 5\}$ events/s; 2) point processes with i. i.d. inter-event intervals taken from the Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$, with mean varied in the set $\mu \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$ s and standard deviation varied in the set $\sigma \in \{0.2, 0.4, 0.6, 0.8, 1.0\}$ s; 3) point processes with i. i.d. inter-event intervals taken from the inverse Gaussian distribution IG (μ, λ), with mean varied in the set $\mu \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$ s and shape parameter varied in the set $\lambda \in \{500, 600, 700, 800, 900\}$ s; 4) point processes with identically distributed history-dependent inter-event intervals taken from the inverse Gaussian distribution, HDIG (μ, λ, θ), where $\mu \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$ and $\lambda \in \{500, 600, 700, 800, 900\}$ are the mean and shape parameters of an IG distribution, respectively, and θ is a vector of parameters that sets the correlations between the inter-event intervals of each process, and makes them it history-dependent (the values of the parameters in θ are described in **Section 3.1.2**).

While the first two distributions are typically used in the simulation of point processes, the IG and HDIG distributions are considered as they constitute the basis for a model that reproduces realistic heartbeat dynamics as presented in the following **Section 3.1.2**. For all these classes of point processes, the ground truth value of the information exchanged dynamically between the two processes is zero ($\dot{I}_{X;Y} = \dot{T}_{X\to Y} = \dot{T}_{Y\to X} = 0$) because the processes are obtained from independent runs of the simulation. This allows to quantify the bias of the adopted MIR estimator, which is equivalent to the median value of the MIR estimated across several realizations of each simulation. In a single simulation, 100 realizations of each pair of uncoupled point processes were generated, each consisting of $N = 300$ events, and the distribution of the MIR measure was computed for each combination of the simulation parameters in the four cases described above.

3.1.2 Simulation 2

In the second simulation the process X, which reproduces the heartbeat times, is generated as a point process following the history-dependent inverse Gaussian (HDIG) model proposed by Barbieri et al. (2005). According to this model, given any event x_i that simulates the occurrence time of a heartbeat, the waiting time until the next event, i.e. the ith inter-event interval w_i, is assumed to be drawn from the probability density function

$$p\left(w_i, X_{x_i}^p, \theta, \lambda\right) = \sqrt{\frac{\lambda}{2\pi w_i^3}} \cdot e^{\frac{-\lambda\left[w_i-\mu\left(X_{x_i}^p,\theta\right)\right]^2}{2\mu\left(X_{x_i}^p,\theta\right)^2 w_i}}, \tag{13}$$

where $\mu(X_{x_i}^p, \theta)$ and λ are the mean and the scale parameter of the inverse Gaussian distribution. In the HDIG model, the mean is dependent on the history of the inter-event intervals up to the current event x_i, $X_{x_i}^p = [w_{i-1}, \ldots, w_{i-p}]$, according to the linear autoregressive (AR) model:

$$\mu\left(X_{x_i}^p, \theta\right) = \theta_0 + \sum_{j=1}^{p} \theta_j w_{i-j}. \tag{14}$$

This model represents, through the parameter vector $\theta = (\theta_0, \theta_1, \ldots, \theta_p)$, the dependence of the present inter-event interval on the past history of the process, and in this application accounts for autonomic influences on heart rate variability (Stein et al., 1994). The setting of the model parameters is performed to reproduce typical point-process patterns of heart rate variability and cardiovascular interactions (Faes et al., 2014; Beda et al., 2017). Specifically, in our simulation we assume that the inter-event intervals exhibit lagged dependencies up to the order $p = 5$, and we set the coefficients $\{\theta_1, \ldots, \theta_5\}$ to obtain oscillations of w_i within the very low frequency (VLF, < 0.04 Hz), low frequency (LF, 0.04–0.15 Hz) and high frequency (HF, 0.15–0.4 Hz) bands, as typically observed in the time series of heart period variability (Stein et al., 1994). This is achieved by simulating for the AR model (14) a transfer function with two complex-conjugate poles with modulus $\rho_{\mathrm{LF}} = 0.8$ and phases $\pm 2\pi \cdot 0.1$ rad, two other complex-conjugate poles with modulus $\rho_{\mathrm{HF}} = 0.92$ and phases $\pm 2\pi \cdot 0.25$ rad, and a real pole with modulus $\rho_{\mathrm{VLF}} = 0.6$ (Beda et al., 2017). The mean and scale parameters of the inverse Gaussian distribution are set to $\theta_0 = 1$ s (average heart period) and $\lambda = 600$ s.

After generating the heartbeat point process X as described above, the point process Y that simulates the blood pressure arrival times is obtained generating its events as

$$y_i = x_i + \tau_i, \tag{15}$$

where each propagation delay τ_i simulates an instance of the pulse arrival time (PAT), i.e. the time interval between the initiation of a cardiac contraction (identified by the electrical depolarization of the ventricles) and the following time of arrival of the blood pressure wave at the body periphery (identified by the time of maximum finger arterial pressure). The propagation delays are modelled as realizations of a second-order AR process defined as

$$\tau_i = a_0 + a_1\tau_{i-1} + a_2\tau_{i-2} + u_i, \tag{16}$$

where a_0 represents the mean PAT set to 300 ms to reproduce the average propagation time of the sphygmic wave from the heart to the body periphery; a_1 and a_2 were set to reproduce a stochastic oscillation at ~ 0.1 Hz by using a transfer function with two complex-conjugate poles with modulus $\rho_{\mathrm{LF}} = 0.8$ and phases $\pm 2\pi \cdot 0.1$ rad, and u_i are random numbers taken from a Gaussian distribution with zero mean. The standard deviation of u_i was adjusted to obtain specific values for the standard deviation of τ_i, which we denote as σ_{PAT}. This important parameter modulates

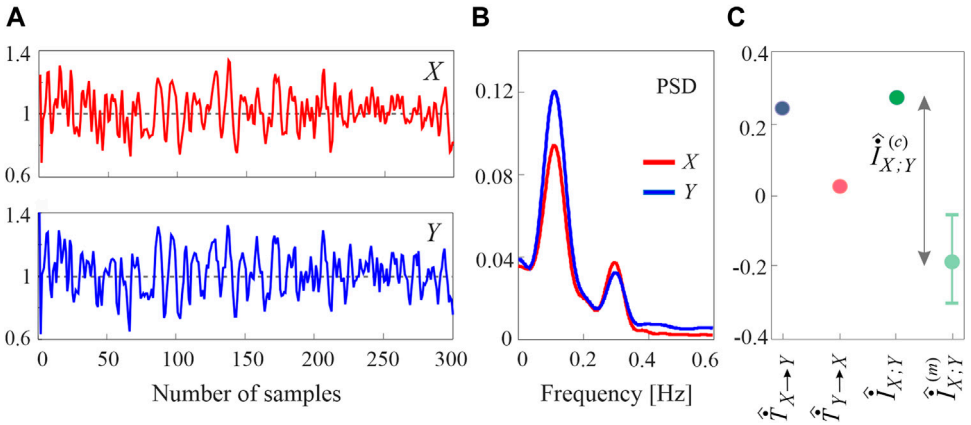


FIGURE 2 | Representative example of the analysis relevant to the second simulation. **(A)** Inter-event intervals obtained for the process X as observations of the history-dependent inverse Gaussian model of **Eqs. 13, 14** and for the process Y as observations of the process described by **Eqs. 15, 16** generated with σ_{PAT} =85 ms; **(B)** power spectral densities of the two inter-event series evidencing LF and HF oscillations at ~ 0.1 Hz and ~ 0.25 Hz; **(C)** corresponding estimates of the TER from X to Y (blue circle) and from Y to X (red circle), of the MIR obtained as the sum of the two TERs (green circle), and distribution (median and percentiles) of the MIR estimated from 100 surrogate event series (light green); the difference between the MIR and the median of its distribution on the surrogates corresponds to the bias-corrected cMIR (gray arrow).

the variability of the arrival times y_i, and in this simulation is inversely related to the strength of the interaction from X to Y; here, σ_{PAT} was varied from 10 to 235 ms with steps of 25 ms.

The inter-event intervals of the simulated heartbeat and blood pressure timings generated by a run of the simulation 2 are reported in **Figure 2A**) along with the respective power spectral densities (PSD, shown in **Figure 2B**), which evidence VLF, LF and HF oscillations in the two processes. The values of the TER estimated along the two directions of interaction, the MIR estimated as the sum of the two TERs, as well as the distribution of the MIR estimated from 100 surrogate series and the corresponding cMIR, are displayed in **Figure 2C**).

3.1.3 Simulation 3

The third simulation is a modification of the second one and is devised to impose a common oscillation in the inter-event intervals of the process X and in the propagation delays τ_i, so as to reproduce a condition in which the same underlying mechanism drives the two point processes. To this end, the HDIG model is retained to simulate the heartbeat intervals as in **Eq. 13**, but with different autocorrelation structure; specifically, an AR model of order $p = 3$ was used in **Eq. 14**, with coefficients $\{\theta_1, \theta_2, \theta_3\}$ set to obtain oscillatory activity within the VLF and HF bands only (i.e., using a transfer function with one real pole with modulus $\rho_{HF} = 0.92$ and two complex conjugate poles with modulus $\rho_{HF} = 0.92$ and phases $\pm\ 2\pi \cdot 0.25$ rad). Starting from the intervals w_i drawn from this HDIG distribution with VLF and HF components, the LF component is introduced by adding to w_i a term equal to $2\tau_i$, where τ_i is the random interval generated by **Eq. 16**. The simulation is then completed as in the previous case, i.e., by generating blood pressure arrival times as in **Eq. 15** with propagation delays given again by **Eq. 16**. In this way, the LF component of the inter-event intervals in X and the propagation delays that contribute to the LF variability of Y are generated from the same random seed u_i and, as a consequence, the parameter σ_{PAT} that determines the variability of both components directly modulates the coupling between the two processes (i.e., we expect that higher values of σ_{PAT} determine higher amounts of information shared between X and Y).

3.2 Simulation Results

In the first simulation, the MIR computed according to **Eq. 8**, where the two TER terms are estimated as in **Eq. 11**, was evaluated in pairs of uncoupled point processes by varying the type of inter-event interval distribution of the processes and the distribution parameters. Since for these processes the true value of the index is $I_{X;Y} = 0$, the values of the MIR estimate $\hat{I}_{X;Y}$ highlight the bias of the estimator. The results reported in **Figure 3** indicate the presence of a negative bias in all simulations, as documented by the negative values of $\hat{I}_{X;Y}$ measured by varying the type and parameters of the distribution of the uncoupled processes.

For Poisson processes, the bias tends to increase with the event rate and with the mismatch between the rates of the two processes (**Figure 3A**). For Gaussian processes, the bias increases when the standard deviation of the inter-event intervals is decreased, and is not substantially affected by the mean (**Figure 3B**). In the case of uncorrelated inverse Gaussian inter-event intervals, the bias is inversely related both to the mean and to the shape parameter of the interval distribution (**Figure 3C**); the dependence on the shape parameter becomes direct when the inverse Gaussian intervals are correlated in HDIG processes (**Figure 3D**). Overall, these results indicate that, in the presence of short realizations of point processes as in the present case where $N = 300$ spikes are simulated, the MIR estimates are strongly biased, and therefore strategies are needed for the compensation of such bias in the practical analysis of the information shared between point processes.

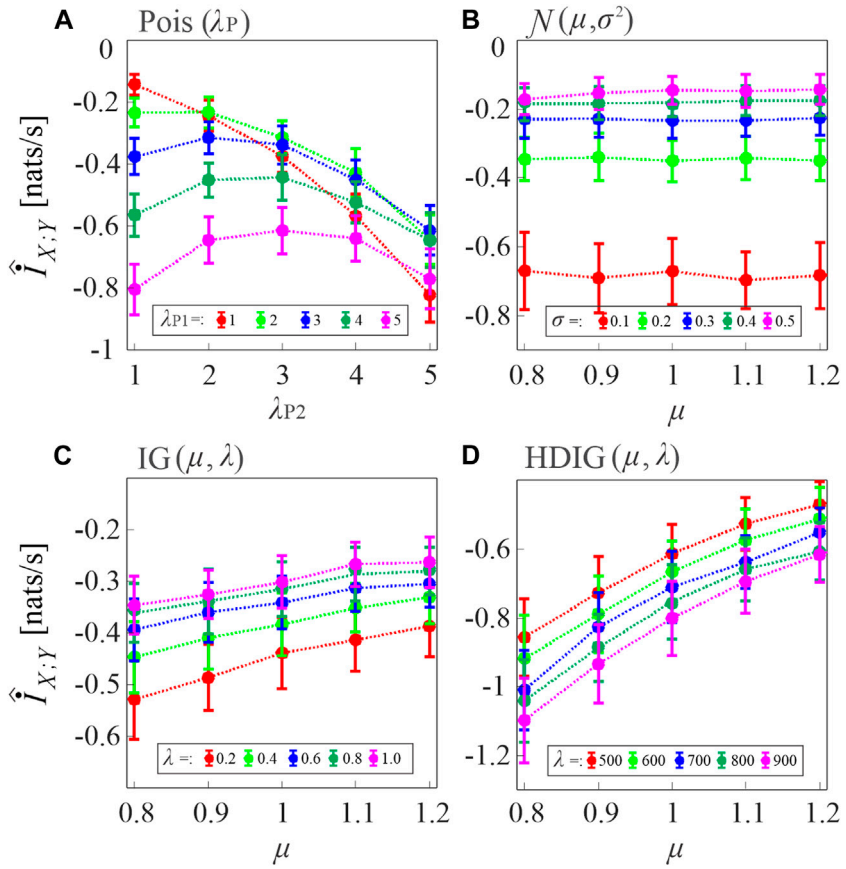


FIGURE 3 | Assessment of the bias of the proposed MIR estimator. Plots depict the distribution (mean ± SD) of the MIR values computed over 100 realizations of uncoupled, short-length processes (N = 300 samples) with inter-event intervals taken from an exponential distribution with parameter λ (**(A)**, Poisson processes), a Gaussian distribution with mean μ and variance σ^2 **(B)**, an inverse Gaussian (IG) distribution with mean μ and shape parameter λ **(C)**, and a history-dependent inverse Gaussian (HDIG) distribution with mean μ and shape parameter λ **(D)**. The history embedding length was set to l = 1 in all computations.

The procedure for compensating the bias of MIR estimates, as well as the performance of the corrected cMIR estimator, are illustrated in **Figure 4** for several runs of simulation 2 generated by varying the intensity of the interaction between the HDIG processes modulated by the parameter σ_{PAT}.

As shown in **Figure 4A**, the progressive de-coupling of the interactions between X and Y obtained by increasing σ_{PAT} is reflected by a progressive decrease of the MIR estimates; this behavior is observed for all the analyzed values of the history embedding length l. However, the analysis also confirms the presence of a substantial bias in the estimates of MIR, which take on negative values when the coupling between the two processes decreases. **Figure 4B** reports the bias-corrected MIR estimates, showing how the correction leads to non-negative values of cMIR even when the processes approach the uncoupled states for high values of σ_{PAT}. The correction brings the cMIR values in the range $0 - 0.6$ nats/s for $l = 1$, which extends to ~ 0.7 nats/s for $l = 5$, and evidences the appropriateness of using higher embedding lengths in the simulated process when the inter-event intervals are modeled by an AR model of order $p = 5$.

The benefit of longer history embeddings is documented also in **Figure 4C**, where we employ the standard procedure for testing coupling significance based on surrogate data. This procedure tests the null hypothesis of uncoupling between the two analyzed point processes and is based on generating, from each pair of original realizations of the processes, a suitable number of pairs of surrogate event series using the local permutation method, and then on deeming the original pair as significantly coupled if the MIR value was above the 95th percentile of the MIR surrogate distribution. The percentage of realizations for which the MIR/cMIR values were detected as statistically significant is reported in **Figure 4C**, showing that the rate of detection of weakly coupled point processes (higher values of σ_{PAT}) increases for higher embedding lengths.

Figure 5 has the same structure of **Figure 4A**, and shows alternative approaches to generate the surrogate data consistent with the null hypothesis of uncoupling between the two analyzed point processes. The figure shows that the analysis of cMIR is rather stable at varying the type of surrogate data. The most remarkable difference is that using the shuffling surrogates, similarly to the local permutation surrogates employed in **Figure 4A** even though with a lower extent, the MIR estimates partially overlap with the distribution of the MIR for the original process realizations when the de-coupling parameter is high

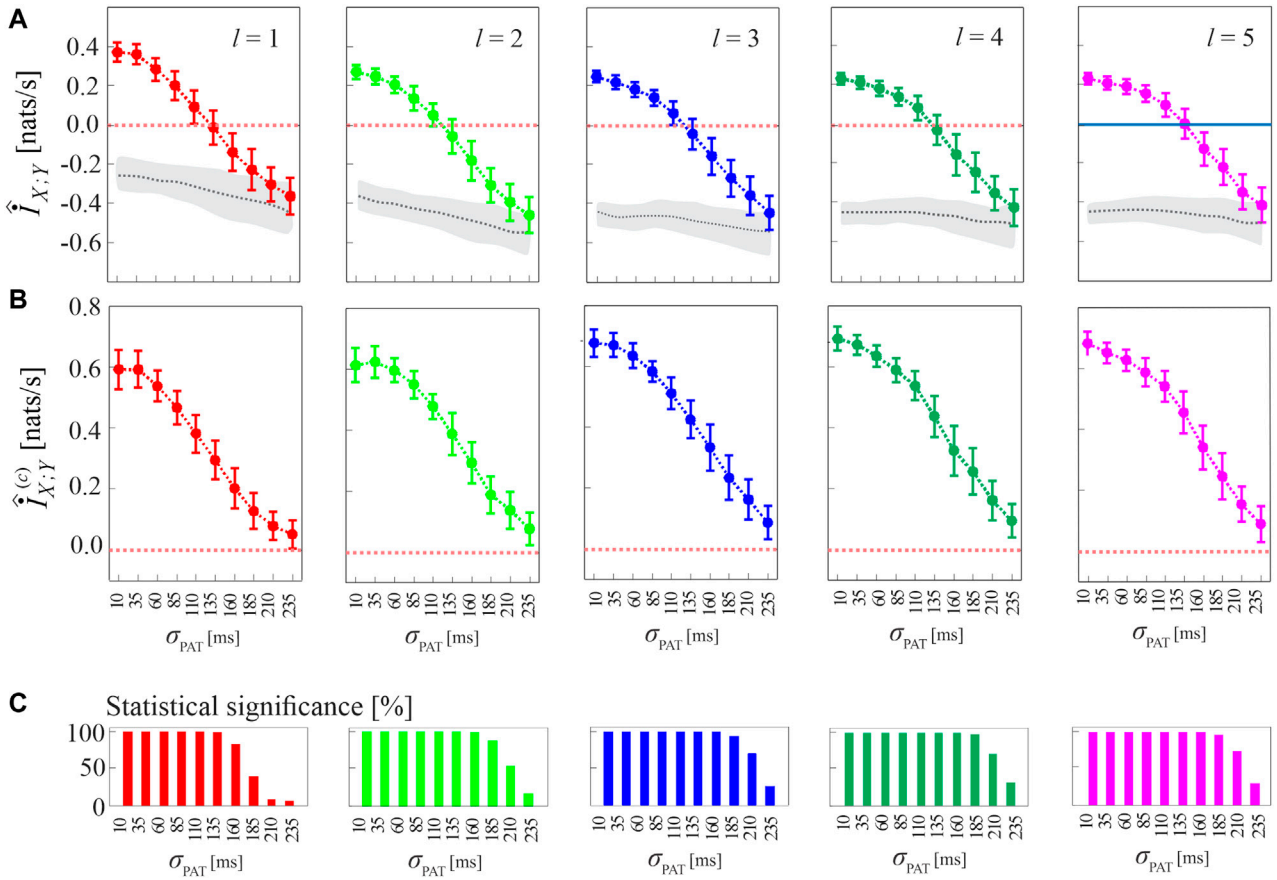


FIGURE 4 | Computation of MIR and cMIR, and assessment of their statistical significance, in simulations of short-length (N =300 events) coupled history-dependent inverse Gaussian (HDIG) processes. **(A)** Distribution (mean ± SD) of the MIR measure, estimated over 100 realizations of simulation 2 as a function of the de-coupling parameter σ_{PAT}, for different values of the history embedding length, $l \in [1, 5]$; gray dotted lines and shades correspond to the median and 5^{th} – 95^{th} percentiles of the distribution over the 100 realizations of the median MIR (for each realization, the median is computed over 100 surrogate pairs obtained using local permutation). **(B)** Distribution (mean ± SD) of the corrected MIR (cMIR) measure, estimated over the same realizations of simulation 2; for each realization, cMIR is computed as the difference between MIR and the median of the MIR distribution assessed over 100 local permutation surrogates. **(C)** Bar plots reporting the number of realizations for which the MIR is detected as statistically significant according to the surrogate data analysis.

(σ_{PAT} = 210 ms and particularly σ_{PAT} = 235 ms); such an effect is not observed using IAAFT and JODI surrogates. This suggests that surrogates which preserve autocorrelation properties of the inter-event intervals are more prone to detect weak but significant amounts of information shared by two point processes and to return higher values of the cMIR measure. On the other hand, the use of the local permutation method for the generation of surrogate time series resulted in higher values of the MIR assessed on the surrogates (see the gray areas in **Figure 4A** vs those in **Figure 5**). This result is expected, as the local permutation method maintains the relationship of the source history embeddings with the history embeddings of the target, thus allowing to keep a low rate of false positive detection of information transfer (Shorten et al., 2021). Thus the comparison between **Figures 4A**, **5** evidences that the local permutation surrogates adopted as a main solution in our work tends to favor specificity in the detection of coupled point process dynamics, while surrogates preserving autocorrelation structure of the inter-event intervals tend to favor sensitivity. These considerations are of practical relevance for the analysis of real-world data.

To show that the bias of the MIR estimates is due to the small sample size of the point process realizations analyzed, in **Figure 6** we show the MIR computed as a function of the decoupling parameter σ_{PAT} for different lengths of the simulated processes, $N \in \{150, 300, 1{,}000, 5{,}000, 10{,}000\}$, together with the cMIR obtained using either the local permutation method or the JODI algorithm to generate surrogate point processes. We observe that increasing the number of simulated events progressively reduces the bias, as documented by the progressively higher values observed for the MIR and by the absence of negative values for $N \geq 5{,}000$. As expected, also the variance of the MIR estimates decreases while increasing N, confirming that larger sample sizes reduce not only the bias, but also the variability of the estimates. We also note that the median of MIR over the surrogate distribution (gray dotted line in **Figures 6A,C**) is not a

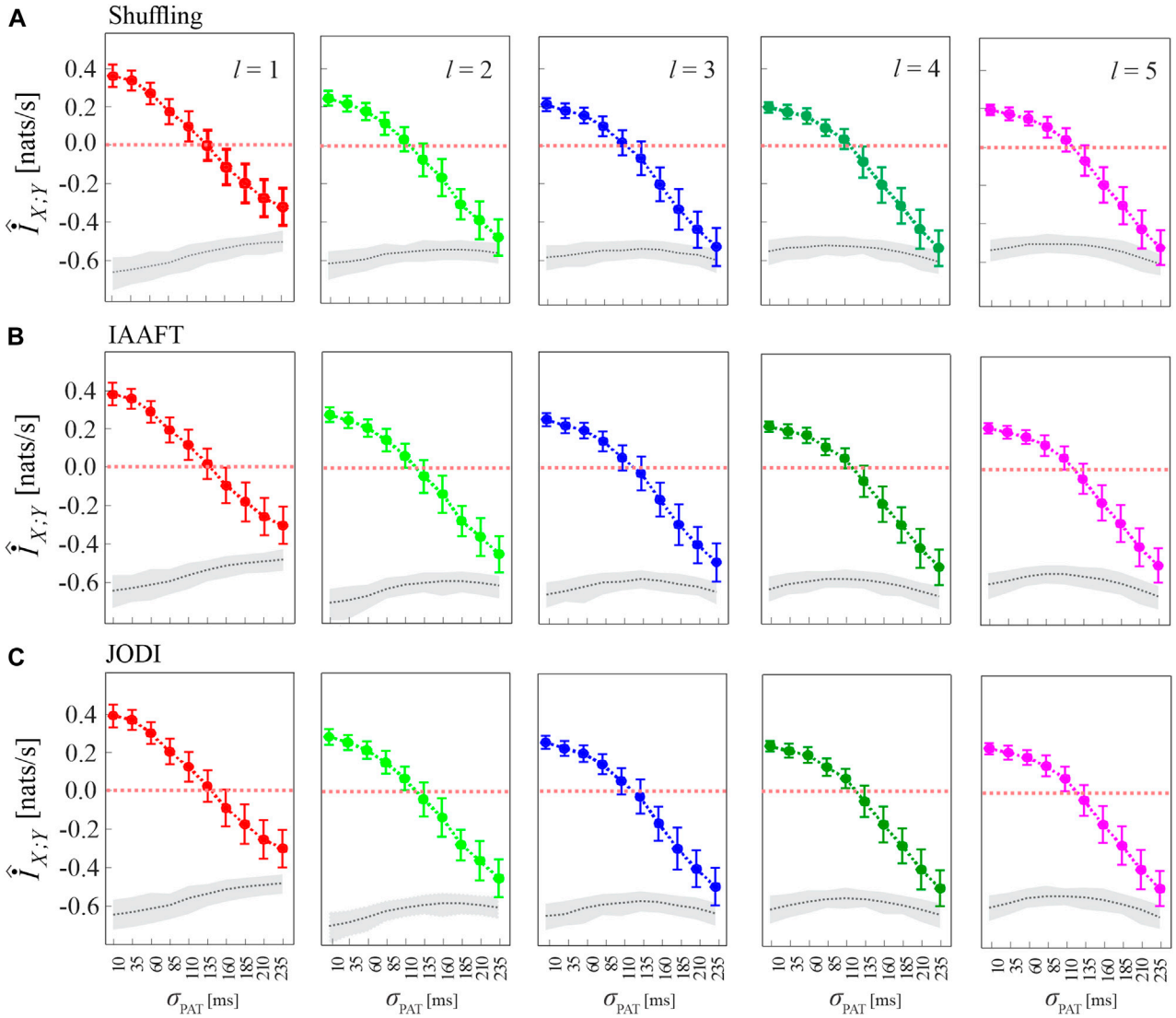


FIGURE 5 | Analysis of MIR using different methods for surrogate data generation. Plots depict the distribution (mean ± SD) of the MIR measure, estimated over 100 realizations of simulation 2 as a function of the de-coupling parameter σ_{PAT}, for different values of the history embedding length, $l \in [1, 5]$. Gray dotted lines and shades correspond to the median and 5^{th} – 95^{th} percentiles of the distribution over the 100 realizations of the median MIR; for each realization, the median is computed over 100 surrogate pairs generated by random shuffling of the inter-event intervals **(A)**, according to the IAAFT algorithm **(B)**, and according to the JODI algorithm **(C)**.

constant function of σ_{PAT} and differs for the two methods for surrogate generation. As a consequence, the cMIR does not represent a simple translation of MIR toward positive values and depends on the adopted surrogates. In particular, the use of local permutation surrogates results in lower values of cMIR compared to that based on JODI surrogates (see **Figures 6B,D**), suggesting a better bias compensation for the latter approach. Moreover, the non-monotonic behavior of MIR estimated for small sample size (N = 150 and N = 300) is accentuated in cMIR when local permutation surrogates are used (**Figures 6A,B**), while it is smoothed when JODI surrogates are used (**Figures 6C,D**).

Figure 7 reports the results of Simulation 3, where coupled HDIG processes are generated so that increasing the variability of the propagation delay from X to Y may also increase the coupling between the two processes. This effect is verified in our simulations by observing that the cMIR measure increases with the parameter σ_{PAT}, which in this case modulates the variability of both the LF component of the inter-event intervals of X and the propagation delays; the increase of the information shared between the two processes at increasing σ_{PAT} is observed consistently for all the analyzed history embedding lengths, $l \in [1, 5]$.

4 APPLICATION TO REAL DATA

This section describes the application of cMIR on experimental point-process data relevant to cardiovascular variability. In the information-theoretic domain, cardiovascular interactions are

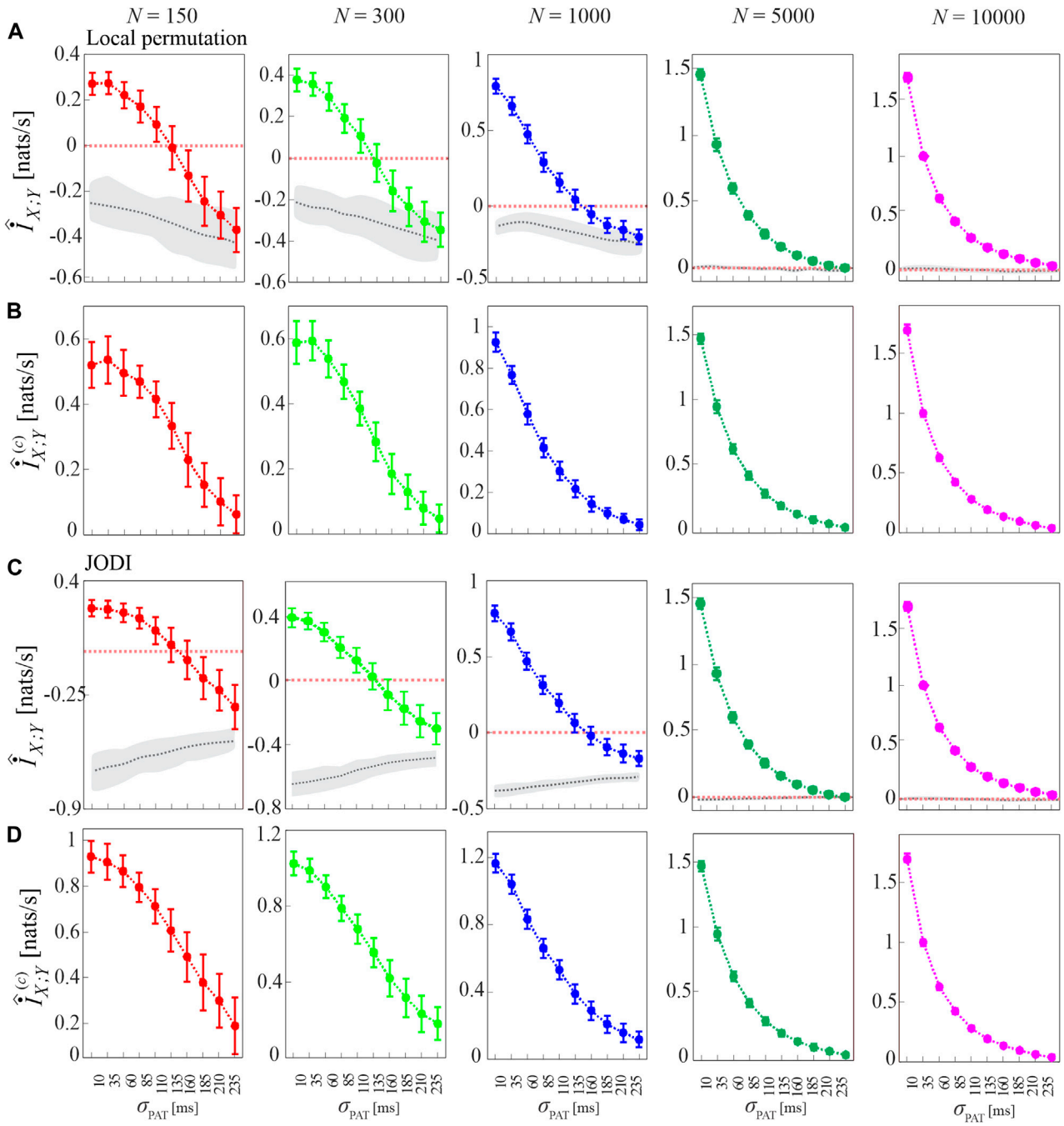


FIGURE 6 | Dependence of MIR and cMIR on the size of the analyzed point processes and on the type of surrogate series used for bias compensation. Plots depict the distribution (mean ± SD) of the MIR **(A,C)**, and the cMIR based on local permutation surrogates **(B)** or JODI surrogates **(D)**, estimated (history embedding length $l = 1$) as a function of the decoupling parameter σ_{PAT} over 100 realizations of simulation 2 of different lengths (number of simulated events $N \in \{150, 300, 1000, 5000, 10\,000\}$). In panels **(A,C)**, gray dotted lines and shades correspond to the median and 5^{th} – 95^{th} percentiles of the distribution over the 100 realizations of the median MIR, where for each realization the median is computed over 100 surrogate pairs obtained by using local permutation surrogates **(A)** or JODI surrogates **(C)**.

commonly studied by means of entropy measures applied to the discrete-time series of heart period and arterial pressure variability (Faes et al., 2012; Faes et al., 2015; Javorka et al., 2017). However, given the intrinsic unevenly sampled nature of human heartbeats (Barbieri et al., 2005), recent studies started to face the analysis of cardiovascular, cardiorespiratory and brain-heart dynamics from the perspective of point processes analyzed also using information measures (Valenza et al., 2018; Greco et al., 2019). Here, with the aim of assessing the potential of MIR analysis in short-term cardiovascular variability as well

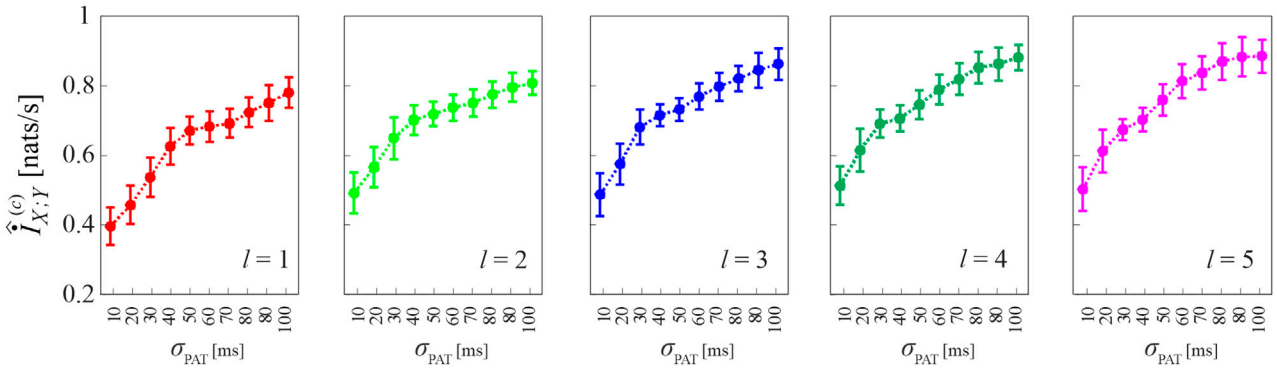


FIGURE 7 | Computation of cMIR for short-length realizations (N =300 events) of simulation 3. Plots depict the distribution (mean ± SD) of the cMIR measure, estimated over 100 realizations of simulation 3 as a function of the parameter σ_{PAT}, for different values of the history embedding length, $l \in [1,5]$. Note that in this simulation in which common oscillations are imposed in the variability of the inter-event intervals of the process X and on the propagation delay from X to Y, σ_{PAT} serves as a coupling parameter.

physiological mechanisms other than those investigated by the traditional information-theoretic measures, we apply our continuous-time approach on the point processes that map the heartbeat and systolic time events measured in healthy humans and monitored under different physiological states.

4.1 Database and Experimental Protocol

The analyzed data belong to an historical database previously used to study the effects of physiological stress and cognitive workload on cardiovascular variability (Javorka et al., 2017; Pernice et al., 2019). The data were acquired on 76 young healthy subjects (age: 18.4 ± 2.7 years, 32 males), normotensive and with a normal body mass index (21.3 ± 2.3 kg/m^2), and consisted of electrocardiographic (ECG) and blood pressure (BP) recordings acquired synchronously with a sampling frequency of 1 kHz. ECG and BP signals were recorded by using CardioFax ECG-9620 (Nihon Kohden, Japan; horizontal bipolar thoracic leads) and the Finometer Pro devices (FMS, Netherlands; volume-clamp continuous BP measurement), respectively. The experimental protocol foresaw the acquisition of the signals in different physiological states, going from resting conditions to different types of stress (orthostatic or mental). For the analyses carried out in this work, we have taken into account the following states: 1) baseline state (B), with subjects resting in the supine position for 15 min; 2) head-up tilt state (T), obtained by passively tilting the subjects by 45° to the upright position and maintaining them in that state for 8 min in order to produce orthostatic stress; 3) mental arithmetic state (M), obtained with subjects in the supine position and by asking them to sum up as fast as possible 3-digit numbers projected on the ceiling until reaching a 1-digit number and to decide whether the resulting number was even or odd (PMT test, Psycho Soft Software, s. r.o., Brno, Czech Republic), where this task was repeated over a period of 6 min to elicit cognitive load. Further details on the experimental protocol can be found in (Javorka et al., 2017; Pernice et al., 2019).

4.2 Data Analysis

The data analyzed consisted of sequences containing the timings of the consecutive R peaks in the ECG (event series of the R times) and of the following maxima in the BP signals (event series of the systolic times), previously extracted by means of LabChart 8 (ECG analysis, blood pressure modules) toolbox from ADInstruments (Javorka et al., 2017; Pernice et al., 2019). Moreover, the time series of the RR and PAT intervals were measured respectively as the sequences of the difference between two consecutive R times, and of the difference between each systolic time and the preceding R time. The event series and time series analyzed for each subject and experimental condition consisted of N = 300 events, which were extracted starting respectively ~8 min after the beginning of the phase B, ~3 min after the beginning of the phase T, and ~2 min after the beginning of the phase M; the corresponding RR and PAT time series were checked for stationarity by using a test targeting a restricted form of weak stationarity (Magagnin et al., 2011).

Starting from the interval series of RR and PAT, the mean and standard deviation of the two series, respectively computed as the average interval duration and the interval variability, were computed for each subject and experimental condition. Starting from the corresponding event series of R times and systolic times, the intervals forming the history embeddings were extracted as displayed in **Figure 1**, and employed as described in **Section 2** to estimate first the TER along the two directions of interaction, then the MIR, and finally the cMIR. To test the statistical significance of the differences in the median of the distributions of each measure (mean, standard deviation and cMIR) evaluated across conditions (B, T, M), we used the non-parametric Kruskal-Wallis test, followed by post-hoc paired Wilcoxon signed rank test to assess pairwise differences (B vs. T, B vs. M, T vs. M) with 5% significance and employing the Bonferroni-Holm correction for multiple comparisons.

4.3 Results and Discussion

The results of the real data analysis are summarized in **Figure 8**, reporting the distributions of the basic statistics (mean and

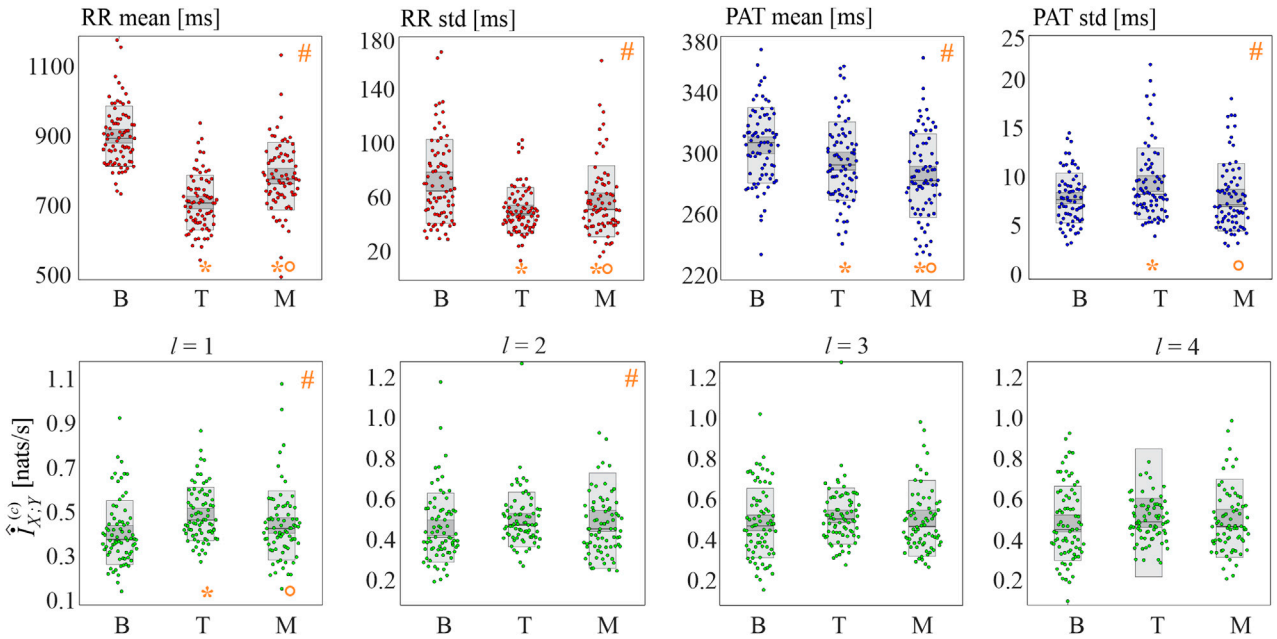


FIGURE 8 | Basic statistics and information shared in the cardiovascular time series. Panels depict the boxplots and individual values of the mean and standard deviation of the RR intervals (red dots) and of the PAT intervals (blue dots), as well as of the cMIR measure estimated for different values of the history embedding length *l* (green dots), computed for all subjects during the three analyzed experimental conditions (baseline (B), head-up tilt (T), and mental arithmetic (M)). Statistical analysis (orange symbols): #, $p < 0.05$, Kruskal-Wallis test; *$p < 0.05$ B vs T or B vs M; °$p < 0.05$: T vs M, Wilcoxon test.

standard deviation of RR and PAT intervals) in the upper panels and of cMIR in the lower panels.

The mean RR interval decreased significantly moving from B to T and from B to M; the effect was more pronounced during tilt than during mental arithmetic. Similarly, both postural stress and mental stress induced a decrease of the variability of the RR intervals, with a larger effect during head-up tilt, as documented by the statistically significant decrease of the standard deviation of the RR intervals moving from B to T and from B to M and by its higher values during M compared to T.

The physiological stressors induced also statistically significant variations in the mean and variability of the propagation delays of the sphygmic wave from the heart to the periphery. Specifically, the mean PAT decreased progressively and significantly while moving from B to T and from T to M, and the standard deviation of PAT increased during T compared to B, and decreased during M compared to T.

The analysis of the cMIR measure indicated that the postural stress tends to increase the information shared between the R times and the systolic times, while mental stress does not have significant effects. In fact, cMIR was significantly higher during T compared to B, and significantly lower during M compared to T when a history embedding $l = 1$ was used. These variations were less evident when $l = 2$, as the Kruskal-Wallis test reported statistically significant differences among the three distributions despite the post-hoc tests did not reach statistical significance (B vs T, $p = 0.070$; B vs M, $p = 0.195$, T vs M, $p = 0.333$), and were reduced to non-significant trends when $l = 3$ and $l = 4$.

The alterations observed in the basic cardiovascular parameters during the two physiological stressors are in agreement with a large body of literature in cardiovascular variability analysis, and document the involvement of several physiological mechanisms in the elicitation of these stressors. In particular, the lower mean and variability of the RR intervals during tilt and mental arithmetic reflect well-known effects such as the tachycardia and the shift of the cardiac autonomic balance towards sympathetic activation and parasympathetic inhibition induced by postural and mental stress (Montano et al., 1994; Carnethon et al., 2002; Garde et al., 2002; Wood et al., 2002; Martinelli et al., 2005; Javorka et al., 2017, 2018; Kim et al., 2018; Pernice et al., 2019). The interpretation of the shortening of PAT and of the increase of its variability observed during tilt is less straightforward. The PAT is composed by the pre-ejection period (PEP), i.e., the interval from the electrical depolarization of the ventricles to the ejection of the blood from the heart, and by the pulse transit time (PTT), i.e., the time that it takes for the blood pressure wave to reach the body periphery; the PEP depends mainly on the strength of left ventricular contraction, influenced by the Frank-Starling law and by sympathetic control (Krohová et al., 2017), while the PTT is mostly affected by arterial compliance, reflecting (on a short time scale) modulation of blood pressure and vasomotion (Mukkamala et al., 2015; Czippelova et al., 2019). In accordance with our previous research in a related database (Krohová et al., 2017), we expect an increase in PEP during orthostasis as an effect of decreased diastolic filling of the heart via the Frank-Starling mechanism leading to a lower strength of the cardiac

contraction. Therefore, the decrease of the mean PAT observed during tilt should reflect mostly a decrease in PTT related to an augmented arterial stiffness caused by peripheral vasoconstriction, which is in turn evoked by the vascular baroreflex response associated with a decrease of blood pressure due to pooling of blood in the lower extremities (Czippelova et al., 2019); the concomitance of these opposite trends (i.e., increase of PEP and decrease of PTT) and the complexity of the related physiological mechanisms including autonomic reflexes and mechanical effects (Rapalis et al., 2017; Czippelova et al., 2019; Pernice et al., 2021) may - together with an increased systolic blood pressure variability associated with tilt - explain the higher variability of PAT observed during tilt. During cognitive load, induced in our protocol by the mental arithmetic task, the more prominent decrease of PAT likely reflects–in addition to vasoconstriction driven by commands stemming from the central nervous system which reduces the PTT–also a reduction of PEP associated with an increased cardiac contractility mediated by the sympathetic nervous system (Martin et al., 2016); in this case, the presence of common trends (i.e., decrease of both PEP and PTT) may explain both the lower PAT and its lower variability measured during mental arithmetic.

According to our results, the physiological mechanisms described above are associated with an increase of the rate of information exchange between the point processes marking the R times of the ECG and the times of arrival of the sphygmic wave in the body periphery. Higher values of MIR are expected when the variations of the propagation delay from one process to another are small, or when such variations occur in phase due to the effect of some common driver mechanism. Since we observe an increase in cMIR simultaneously with a shortening of the mean PAT and an increase of the PAT variability, we conclude that the presence of a common driver oscillation is the mechanism underlying the higher exchange of information. This mechanism was synthetically reproduced in our third simulation (see **Figure 7**), and can be physiologically explained by the sympathetic activation induced by head-up tilt (Montano et al., 1994; Carnethon et al., 2002; Martinelli et al., 2005). The "common driver" nature of this mechanism can be explained by observing that during postural stress the sympathetic activation is related to the baroreflex mechanism and, as such, it simultaneously involves the variability of the heart period (and thus that of the R times) and the variability of the arterial pressure (and thus that of the PAT) (Porta et al., 2011; Faes et al., 2013), thereby determining a more intense exchange of information between the two processes. In fact, vasoconstriction in the arterioles in systemic circulation is modulated almost exclusively by the sympathetic part of the autonomic nervous system (Krohova et al., 2020) whose oscillations mostly occur in the LF band; a similar effect is mimicked in our simulations in **Section 3**. On the other hand, the less evident variations of cMIR observed during the mental arithmetic test may be associated with the fact that the sympathetic activation evoked by mental stress is of a different type, likely involving central commands from the upper brain centers (cortex) which control more independently the heartbeat and the arterial compliance without prominent synchronization effects related to the baroreflex (Fauvel et al., 2000).

The observation of statistically significant differences across conditions of the cMIR index only for small values of the history embedding length (variations from B to T and from T to M are detected for $l = 1$ and, to a lower extent, for $l = 2$) suggests that the cardiovascular interactions altered by physiological stress occur mostly as a consequence of the variability of the propagation time of the sphygmic wave from the heart to the body periphery, and that the use of longer memory effects may confound the detection of such altered interactions. This result can be expected by considering that the largest part of the analyzed type of interactions is due to the PAT, whose effects are fully captured with $l = 1$ (note that, within the point process framework, effects explained with $l = 1$ are not immediate but rather indicative of time-lagged effects with short memory). The result is in agreement with previous observations reporting that the latency of cardiovascular information transfer is typically limited to zero-lag or one-beat interactions, especially during postural stress (Faes et al., 2014). Nevertheless, we remark that the type of cardiovascular interactions studied using time-series based methods (Faes et al., 2014; Porta and Faes, 2015) reflect different mechanisms than those reflected by the event-based method employed here, the former being related mainly to the baroreflex control of heart rate, while the latter being related to blood pulse propagation and arterial contractility.

5 CONCLUDING REMARKS

This study reports the first application to cardiovascular dynamics of the continuous-time estimator of the information exchanged dynamically between point processes introduced in (Shorten et al., 2021) to compute the TER and employed in (Mijatovic et al., 2021a) to compute the MIR. In the reported application context where the direction of interaction is determined by the cardiac pacemaker that triggers the propagation of the sphygmic waves through the arterial bed, studying causal interactions through the TER is less relevant than assessing the coupling between the heartbeat and systolic times through the MIR. Moreover, this application context is particularly challenging with regard to the computation of information rates, because the cardiovascular regulation operates mostly through short-term control mechanisms and needs to be performed over short stationary series including a few hundred heartbeats at most (Cohen and Taylor, 2002). The adopted estimator combines the property that for point processes the MIR can be formulated in terms of the TER (Mijatovic et al., 2021a), and exploits the approach based on representing dynamic states of point processes in terms of inter-event intervals to efficiently capture information flows (Shorten et al., 2021). In this work we investigate the small sample properties of the MIR estimator, finding the presence of a negative bias which is significant in almost all the scenarios simulated between uncoupled point processes (**Figure 3**). A similar bias, even though considerably smaller, was described in the work that first introduced the TER and MIR estimators (Mijatovic et al.,

2021a; Shorten et al., 2021). As opposed to previous applications in neuroscience, cardiovascular interactions feature conditions of strongly auto-correlated processes and of short data sequences, which can be responsible of the strong bias that very often leads to meaningless negative values of MIR, thus justifying the adoption of countermeasures to prevent such bias. In Shorten et al. (2021), this bias was associated with a violation of the assumption of local uniformity of the probability density within the range of the k nearest neighbors used for entropy estimation. While methods for reducing the bias of nearest neighbor information estimators which address specifically cases where local uniformity does not apply can be devised (Gao et al., 2015), in this work we resort to an empirical approach that reproduces the bias of MIR estimated over uncoupled surrogate time series, and then subtracts this bias from the MIR computed for the original series. This empirical approach has the advantage of generality, since surrogates mimic the data distribution and are in principle able to reproduce diverse sources of bias and to compensate them in the corrected measure (Papana et al., 2011). We find that different procedures for surrogate data generation have a different impact on the detection coupling and on the compensation of the bias, with differences being emphasized as the size of the analyzed event series decreases. A main advantage of the resulting cMIR measure is that it establishes the statistical significance of the information shared by the two processes, meaning that it does not indicate significant coupling when the coupling is indeed absent (Papana et al., 2011); this aspect has been verified in our simulations showing that cMIR tends to zero when the studied processes approach the uncoupled regime. A drawback of the proposed correction stands in the fact that it reproduces the bias for uncoupled processes, which can be different than that occurring in the case of coupling. As a consequence, since the MIR for truly coupled signals can be affected by a different bias than that observed for uncoupled signals, our approach does not provide a rigorous correction of the bias when the coupling is nonzero and some residual bias possibly remains also after the correction. Moreover, some applications of cMIR to networks with several connections to be estimated can become computationally unfeasible since the generation and information-theoretic analysis of surrogate point processes is a time-consuming procedure.

The proposed approach to estimate MIR in the presence of short and possibly noisy point process data is recommended for applications in the field of Network Physiology, where the estimation of organ system interactions is typically challenged by the inherently complex nature of human physiological signals (Lehnertz et al., 2020). In our work, where complex point process interactions between the heartbeat timings and the arrival times of the sphygmic wave on the body periphery have been analyzed, we detected significant coupling between the two processes in all subjects and experimental conditions. Moreover, the statistically significant variations of cMIR observed during physiological stress suggest that the index can reflect the neuroautonomic modulation of the heartbeat and vascular dynamics. This conclusion is supported by previous studies performed by using different approaches working in discrete time on interval time series, which suggests that the differences between heart rate and pulse rate variability are due not only to measurement noise, but also to physiological factors (Schäfer and Vagedes, 2013; Pernice et al., 2019). These factors are related to the physiological modulation of the two time intervals that compose the PAT, i.e. the PEP and the PTT. According to our present findings and previous research (Krohová et al., 2017; Czippelova et al., 2019; Pernice et al., 2019), the increased variability of PAT observed during postural stress arises from an increased variability of PEP related to sympathetic influence on cardiac contractility, an increased variability of systolic blood pressure leading to increased PTT variability, and an increased variability in the vascular tone related to sympathetic vasomotor control. These effects are manifested mainly in the LF band (0.04–0.15 Hz) of the spectrum, which is the frequency range where dominant oscillations of the blood pressure and the heart rate are observed during head-up tilt (Montano et al., 1994; Pernice et al., 2021). Accordingly, we ascribe the increase of cMIR observed during postural stress to the activation of the sympathetic nervous system and to the increased chronotropic baroreflex coupling occurring with tilt, which are likely responsible of the synchronous modulation of the LF variability of heart rate and PAT. Whilst we support this interpretation with our simulation, a recent study showed that heart rate and PAT variability are more correlated at the frequency of the Mayer waves (~ 0.1 Hz) (Peng et al., 2021). On the other hand, the smaller changes of cMIR observed during mental arithmetic suggest that mental stress evokes a different type of sympathetic activation, possibly more of central origin than related to common modulation of heart rate and vascular tone (Javorka et al., 2017). Future studies should address the separate role of PEP and PTT variability in the changes of the coupling between heartbeat and systolic time dynamics, and investigate the clinical value (e.g., in relation to the alterations of the arterial compliance observed with aging or hypertension) of the novel measures computed in this work.

In summary, the method for MIR computation presented in this work constitutes a viable approach to assess the rate of information exchanged dynamically between pairs of point processes from short realizations of event-based data. Our approach, which explicitly considers the point-process structure of human heartbeats, is alternative to existing model-free information measures developed in discrete time and working on amplitudes rather than on events (Porta and Faes, 2015), as well as to existing model-based parametric models developed in the point process framework (Barbieri et al., 2005; Valenza et al., 2018; Greco et al., 2019). As such, it holds the potential to disclose different physiological mechanisms than those investigated by traditional cardiovascular variability approaches.

DATA AVAILABILITY STATEMENT

The experimental data used for this article are available upon request to the corresponding Author. The software packages relevant to MIR and cMIR estimation, and to different algorithms for generation of surrogates of event sequences, are

available for free download from the GitHub repositories https://github.com/mijatovicg/TEMI, and https://github.com/LeonardoRicci/SpiSeMe.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Jessenius Faculty of Medicine, Comenius University, Martin, Slovakia. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LF contributed to conceptualization and supervision. GM, AP, YA, LR, and LF contributed to methodology. GM and YA contributed to software and validation. MJ and RP contributed to data curation. GM, RP, and LF contributed to writing—original draft preparation. LR, MJ, and AB contributed to writing—review and editing. GM, RP, and AP contributed to visualization. GM, MJ, and AB contributed to funding acquisition. All authors have read and agreed to the published version of the manuscript.

FUNDING

Authors acknowledge support from the Ministry of Education, Science and Technological Development of Serbia, project no. 451–03–68/2 020–14/200 156: "Innovative scientific and artistic research from the FTS (activity) domain", the European Union's Horizon 2020 research and innovation programme under Grant Agreement number 856 967, by the grants no. VEGA 1/0 283/21, VEGA 1/0 199/19, VEGA 1/0 200/19, and by the Italian MIUR, project PRIN 2017 (PRJ-0167) 2017WZFTZP "Stochastic forecasting in complex systems". RP is supported by the Italian MIUR PON R&I 2014–2020 AIM project no. AIM1851228-2.

REFERENCES

Amblard, P.-O., and Michel, O. J. (2013). The Relation between granger Causality and Directed Information Theory: A Review. *Entropy* 15, 113–143. doi:10.3390/e15010113

Barbieri, R., Matten, E. C., Alabi, A. A., and Brown, E. N. (2005). A point-process Model of Human Heartbeat Intervals: New Definitions of Heart Rate and Heart Rate Variability. *Am. J. Physiol.-Heart Circ. Physiol.* 288, H424–H435. doi:10.1152/ajpheart.00482.2003

Beda, A., Simpson, D. M., and Faes, L. (2017). Estimation of Confidence Limits for Descriptive Indexes Derived from Autoregressive Analysis of Time Series: Methods and Application to Heart Rate Variability. *PLoS One* 12, e0183230. doi:10.1371/journal.pone.0183230

Carnethon, M. R., Liao, D., Evans, G. W., Cascio, W. E., Chambless, L. E., and Heiss, G. (2002). Correlates of the Shift in Heart Rate Variability with an Active Postural Change in a Healthy Population Sample: The Atherosclerosis Risk in Communities Study. *Am. Heart J.* 143, 808–813. doi:10.1067/mhj.2002.121928

Chan, G. S. H., Middleton, P. M., Celler, B. G., Wang, L., and Lovell, N. H. (2007). Change in Pulse Transit Time and Pre-ejection Period during Head-Up Tilt-Induced Progressive central Hypovolaemia. *J. Clin. Monit. Comput.* 21, 283–293. doi:10.1007/s10877-007-9086-8

Cohen, M. A., and Taylor, J. A. (2002). Short-Term Cardiovascular Oscillations in Man: Measuring and Modelling the Physiologies. *J. Physiol.* 542, 669–683. doi:10.1113/jphysiol.2002.017483

Cover, T. M. (1999). *Elements of Information Theory*. Hoboken, New Jersey: John Wiley & Sons.

Czippelova, B., Turianikova, Z., Krohova, J., Wiszt, R., Lazarova, Z., Pozorciakova, K., et al. (2019). Arterial Stiffness and Endothelial Function in Young Obese Patients - Vascular Resistance Matters. *J. Atheroscler. Thromb.* 26, 1015–1025. doi:10.5551/jat.47530

Duncan, T. E. (1970). On the Calculation of Mutual Information. *SIAM J. Appl. Math.* 19, 215–220. doi:10.1137/0119020

Faes, L., Pinna, G. D., Porta, A., Maestri, R., and Nollo, G. D. (2004). Surrogate Data Analysis for Assessing the Significance of the Coherence Function. *IEEE Trans. Biomed. Eng.* 51, 1156–1166. doi:10.1109/tbme.2004.827271

Faes, L., Nollo, G., and Porta, A. (2012). Non-uniform Multivariate Embedding to Assess the Information Transfer in Cardiovascular and Cardiorespiratory Variability Series. *Comput. Biol. Med.* 42, 290–297. doi:10.1016/j.compbiomed.2011.02.007

Faes, L., Nollo, G., and Porta, A. (2013). Mechanisms of Causal Interaction between Short-Term Rr Interval and Systolic Arterial Pressure Oscillations during Orthostatic challenge. *J. Appl. Physiol.* 114, 1657–1667. doi:10.1152/japplphysiol.01172.2012

Faes, L., Marinazzo, D., Montalto, A., and Nollo, G. (2014). Lag-specific Transfer Entropy as a Tool to Assess Cardiovascular and Cardiorespiratory Information Transfer. *IEEE Trans. Biomed. Eng.* 61, 2556–2568. doi:10.1109/tbme.2014.2323131

Faes, L., Kugiumtzis, D., Nollo, G., Jurysta, F., and Marinazzo, D. (2015). Estimating the Decomposition of Predictive Information in Multivariate Systems. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 91, 032904. doi:10.1103/PhysRevE.91.032904

Fauvel, J. P., Cerutti, C., Quelin, P., Laville, M., Gustin, M. P., Paultre, C. Z., et al. (2000). Mental Stress-Induced Increase in Blood Pressure Is Not Related to Baroreflex Sensitivity in Middle-Aged Healthy Men. *Hypertension* 35, 887–891. doi:10.1161/01.hyp.35.4.887

Fiedor, P. (2014). Networks in Financial Markets Based on the Mutual Information Rate. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 89, 052801. doi:10.1103/PhysRevE.89.052801

Gao, S., Ver Steeg, G., and Galstyan, A. (2015). "Efficient Estimation of Mutual Information for Strongly Dependent Variables," in *Artificial Intelligence and Statistics* (PMLR), 277–286.

Garde, A., Laursen, B., Jørgensen, A., and Jensen, B. (2002). Effects of Mental and Physical Demands on Heart Rate Variability during Computer Work. *Eur. J. Appl. Physiol.* 87, 456–461. doi:10.1007/s00421-002-0656-7

Gray, R. M. (2011). *Entropy and Information Theory*. Hoboken, New Jersey: Springer Science & Business Media.

Greco, A., Faes, L., Catrambone, V., Barbieri, R., Scilingo, E. P., and Valenza, G. (2019). Lateralization of Directional Brain-Heart Information Transfer during Visual Emotional Elicitation. *Am. J. Physiol. -Regul., Integr. Comp. Physiol.* 317, R25–R38. doi:10.1152/ajpregu.00151.2018

Javorka, M., Krohova, J., Czippelova, B., Turianikova, Z., Lazarova, Z., Javorka, K., et al. (2017). Basic Cardiovascular Variability Signals: Mutual Directed Interactions Explored in the Information Domain. *Physiol. Meas.* 38, 877–894. doi:10.1088/1361-6579/aa5b77

Javorka, M., Krohova, J., Czippelova, B., Turianikova, Z., Lazarova, Z., Wiszt, R., et al. (2018). Towards Understanding the Complexity of Cardiovascular Oscillations: Insights from Information Theory. *Comput. Biol. Med.* 98, 48–57. doi:10.1016/j.compbiomed.2018.05.007

Kim, H.-G., Cheon, E.-J., Bai, D.-S., Lee, Y. H., and Koo, B.-H. (2018). Stress and Heart Rate Variability: a Meta-Analysis and Review of the Literature. *Psychiatry Investig.* 15, 235–245. doi:10.30773/pi.2017.08.17

Kozachenko, L., and Leonenko, N. N. (1987). Sample Estimate of the Entropy of a Random Vector. *Probl. Peredachi Inf.* 23, 9–16.

Krohová, J., Czippelová, B., Turianiková, Z., Lazarová, Z., Tonhajzerová, I., and Javorka, M. (2017). Preejection Period as a Sympathetic Activity index: a Role of Confounding Factors. *Physiol. Res.* 66, S265–S275. doi:10.33549/physiolres.933682

Krohova, J., Faes, L., Czippelova, B., Pernice, R., Turianikova, Z., Wiszt, R., et al. (2020). Vascular Resistance Arm of the Baroreflex: Methodology and Comparison with the Cardiac Chronotropic Arm. *J. Appl. Physiol.* 128, 1310–1320. doi:10.1152/japplphysiol.00512.2019

Lancaster, G., Iatsenko, D., Pidde, A., Ticcinelli, V., and Stefanovska, A. (2018). Surrogate Data for Hypothesis Testing of Physical Systems. *Phys. Rep.* 748, 1–60. doi:10.1016/j.physrep.2018.06.001

Lehnertz, K., Bröhl, T., and Rings, T. (2020). The Human Organism as an Integrated Interaction Network: Recent Conceptual and Methodological Challenges. *Front. Physiol.* 11, 1694. doi:10.3389/fphys.2020.598694

Magagnin, V., Bassani, T., Bari, V., Turiel, M., Maestri, R., Pinna, G. D., et al. (2011). Non-stationarities Significantly Distort Short-Term Spectral, Symbolic and Entropy Heart Rate Variability Indices. *Physiol. Meas.* 32, 1775–1786. doi:10.1088/0967-3334/32/11/s05

Marschinski, R., and Kantz, H. (2002). Analysing the Information Flow between Financial Time Series. *Eur. Phys. J. B* 30, 275–281. doi:10.1140/epjb/e2002-00379-2

Martin, S. L., Carek, A. M., Kim, C. S., Ashouri, H., Inan, O. T., Hahn, J. O., et al. (2016). Weighing Scale-Based Pulse Transit Time Is a superior Marker of Blood Pressure Than Conventional Pulse Arrival Time. *Sci. Rep.* 6, 39273–39278. doi:10.1038/srep39273

Martinelli, F. S., Chacon-Mikahil, M. P. T., Martins, L. E. B., Lima-Filho, E. C., Golfetti, R., Paschoal, M. A., et al. (2005). Heart Rate Variability in Athletes and Nonathletes at Rest and during Head-Up Tilt. *Braz. J. Med. Biol. Res.* 38, 639–647. doi:10.1590/s0100-879x2005000400019

Mijatovic, G., Antonacci, Y., Turukalo, T. L., Minati, L., and Faes, L. (2021a). An Information-Theoretic Framework to Measure the Dynamic Interaction between Neural Spike Trains. *IEEE Trans. Biomed. Eng.* 68, 3471–3481. doi:10.1109/tbme.2021.3073833

Mijatovic, G., Loncar-Turukalo, T., Bozanic, N., Milosavljevic, N., Storchi, R., and Faes, L. (2021b). A Measure of Concurrent Neural Firing Activity Based on Mutual Information. *Neuroinformatics* 19, 719–735. doi:10.1007/s12021-021-09515-w

Minati, L., Faes, L., Frasca, M., Oświęcimka, P., and Drożdż, S. (2018). Apparent Remote Synchronization of Amplitudes: A Demodulation and Interference Effect. *Chaos* 28, 063124. doi:10.1063/1.5026980

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

Mukkamala, R., Hahn, J.-O., Inan, O. T., Mestha, L. K., Kim, C.-S., Toreyin, H., et al. (2015). Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice. *IEEE Trans. Biomed. Eng.* 62, 1879–1901. doi:10.1109/tbme.2015.2441951

Okada, M., Matsuto, T., Satoh, S., Igarashi, S., Baba, M., Sugita, O., et al. (1996). Role of Pulse Wave Velocity for Assessing Autonomic Nervous System Activities in Reference to Heart Rate Variability. *Med. Inform.* 21, 81–90. doi:10.3109/14639239609009013

Papana, A., Kugiumtzis, D., and Larsson, P. G. (2011). Reducing the Bias of Causality Measures. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 83, 036207. doi:10.1103/PhysRevE.83.036207

Pasquale, V., Massobrio, P., Bologna, L. L., Chiappalone, M., and Martinoia, S. (2008). Self-organization and Neuronal Avalanches in Networks of Dissociated Cortical Neurons. *Neuroscience* 153, 1354–1369. doi:10.1016/j.neuroscience.2008.03.050

Peng, R. C., Li, Y., and Yan, W. R. (2021). A Correlation Study of Beat-To-Beat R-R Intervals and Pulse Arrival Time under Natural State and Cold Stimulation. *Sci. Rep.* 11, 1–10. doi:10.1038/s41598-021-90056-2

Perinelli, A., Castelluzzo, M., Minati, L., and Ricci, L. (2020). Spiseme: A Multi-Language Package For Spike Train Surrogate Generation. *Chaos* 30, 073120. doi:10.1063/5.0011328

Perinelli, A., Castelluzzo, M., Tabarelli, D., Mazza, V., and Ricci, L. (2021). Relationship between Mutual Information and Cross-Correlation Time Scale of Observability as Measures of Connectivity Strength. *Chaos* 31, 073106. doi:10.1063/5.0053857

Pernice, R., Javorka, M., Javorka, M., Krohova, J., Czippelova, B., Turianikova, Z., et al. (2019). Comparison of Short-Term Heart Rate Variability Indexes Evaluated through Electrocardiographic and Continuous Blood Pressure Monitoring. *Med. Biol. Eng. Comput.* 57, 1247–1263. doi:10.1007/s11517-019-01957-4

Pernice, R., Sparacino, L., Nollo, G., Stivala, S., Busacca, A., and Faes, L. (2021). Comparison of Frequency Domain Measures Based on Spectral Decomposition for Spontaneous Baroreflex Sensitivity Assessment after Acute Myocardial Infarction. *Biomed. Signal Process. Control.* 68, 102680. doi:10.1016/j.bspc.2021.102680

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., Catai, A. M., Takahashi, A. C. M., Magagnin, V., Bassani, T., Tobaldini, E., et al. (2011). Causal Relationships between Heart Period and Systolic Arterial Pressure during Graded Head-Up Tilt. *Am. J. Physiol.-Regul., Integr. Comp. Physiol.* 300, R378–R386. doi:10.1152/ajpregu.00553.2010

Rapalis, A., Janušauskas, A., Marozas, V., and Lukoševičius, A. (2017). Estimation of Blood Pressure Variability during Orthostatic Test Using Instantaneous Photoplethysmogram Frequency and Pulse Arrival Time. *Biomed. Signal Process. Control.* 32, 82–89. doi:10.1016/j.bspc.2016.10.014

Ricci, L., Castelluzzo, M., Minati, L., and Perinelli, A. (2019). Generation of Surrogate Event Sequences via Joint Distribution of Successive Inter-event Intervals. *Chaos* 29, 121102. doi:10.1063/1.5138250

Schäfer, A., and Vagedes, J. (2013). How Accurate Is Pulse Rate Variability as an Estimate of Heart Rate Variability? A Review on Studies Comparing Photoplethysmographic Technology with an Electrocardiogram. *Int. J. Cardiol.* 166, 15–29. doi:10.1016/j.ijcard.2012.03.119

Schreiber, T., and Schmitz, A. (1996). Improved Surrogate Data for Nonlinearity Tests. *Phys. Rev. Lett.* 77, 635–638. doi:10.1103/physrevlett.77.635

Schreiber, T. (2000). Measuring Information Transfer. *Phys. Rev. Lett.* 85, 461–464. doi:10.1103/physrevlett.85.461

Shorten, D. P., Spinney, R. E., and Lizier, J. T. (2021). Estimating Transfer Entropy in Continuous Time between Neural Spike Trains or Other Event-Based Data. *Plos Comput. Biol.* 17, e1008054. doi:10.1371/journal.pcbi.1008054

Spinney, R. E., and Lizier, J. T. (2018). Characterizing Information-Theoretic Storage and Transfer in Continuous Time Processes. *Phys. Rev. E* 98, 012314. doi:10.1103/PhysRevE.98.012314

Spinney, R. E., Prokopenko, M., and Lizier, J. T. (2017). Transfer Entropy in Continuous Time, with Applications to Jump and Neural Spiking Processes. *Phys. Rev. E* 95, 032319. doi:10.1103/PhysRevE.95.032319

Stein, P. K., Bosner, M. S., Kleiger, R. E., and Conger, B. M. (1994). Heart Rate Variability: a Measure of Cardiac Autonomic Tone. *Am. Heart J.* 127, 1376–1381. doi:10.1016/0002-8703(94)90059-0

Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P., and Brown, E. N. (2005). A point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural Ensemble, and Extrinsic Covariate Effects. *J. Neurophysiol.* 93, 1074–1089. doi:10.1152/jn.00697.2004

Valderas, M. T., Bolea, J., Laguna, P., Bailón, R., and Vallverdú, M. (2019). Mutual Information between Heart Rate Variability and Respiration for Emotion Characterization. *Physiol. Meas.* 40, 084001. doi:10.1088/1361-6579/ab310a

Valenza, G., Faes, L., Citi, L., Orini, M., and Barbieri, R. (2018). Instantaneous Transfer Entropy for the Study of Cardiovascular and Cardiorespiratory Nonstationary Dynamics. *IEEE Trans. Biomed. Eng.* 65, 1077–1085. doi:10.1109/TBME.2017.2740259

Vicente, R., Wibral, M., Lindner, M., and Pipa, G. (2011). Transfer Entropy-A Model-free Measure of Effective Connectivity for the Neurosciences. *J. Comput. Neurosci.* 30, 45–67. doi:10.1007/s10827-010-0262-3

Wood, R., Maraj, B., Lee, C. M., and Reyes, R. (2002). Short-term Heart Rate Variability during a Cognitive challenge in Young and Older Adults. *Age ageing* 31, 131–135. doi:10.1093/ageing/31.2.131

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.

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.

APPENDIX

This section reports the derivation of the decomposition of the MIR presented in **Eq. 6**. We start considering two discrete-time processes $X_\Delta = \{X_{t_n}\}$ and $Y_\Delta = \{Y_{t_n}\}$ defined at the discrete time instants $t_n = n\Delta t, n \in \mathbb{Z}$, where Δt is the time interval between samples expressed in unit of time; the derivation in continuous time, i.e., for $\Delta t \to 0$, will intuitively follow. In discrete time, the MIR is defined as

$$\dot{I}_{X_\Delta;Y_\Delta} = \lim_{n\to\infty} \frac{1}{n\Delta t} I\left(X_{t_1:\, t_n}; Y_{t_1:\, t_n}\right), \tag{17}$$

where $X_{t_1:\, t_n} = \{X_{\Delta t}, X_{2\Delta t}, \ldots, X_{n\Delta t}\}$ and $Y_{t_1:\, t_n} = \{Y_{\Delta t}, Y_{2\Delta t}, \ldots, Y_{n\Delta t}\}$ are n-dimensional vectors of consecutive random variables taken from the two processes. According to basic information-theoretic rules we have that $I\left(X_{t_1:\, t_n}; Y_{t_1:\, t_n}\right) = H\left(X_{t_1:\, t_n}\right) + H\left(Y_{t_1:\, t_n}\right) - H\left(X_{t_1:\, t_n}, Y_{t_1:\, t_n}\right)$, from which **Eq. 17** can be expressed in terms of entropy rates as

$$\dot{I}_{X_\Delta;Y_\Delta} = \dot{H}_{X_\Delta} + \dot{H}_{Y_\Delta} - \dot{H}_{X_\Delta,Y_\Delta}. \tag{18}$$

Then, recalling the equivalent definitions of entropy rate for sequences of identically distributed random variables (Cover, 1999), the entropy rate for the process X_Δ can be written as

$$\begin{aligned}\dot{H}_{X_\Delta} &= \frac{1}{\Delta t} \lim_{n\to\infty} \frac{H\left(X_{t_1:\, t_n}\right)}{n} \equiv \frac{1}{\Delta t} \lim_{n\to\infty} H\left(X_{t_i} | X_{t_{i-n}:\, t_{i-1}}\right) \\ &= \frac{1}{\Delta t} H\left(X_{t_i} | X^-_{t_i}\right),\end{aligned} \tag{19}$$

where the last term is independent on t_i due to stationarity and is written in compact form evidencing the past history of the process, $X^-_{t_i} = X_{t_{i-n}:\, t_{i-1}}$ with $n \to \infty$. By using **Eqs. 18, 19** can be formulated evidencing a difference of MI terms as

$$\begin{aligned}\dot{I}_{X_\Delta;Y_\Delta} &= \frac{1}{\Delta t}\left(H\left(X_{t_i} | X^-_{t_i}\right) + H\left(Y_{t_i} | Y^-_{t_i}\right)\right) - H\left(X_{t_i}, Y_{t_i} | X^-_{t_i}, Y^-_{t_i}\right) \\ &= \frac{1}{\Delta t}\left(H\left(X_{t_i}, X^-_{t_i}\right) - H\left(X^-_{t_i}\right) + H\left(Y_{t_i}, Y^-_{t_i}\right) - H\left(Y^-_{t_i}\right)\right. \\ &\quad \left. - H\left(X_{t_i}, Y_{t_i}, X^-_{t_i}, Y^-_{t_i}\right) + H\left(X^-_{t_i}, Y^-_{t_i}\right)\right) \\ &= \frac{1}{\Delta t}\left(I\left(X_{t_i}, X^-_{t_i}; Y_{t_i}, Y^-_{t_i}\right) - I\left(X_{t_i}; Y^-_{t_i}\right)\right).\end{aligned} \tag{20}$$

Moreover, application of the chain rule for mutual information allows to expand the first MI term in the last line of **Eq. 20** as

$$\begin{aligned}I\left(X_{t_i}, X^-_{t_i}; Y_{t_i}, Y^-_{t_i}\right) &= I\left(X^-_{t_i}; Y_{t_i}, Y^-_{t_i}\right) + I\left(X_{t_i}; Y_{t_i}, Y^-_{t_i} | X^-_{t_i}\right) \\ &= I\left(X^-_{t_i}; Y^-_{t_i}\right) + I\left(Y_{t_i}, X^-_{t_i} | Y^-_{t_i}\right) \\ &\quad + I\left(X_{t_i}; Y_{t_i}, Y^-_{t_i} | X^-_{t_i}\right) \\ &= I\left(X^-_{t_i}; Y^-_{t_i}\right) + I\left(Y_{t_i}, X^-_{t_i} | Y^-_{t_i}\right) + I\left(X_{t_i}; Y^-_{t_i} | X^-_{t_i}\right) \\ &\quad + I\left(X_{t_i}; Y_{t_i} | X^-_{t_i}, Y^-_{t_i}\right).\end{aligned} \tag{21}$$

Finally, substituting **Eqs. 20, 21** leads to express the MIR as

$$\begin{aligned}\dot{I}_{X_\Delta;Y_\Delta} &= \frac{1}{\Delta t}\left(I\left(Y_{t_i}, X^-_{t_i} | Y^-_{t_i}\right) + I\left(X_{t_i}; Y^-_{t_i} | X^-_{t_i}\right) + I\left(X_{t_i}; Y_{t_i} | X^-_{t_i}, Y^-_{t_i}\right)\right) \\ &= \frac{1}{\Delta t}\left(T_{X_\Delta \to Y_\Delta} + T_{Y_\Delta \to X_\Delta} + I^0_{X_\Delta;Y_\Delta}\right),\end{aligned} \tag{22}$$

where the transfer entropies $T_{X_\Delta \to Y_\Delta}$ and $T_{Y_\Delta \to X_\Delta}$ (Schreiber, 2000) and the instantaneous information exchanged between the two processes, $I^0_{X_\Delta;Y_\Delta}$ (Amblard and Michel, 2013), are put in evidence. Taking the limit $\Delta t \to 0$ in **Eq. 22** leads to **Eq. 6**, which is valid when the processes X and Y are continuous.

frontiers in Network Physiology

ORIGINAL RESEARCH
published: 16 February 2022
doi: 10.3389/fnetp.2022.837860

Dynamic Temporal Relationship Between Autonomic Function and Cerebrovascular Reactivity in Moderate/Severe Traumatic Brain Injury

Logan Froese[1], Alwyn Gomez[2,3], Amanjyot Singh Sainbhi[1], Carleen Batson[3], Kevin Stein[2], Arsalan Alizadeh[2] and Frederick A. Zeiler[1,2,3,4,5]*

[1]Biomedical Engineering, Faculty of Engineering, University of Manitoba, Winnipeg, MB, Canada, [2]Section of Neurosurgery, Department of Surgery, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada, [3]Department of Human Anatomy and Cell Science, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada, [4]Centre on Aging, University of Manitoba, Winnipeg, MB, Canada, [5]Division of Anaesthesia, Department of Medicine, Addenbrooke's Hospital, University of Cambridge, Cambridge, United Kingdom

OPEN ACCESS

Edited by:
Alberto Porta,
University of Milan, Italy

Reviewed by:
Beatrice Cairo,
University of Milan, Italy
Federico Aletti,
Universidade Federal de São Paulo, Brazil
Teodor Buchner,
Warsaw University of Technology, Poland

***Correspondence:**
Logan Froese
log.froese@gmail.com

Specialty section:
This article was submitted to Networks in the Cardiovascular System, a section of the journal Frontiers in Network Physiology

Received: 17 December 2021
Accepted: 28 January 2022
Published: 16 February 2022

Citation:
Froese L, Gomez A, Sainbhi AS, Batson C, Stein K, Alizadeh A and Zeiler FA (2022) Dynamic Temporal Relationship Between Autonomic Function and Cerebrovascular Reactivity in Moderate/Severe Traumatic Brain Injury. Front. Netw. Physiol. 2:837860. doi: 10.3389/fnetp.2022.837860

There has been little change in morbidity and mortality in traumatic brain injury (TBI) in the last 25 years. However, literature has emerged linking impaired cerebrovascular reactivity (a surrogate of cerebral autoregulation) with poor outcomes post-injury. Thus, cerebrovascular reactivity (derived through the pressure reactivity index; PRx) is emerging as an important continuous measure. Furthermore, recent literature indicates that autonomic dysfunction may drive impaired cerebrovascular reactivity in moderate/severe TBI. Thus, to improve our understanding of this association, we assessed the physiological relationship between PRx and the autonomic variables of heart rate variability (HRV), blood pressure variability (BPV), and baroreflex sensitivity (BRS) using time-series statistical methodologies. These methodologies include vector autoregressive integrative moving average (VARIMA) impulse response function analysis, Granger causality, and hierarchical clustering. Granger causality testing displayed inconclusive results, where PRx and the autonomic variables had varying bidirectional relationships. Evaluating the temporal profile of the impulse response function plots demonstrated that the autonomic variables of BRS, ratio of low/high frequency of HRV and very low frequency HRV all had a strong relation to PRx, indicating that the sympathetic autonomic response may be more closely linked to cerebrovascular reactivity, then other variables. Finally, BRS was consistently associated with PRx, possibly demonstrating a deeper relationship to PRx than other autonomic measures. Taken together, cerebrovascular reactivity and autonomic response are interlinked, with a bidirectional impact between cerebrovascular reactivity and circulatory autonomics. However, this work is exploratory and preliminary, with further study required to extract and confirm any underlying relationships.

Keywords: cerebrovascular reactivity (CVRx), time series analysis, traumatic brain injury, causality testing, autonomic nervous system responses

1 INTRODUCTION

There has been little change in morbidity and mortality in moderate and severe traumatic brain injury (TBI) over the past 25 years (Carney et al., 2017; Maas et al., 2017; Donnelly et al., 2019; Steyerberg et al., 2019). TBI remains one of the leading burdens to global health (Maas et al., 2015), thus there is a need to investigate new ways to improve TBI care. Secondary injury mechanisms dictate ongoing neural injury during the acute phase of TBI care and take various forms. Such secondary injury pathways are essential targets for therapeutic intervention in moderate/severe TBI care. However, before developing precision therapeutics aimed at specific secondary injury mechanisms, we require a more comprehensive understanding of the inter-relationships between different aspects of cerebral physiology post-injury.

Impaired cerebral autoregulation in moderate/severe TBI is a secondary injury mechanism that leads to ongoing neural insult (Zeiler et al., 2020a). Literature has emerged demonstrating impaired cerebral autoregulation, assessed through cerebrovascular reactivity indices (surrogate measures for cerebrovascular autoregulation) after TBI are independently associated with mortality and poor functional outcome at 6 and 12 months post-injury (Czosnyka et al., 1997; Sorrentino et al., 2012; Zeiler et al., 2018a; Donnelly et al., 2019; Zeiler et al., 2019; Bennis et al., 2020; Zeiler et al., 2020b; Åkerlund et al., 2020). The pressure reactivity index (PRx) has emerged as one of the most commonly utilized methods for assessing cerebrovascular reactivity in the TBI literature (Zeiler et al., 2017). Despite the growing body of literature supporting the association of PRx with outcome (Czosnyka et al., 1997; Sorrentino et al., 2012; Zeiler et al., 2018a; Donnelly et al., 2019; Zeiler et al., 2019; Bennis et al., 2020; Zeiler et al., 2020b; Åkerlund et al., 2020), emerging literature suggests current guideline-based therapeutic interventions in moderate/severe TBI have little-to-no impact on the degree of impaired cerebrovascular reactivity seen (Donnelly et al., 2019; Froese et al., 2020a; Froese et al., 2020b; Zeiler et al., 2020b; Froese et al., 2020c). As such, more work is required to uncover the driving factors of impaired cerebrovascular reactivity.

Autonomic dysfunction after moderate/severe TBI has been well documented and is associated with poor global outcome (Hasen et al., 2019; Tymko et al., 2019; Fedriga et al., 2021a; Fedriga et al., 2021b). Furthermore, it is clear that autonomics and cerebrovascular function intersect (Ogoh et al., 2005; Hasen et al., 2019; Tymko et al., 2019; Fedriga et al., 2021a; Fedriga et al., 2021b). Recent literature demonstrates that PRx has an association with heart rate variability (HRV), including low frequency HRV (HRV_LF) and high frequency HRV (HRV_HF) (Lavinio et al., 2009; Sykora et al., 2016), with PRx also being connected to the baroreflex sensitivity (BRS) (Sykora et al., 2016). However, these studies had only a limited correlation between the autonomic variables and PRx, and did not examine the temporal profiles of autonomic and PRx measures. As such, a knowledge gap regarding the temporal and causal relationship between autonomic function and cerebrovascular reactivity exists.

Understanding the relationship between cerebrovascular reactivity and autonomic function is an important step to improve TBI care. The directional relationship between PRx and autonomic response portends to future targeted therapeutic development that is aimed at the prevention and reduction of secondary injury insult burden. Past work has shown that autonomic response drives factors associated with cerebrovascular reactivity, however dysautonomia has shown to be exacerbated by increases in intracranial pressure (Baguley et al., 1999; Baguley et al., 2008; Hasen et al., 2019). Beyond this, a deeper understanding of this relationship may enable the development of a more complete and accurate prognostic model that accounts for both cerebral autoregulatory and autonomic dysfunction. Ultimately, a robust understanding of how these secondary factors interconnect will improve our ability to predict patients at risk for cerebral autoregulation failure and ANS dysfunction.

Thus, using the prospectively maintained high-resolution data set from the Winnipeg Acute TBI Laboratories, we aim to examine the temporal and causal relationship between PRx and autonomic functionality in more detail using advanced time-series methodologies. The goal of this project is to comprehensively evaluate the time-series statistical properties of cerebrovascular reactivity and autonomics, focusing on the impact they have on each other. This will leverage the fact that circulatory phenomena respond in a fashion which may be assessed using the approach of linear interdependent time-series. Thus, using time-series analysis allows us to comment on what aspects of autonomic function drives cerebrovascular reactivity and gives a more complete picture of physiological response.

2 METHODS AND MATERIALS

2.1 Patients

Data were accessed retrospectively from the maintained TBI database at the Winnipeg Acute TBI Laboratories, University of Manitoba. For this study, patient data were collected from June 2018 up to December 2020. All patients suffered from moderate to severe TBI (moderate = Glasgow Coma Score (GCS) 9—12, and severe = GCS of 8 or less). All patients in this cohort were admitted to the intensive care unit where they were sedated, intubated and were on volume-controlled mode of ventilation (with constant PEEP), during the course of cerebral physiologic data collection. All patients had both invasive intracranial pressure (ICP) and arterial blood pressure (ABP) monitoring conducted, per the Brain Trauma Foundation guidelines (Carney et al., 2017).

2.2 Ethics

Data were collected following a full approval by the University of Manitoba Health Research Ethics Board (H2017:181, H2017:188, and H2020:118).

2.3 Data Collection

For this study, admission demographic information was extracted following the existing prognostic models in TBI (Dijkland et al.,

2021). Such demographic data collected was: age, sex, admission pupillary response (bilaterally reactive, unilaterally reactive, bilaterally unreactive) and admission GCS (both total and motor).

All patients had high-frequency digital signals recorded throughout their ICU stay. ABP was obtained through radial or femoral arterial lines connected to pressure transducers (Baxter Healthcare Corp. CardioVascular Group, Irvine, CA, or similar devices). ICP was acquired *via* an intra-parenchymal strain gauge probe (Codman ICP MicroSensor; Codman and Shurtlef Inc., Raynham, MA). All signals were captured simultaneously, synchronized and digitized via an A/D converter (DT9804; Data Translation, Marlboro, MA), sampled at a frequency of 100 Hertz (Hz) or higher using the Intensive Care Monitoring (ICM+) software (Cambridge Enterprise Ltd., Cambridge, United Kingdom, http://icmplus.neurosurg.cam.ac.uk). Signal artifacts were removed using manual methods before further processing and analysis. This ensured that all analyzed blood pressure had the distinct full wave beat, with all other data removed.

2.4 Signal Processing

2.4.1 Cerebrovascular Reactivity

Post-acquisition processing of the above signals was conducted using ICM+, in keeping with our previously published methodology (Froese et al., 2020b; Froese et al., 2020c). First, 10-s moving averages (updated every 10 s to avoid data overlap) were calculated for all recorded signals: ICP and ABP (which produced MAP). PRx was derived *via* the moving correlation coefficient between 30 consecutive 10-s mean windows of the parent signals (ICP and MAP), updated every minute according to previously validated methods (Czosnyka et al., 1997; Sorrentino et al., 2012; Zeiler et al., 2018b; Donnelly et al., 2019; Depreitere et al., 2021).

2.4.2 Autonomic Response Variables

To determine autonomic functionality, we used three categories of autonomic response that can be derived from a continuous ABP waveform. The categories were heart rate variability (HRV), blood pressure variability (BPV) and the baroreflex sensitivity (BRS). Each of these autonomic response variables (ARVs) were determined for each minute, calculated over both a 5 and 15-min window. These time windows were chosen because 5 min is a common time window for HRV (30) and a 15-min window is the minimum for a short BPV (Mena et al., 2005; Höcht, 2013; Parati et al., 2013). For spectral BPV the Lomb-Scargle periodogram was used to calculate spectral power of the ABP waveform over the 5 and 15-min window to derive the subsequent minute-by-minute updated ARV values (note these result in the power of $mmHg^2$) (Sykora et al., 2016; Szabo et al., 2018). For HRV, the original ABP was processed through a peak detection algorithm based on Pan-Tomkins method (Pan and Tompkins, 1985; Sykora et al., 2016; Szabo et al., 2018). This results in an irregularly sampled peak-to-peak time series value over the 5 and 15-min window. From this the Lomb-Scargle periodogram was used over the window to calculate spectral power values for the resulting HRVs, updated every minute (note these result in the power of $milliseconds^2$) (Electrophysiology, 1996; Sykora et al., 2016; Szabo et al., 2018). Thus, each variable can be time series linked to the minute-by-minute update interval of PRx and allows the implementation of times series methodologies of these spectral variables.

Due to the similarity in the final results between the time windows, all further data demonstrated will be of the 15-min windows.

HRV was derived from ABP by finding the power in 3 bandwidth categories; very low frequency (HRV_VLF; frequency less than 0.04 Hz), low frequency (HRV_LF; frequency of 0.04–0.15 Hz) and high frequency (HRV_HF; frequency of 0.15–0.4 Hz) (Electrophysiology, 1996; Berntson et al., 1997; Shaffer and Ginsberg, 2017). The interpretation of these frequencies is still up for debate (Electrophysiology, 1996; Berntson et al., 1997; Hayano and Yuda, 2019), although common interpretations are; HRV_VLF reflects slow mechanisms of sympathetic activity (though this warrants further elucidation), HRV_LF is a marker of sympathetic modulation or parameter that includes both sympathetic and vagal influences, and HRV_HF reflects parasympathetic (vagal) activity (Electrophysiology, 1996; Shaffer and Ginsberg, 2017). Due to the nature of spectral analysis of ABP waveforms the individual variables can be influenced by physiological responses that are adjunct or entirely separate from autonomic response (Electrophysiology, 1996; Berntson et al., 1997; Hayano and Yuda, 2019). Thus, any correlations must be taken as interpretations more than direct responses.

We also calculated the ratio between low and high frequency, (HRV_LF_HF; HRV_LF divided by HRV_HF) which represents minor sympathetic vagal balance or sympathetic modulations (though further investigation is still required) (Electrophysiology, 1996; Hayano and Yuda, 2019). The root mean square differences between consecutive heart beat period a heartbeat waveform (HRV_RMS) was found, which estimates the vagally mediated changes in autonomics (Electrophysiology, 1996; Hasen et al., 2019; Hayano and Yuda, 2019). The total power (HRV_TOT; which is the sum of the three spectral bands power) is a non-specific variable that reflects the overall autonomic activity (Electrophysiology, 1996; Shaffer and Ginsberg, 2017).

There were two methods of BPV found; the standard deviation of BPV in the time domain and the spectral domain analysis of BPV. The standard deviation of BPV was found in three main groups; mean blood pressure (BPV_M), systolic blood pressure (BPV_S) and diastolic blood pressure (BPV_D) over the moving time window (Höcht, 2013).

Furthermore, we assessed the spectral domain of the systolic blood pressure variability in three domain frequency ranges: low frequency (SBPV _LF; frequency of 0.077–0.15 Hz), high frequency (SBPV _HF; frequency of 0.15–0.4 Hz) and total (SBPV _TOT; total power over the full frequency range) (Höcht, 2013). Though these variables have a limited understanding, current assessments show the following: SBPV_LF variability is modulated by the sympathetic/baroreflex of vascular/vasomotor tone, total peripheral resistance and the Mayer wave (Stauss, 2007; Aletti et al., 2009; Aletti et al., 2012; Aletti et al., 2013) and SBPV_HF variability is mainly influenced by changes in cardiac output,

parasympathetic and respiration action (Janssen et al., 1995; Aletti et al., 2009; Aletti et al., 2013).

Finally, the baroreflex sensitivity (BRS) was calculated using a modification of the sequential cross-correlation method. ABP was used to find systolic peaks and the heart beat period, then a linear regression between the 10-s series of heart beat period and the corresponding 10-s series of systolic blood pressure over the time window results in the BRS (Westerhof et al., 2004). BRS may provide a useful synthetic index of neural regulation at the sinus atrial node, though limitations still exist with this interpretation (Kardos et al., 2001; La Rovere et al., 2008; Pinna et al., 2015).

2.5 Time-Series Analyses

We implemented a wide variety of time-series tests and models to assess both functionality and associated causality between the individual ARVs and PRx. To help elucidate the uses of these methods as well as the limitations and pitfalls of such techniques, we will give an overview of the methods. However, a full conceptual understanding of these individual methodologies can be found from their respective literature (McQuitty, 1966; Granger, 1969; Lütkepohl, 2005; Rokach et al., 2005; Murtagh and Legendre, 2014; Kilian and Lütkepohl, 2017; Chatfield and Xing, 2019; Thelin et al., 2019).

The three major methods used were: vector autoregressive integrated moving average (VARIMA) impulse response function plots (IRF), Granger Causality Testing and Hierarchical Clustering. These methods were chosen for their exploratory nature and their use in previous analyses of temporal physiology within TBI (Zeiler et al., 2018c; Zeiler et al., 2018d; Thelin et al., 2019).

2.5.1 Vector Autoregressive Integrative Moving Average Impulse Response Functions Analysis

IRF are used to graphically demonstrate the causal effect of an impulse on a system. For our uses we created VARIMA models to represent the relationship between PRx and each ARV. Then we created a simulated impulse on the VARIMA model from each respective ARV on PRx and vice versa. In this way we graphically demonstrated the patient specific interaction between each ARV and PRx.

2.5.1.1 Autoregressive Integrative Moving Average Structure Analysis

In order to derive a VARIMA model in an effective manner (due to the heavy computational requirements of such a method) but also to evaluate the accuracy of a VARIMA model, we performed a Box-Jenkin's autoregressive integrative moving average (ARIMA) model for each patients' PRx and all ARVs (Lütkepohl, 2005; Zeiler et al., 2018c; Chatfield and Xing, 2019).

Initially, PRx and ARVs were evaluated for time stationarity using the: autocorrelation function (ACF) plots, partial autocorrelation function (PACF) plots and Kwiatkowski–Phillips–Schmidt–Shin (KPSS). The augmented Dickey-Fuller (ADF) was used to test for root trend. Note we assume that all variables have some aspect that is both interdependent and linear, due to the interconnection of circulatory/vascular function.

Next, the optimal ARIMA structure for PRx and ARVs were derived for each patient. Initially, the *auto.arima* (a pre-built R function) was used to determine the upper order limit for tested ARIMA models (auto.arima, 2021). Based on this, autoregressive order (p) was varied from 1 to 10, and the moving average order (q) was varied from 0 to 10, while the integrative order (d) was held at 0. The integrative order was held at 0 as the tests of ACF, PACF, KPSS, and ADF suggested that all the signals were stationary. This is in keeping with previous time-series work in TBI literature (Zeiler et al., 2018c; Zeiler et al., 2018d; Thelin et al., 2019). All the permutations of the ARIMA orders were assessed using the Akaike Information Criterion (AIC), and Log-Likelihood (LL) recorded for every model.

Using the AIC and LL, the optimal ARIMA structures for PRx and ARVs were compared in the datasheets, with the lowest AIC and highest LL values indicating superior models. A general Box-Jenkin's ARMA model for PRx can be found in **Supplementary Appendix SA1**.

A patient example of the serial ARIMA model testing with AIC and LL outputs can be found in **Supplementary Appendix SB** in Supplementary Material. Similarly, an example in **Figure 1** is given of the raw signal ACF and PACF plots, followed by the plots for the residuals of the optimal ARIMA model (found through LL and AIC), indicating that the autocorrelative structure has been adequately accounted for.

2.5.1.2 Vector Autoregressive Integrative Moving Average (VARIMA) Models

Next, we derived multi-variate VARIMA models to evaluate the impulse response of ARVs on PRx and vice versa. These models explore the behavior of two-time series variables, recorded simultaneously, and are derived through the extension of the standard Box-Jenkin's models into multi-variate systems (further descriptions can be found in the references) (Lütkepohl, 2005; Kilian and Lütkepohl, 2017; Chatfield and Xing, 2019). A formula representing the vector autoregressive moving average model (VARMA) of PRx and an ARV can be found in **Supplementary Appendix SA2**, which is a VARIMA with the integrative order held at 0.

Since the ACF and PACF did not indicate any cyclical trends in the variables, with ADF and KPSS indicating that all variables were stationary, we employed basic VARMA models with autoregressive order of four and moving average order of four. This was based on the findings from individual patient ARIMA models of ARVs for each patient and past work evaluating the ARIMA models of ICP and MAP (Thelin et al., 2019; Zeiler et al., 2020c; Zeiler et al., 2020d).

A VARIMA model autoregressive order of four was chosen given the optimal ARIMA models for many ARVs was less than two, as well as previous studies found that most patient's ICP and MAP had an autoregressive order of two (Thelin et al., 2019; Zeiler et al., 2020c; Zeiler et al., 2020d). Thus, as suggested by Helmut Lutkepohl, taking the product of the ARIMA autoregressive orders for VARIMA modelling is a method to ensure adequate model structure (Lütkepohl, 2005), and thus the order of four. For the moving average order for the VARIMA model, a value of 4 was chosen based on the previous study of ICP and MAP (52,59,60) and ARVs optimal order being below two in

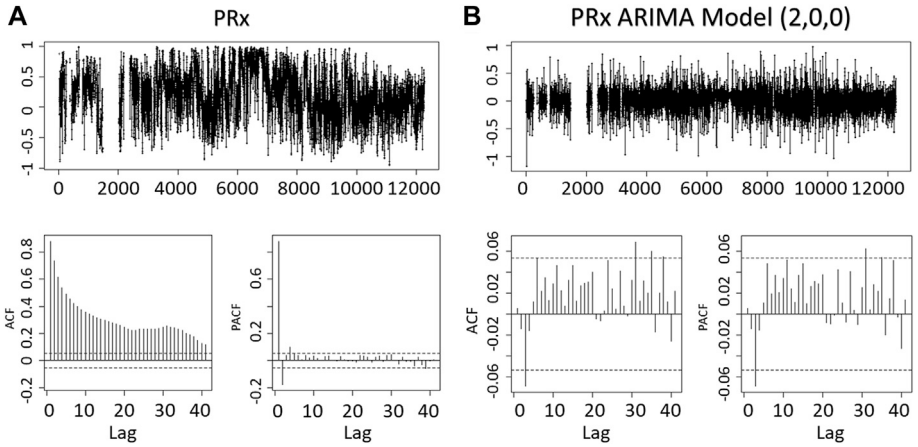


FIGURE 1 | The Residuals, ACF, PACF of the PRx and PRx ARIMA Model. The figures demonstrate the minute-by-minute PRx data and an optimal ARIMA model (autoregressive order of 2, integrative order of 0 and moving average order of 0) obtained from the Akaike Information Criteria and Log Likelihood. The reduced significant lags in ACF or PACF of the ARIMA model shows that the ARIMA model in part accounts for the residual relationship between the PRx data. ACF, autocorrelation function; PACF, partial autocorrelation function; PRx, pressure reactivity index; ARIMA, autoregressive integrated moving average.

many patients. Thus, the sum of the two ARIMA moving average orders was used to get a final VARIMA moving average order of 4.

2.5.1.3 Impulse Response Function Analysis

Next, the coefficients derived from these VARIMA models were employed to derive the IRF plots between PRx and an individual ARV. The IRF plots provide a descriptive graphical representation of the impact of PRx on an ARV, and an ARV on PRx, by using the previously generated VARIMA models and modelling a one unit orthogonal impulse of one variable on the other, and vice versa (Kilian and Lütkepohl, 2017). Depicted in the plot is how much one variable will fluctuate as a response to the impulse from the other variable. A bootstrapping method was used to derive the confidence interval of the population using sampled data. Bootstrapping involves using only part of the sample data for a run, then comparing all the runs, in this design a standard percentile bootstrap interval of 100 runs was used (Efron and Tibshirani, 1993).

Due to the high variability in the IRF plots and the difficulty that arises with simple graphical interpretations of large datasets, a simple method to help separate models into two categories ("more responsive" verses "less responsive") was used. To do this the impulse response was normalized with respect to its original variable data. Then identified if the response was greater than an absolute value of 0.001 threshold (chosen as at least a 0.1% change in the normalized response) after 10 min (one complete cycle of PRx after the initial impulse, i.e., 5 min post PRx calculation window). In this way, we could infer if the impulse created a stronger response within the subsequent variables and differentiate responses.

2.5.2 Granger Causality Testing

Granger causality testing is used to identify the assistance of one interdependent variable to predict another interdependent variable, beyond the degree to which the variable predicts itself (Barnett et al., 2009). In this case, the ability for PRx to predict an ARV (beyond the ability for the ARV to predict itself) and vice versa. Thus, with the minute-by-minute time series linked ARV data to PRx we could perform a Granger causality test between these interdependent variables.

For the Granger causality test, we recorded the response for every patient, both F-test statistic value and *p*-values for all ARVs vs. PRx (Granger, 1969). The Granger causality responses were assessed to identify the reciprocal influences between PRx and ARVs.

2.5.3 Co-Variance Cluster Analysis

Finally, to confirm our findings regarding the relationship between ARVs and PRx, a hierarchical clustering method was used on each patient to identify which ARVs and PRx was most closely associated. Using a divisive method, we separated the variables using the Euclidean distance of the normalized variables and the *hclust* (a prebuilt R function) (McQuitty, 1966; Rokach et al., 2005; Murtagh and Legendre, 2014).

2.5.4 Sub-Group Analysis

The entire database was subdivided based on some simple parameters and re-evaluated to see if the VARIMA IRF analysis or Granger causality test displayed any outlying groups. Parameters included were age (<60 vs. age ≥60; moderate vs old age), Glasgow outcome scale extended (GOSE) at 6 months (<2 vs. ≥ 2; dead vs alive), Marshall computer tomography (CT) score (<4 vs. ≥ 4; mass lesion vs diffuse injury), sex (male vs. female), the first 24 h only and the first 72 h only. Due to the similarity between the first 24 h only, 72 h only and the full time; all data presented will be of the full time.

2.6 Statistics

All statistical analyses were conducted using R (R Core Team (2016). R: A language and environment for statistical computing.

TABLE 1 | 47 Patient demographics.

Demographics	Mean (Interquartile range)
Age	38 (28.5–51)
Sex (% Male)	80.9%
Best admission GCS—total	6 (3.5–9)
Best admission GCS—motor	4 (1.5–5)
Number with hypoxia episode	20
Number with hypotension episode	5
Number with traumatic SAH	45
Number with epidural hematoma	5
Pupils	—
Bilateral unreactive	6
Unilateral unreactive	10
Bilateral reactive	31
Admission marshall CT	—
V	12
IV	8
III	14
II	3
1 month GOSE	6 (4.5–6)
30-Day mortality	27.7%
Average ICU stay (days)	6.34 (5.32–8.25)

CT, computerized tomography; GOSE, extended Glasgow outcome scale; GCS, Glasgow comma score; ICU, intensive care unit; SAH, subarachnoid hemorrhage.

R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/). The alpha was set at 0.05 for significance. No multiple error correction test was performed at this time as the analysis is in the preliminary exploratory phase. The patient population was summarized using simple descriptive statistics, including median/mean and standard deviation/IQR where applicable.

For the Granger causality test, we performed a Mann-Whitney U comparison test between the F-statistics of PRx on an ARV and an ARV on PRx over the entire population. As well we bar plotted the number of significant *p*-values for each variable. This gives the overarching relationship between the PRx and each ARV.

3 RESULTS

3.1 Patient Demographics

A total of 47 patients were included in this study. The mean age was 43.5 ± 23.5 years, with 37 (80.9%) being males. The median admission total GCS score was 6 (IQR: 3.5–9), and motor sub-scores were 4 (IQR: 1.5–5). Six patients (12.8%) presented with bilaterally unreactive pupils, and ten (21.3%) unilaterally unreactive pupils. The median duration of digital signal recording was 2.67 (IQR: 1.35–5.74) days. **Table 1** displays the patient admission demographics and injury information. **Supplementary Appendix SC** shows the median value for each measured variable over the full data set.

3.2 VARIMA and Impulse Response

To assess the relationship between PRx and ARVs, we employed VARIMA modelling with an IRF. VARIMA models of autoregressive order 4 and moving average order 4 were employed for each patient for the first 24 h only, 72 h only and the full data. IRF plots provide a descriptive visualization of the relationship between each ARV and PRx (examples seen in **Supplementary Appendix SD**). These IRF plots allowed us to visually determine the temporal relationship between PRx and ARVs, assessing the impact of one unit impulse on the respective variable.

Overall, there was high variability in absolute changes in the PRx and ARVs. However, ARVs demonstrated a higher magnitude in impulse response in PRx than the alternative. This was explored through the number of patients that exceeded an absolute value threshold of 0.001, with PRx responses being greater in amplitude for an ARV orthogonal impulse than the converse.

BRS demonstrated the most consistent number of patients in the "more responsive" cohort, with BRS impulse on PRx resulting in 9 "more responsive" patients vs PRx impulse on BRS resulting in 7 "more responsive" patients. Other ARVs on PRx that had at least 10% of the population (over five patients) in the "more responsive" category were; HRV_VLF, SBPV_HF, SBPV_TOT, BPV_M, and BPV_D. Likewise for PRx on APVs demonstrating over five patients in the "more responsive" category only had BRS and HRV_VLF.

3.3 Granger Causality

To assess the directional response between ARVs and PRx, we performed a Granger causality test comparing data sets for all patients. The Mann-Whitney U comparisons test between PRx and ARVs are in **Supplementary Appendix SE**, with **Figure 2** showing the number of patients with significant *p*-values (demonstrating a causal connection). In general, across the population, we found that the Granger causality test was inconclusive, with bidirectional causal features between PRx and ARVs seen across the cohort. In addition, some patients favored PRx impacting ARVs, though some showed the alternative causal relationship and not all directional relationships reached significance. **Supplementary Appendix SF** provides the Granger test results, including F-test and *p*-values, for every patient using the entire data set of full time. Of note, the causal direction of the relationship was not significantly changed when evaluating only the first 24 or 72 h of data.

The only variable that had a significant *p*-value was HRV_RMS in the first 24 and 72 h only. From **Figure 2** variables of HRV_VLF, HRV_LF_HF, BPV_S and BRS all had a moderate reduction in *p*-value significant patients from PRx on APV to APV on PRx. This may indicate a more impactful response from PRx in these relationships.

From the VARIMA IRF, PRx impulse most often caused a decrease in BRS. All other variables failed to have a consistent common PRx impulse response.

3.4 Hierarchical Cluster

Finally, the hierarchical clustering analysis helped confirm the connection that the variables had with one another. Though there was significant heterogeneity in the individual plots, PRx was most associated with BRS and HRV_LF_HF in a majority of patients. For an example plot, **Figure 3**. With the full data

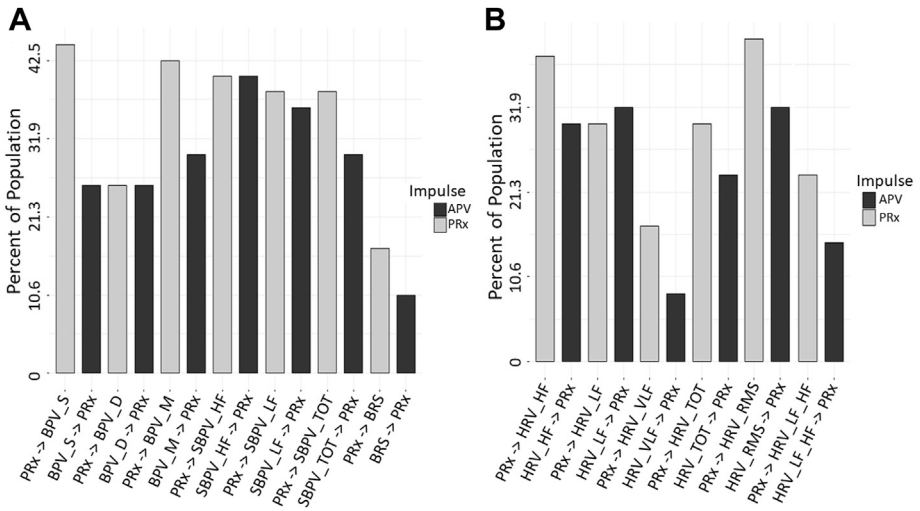


FIGURE 2 | Granger Causality *p*-value Comparisons ($n = 47$). The bar graphs show the number of significant Granger Causality *p*-values of PRx on an ARV and an ARV on PRx, in this way the impulse that has more significant responses may be considered to have a greater influence on causality. Image **(A)** shows all BPV and Baroreflex, **(B)** shows all HRV. ARV, autonomic response variable; BPV_D, standard deviation of diastolic blood pressure variability; BPV_M, standard deviation of mean blood pressure variability; BPV_S, standard deviation of systolic blood pressure variability; HRV, heart rate variability; HRF_HF, heart rate variability high frequency; HRV_LF, heart rate variability low frequency; HRV_LF_HF, heart rate variability ratio between low/high frequency; HRV_RMS, heart rate variability root mean square; HRV_TOT, heart rate variability total; HRV_VLF, heart rate variability very low frequency; PRx, pressure reactivity; SBPV_HF, spectral blood pressure variability high frequency; SBPV_LF, spectral blood pressure variability low frequency; SBPV_TOT, spectral blood pressure variability total.

cophenetic correlation and dendrogram displayed in **Supplementary Appendix SG**.

3.5 Age, GOSE and Significant Patients

Sub-group analysis failed to demonstrate any significant findings or finding that would be consider different from the full data. Dividing patients into sex, <60 vs. ≥60 age groups and six-month GOSE of <2 vs. ≥2 did not identify any significant outliers or an increased percentage of responsive patients. However, nearly all VARIMA IRF plots with a significant response (either PRx on an ARV or vice versa) had a Marshall CT score of 4 or higher.

4 DISCUSSION

We compared cerebrovascular reactivity as measured through PRx with various ARVs like HRV, BPV and BRS. We evaluated underlying behaviors by employing complex time-series analyses, like: ARIMA, VARIMA, IRF plots, Granger causality testing, and hierarchical clustering. Through this evaluation of PRx and ARVs, some exploratory insights into the association in physiological responses were illustrated. The preliminary evidence supports the idea that certain ARVs may have a clinically relevant association with impaired PRx. Further, results here help outline the framework for the role that systemic autonomic response has on cerebral autoregulation. However, the insights here must be tempered by the following limitations inherent with the use of any continuous ABP waveform method but in particular those of a spectral nature.

Foremost is the fact that significant changes to HRV, BPV or BRS may have occurred due to concomitant incidents outside of physiological autonomic response. In order to limit this, we followed the outlined method to achieve an optimal HRV, which are; signal capturing should allow signal reconstruction without amplitude and phase distortion, individual subjects should be recorded under fairly similar conditions and environments, and complete signals should be carefully edited using visual checks (Electrophysiology, 1996). Patients were all in the ICU with TBI, and were sedated, intubated and were on volume control mode of ventilation, with constant PEEP, during the course of cerebral physiologic data collection. The sympathetic nervous system seems to attenuate the CO_2-induced increases in CBF (Jordan et al., 2000), which can be mediated by ventilation in the ICU. Though it has been noted that powerful actions of mechanical ventilation induces periodical modifications of the intrathoracic pressure, modulating venous return which has shown to alter cardiovascular and cerebrovascular interactions (Innes et al., 1993; Elstad et al., 2011; Porta et al., 2021). With literature assessing the relationship between PEEP and cerebral reactivity in pigs, demonstrating that static PEEP improved the assessment of impaired/intact cerebrovascular reactivity (Brady et al., 2012; Fraser et al., 2013). Furthermore, the ABP changes slower then 30 s were linked to PRx, with improved consistency, when PEEP was constant (Fraser et al., 2013). Taken together, this highlights the need for more study between respiratory control and vascular influences. Finally, when the rate of change in blood pressure or cardiac output is rapidly altered, cerebral autoregulation has a reduced ability to regulate CBF (Levine et al., 1994; Zhang et al., 2002; Ogoh et al., 2005; Ogoh et al., 2007). Though it is impossible to limit all situations of extreme systemic circulatory response in the ICU, as part of critical TBI targets, ICP and cerebral prefusion

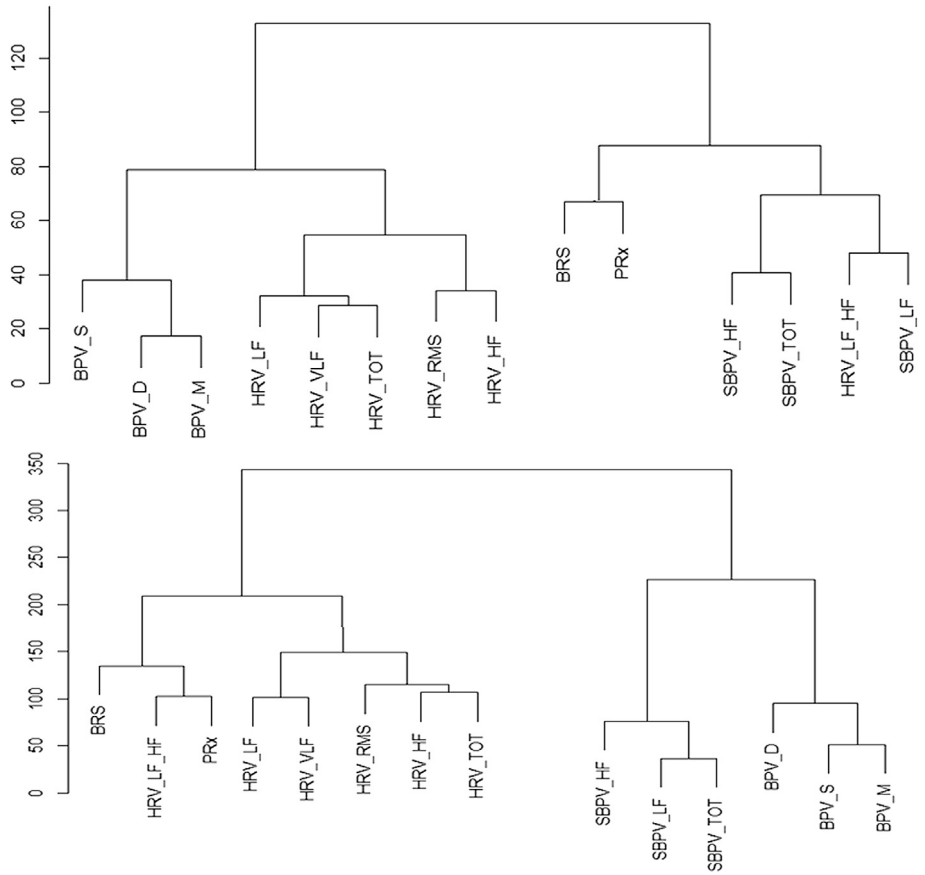


FIGURE 3 | Example of Hierarchical Clusters of Two Patients. Hierarchy cluster example from two patients, the data is normalized, and the distance between the variables is Euclidean. The distance between two variables shows the height of the dendrogram where the two branches merge into a single branch, thus two variables that diverge on the last branch may be more closely linked. BPV_D, standard deviation of diastolic blood pressure variability; BPV_M, standard deviation of mean blood pressure variability; BPV_S, standard deviation of systolic blood pressure variability; HRF_HF, heart rate variability high frequency; HRV_LF, heart rate variability low frequency; HRV_LF_HF, heart rate variability ratio between low/high frequency; HRV_RMS, heart rate variability root mean square; HRV_TOT, heart rate variability total; HRV_VLF, heart rate variability very low frequency; PRx, pressure reactivity; SBPV_HF, spectral blood pressure variability high frequency; SBPV_LF, spectral blood pressure variability low frequency; SBPV_TOT, spectral blood pressure variability total.

pressure are kept in moderate ranges (Carney et al., 2017). With this in mind the results can be summarized as the following.

First PRx, appears to causally impact ARVs as displayed through Granger causality, with the IRF plots of ARVs on PRx demonstrating a greater change then the inverse. This implies that derangements in cerebrovascular reactivity may, in turn, cause fluctuations in systemic circulatory autonomic responses. However, based on IRF results, any fluctuations in systemic circulatory autonomic responses result in larger derangements to cerebrovascular reactivity than the inverse. In a past study that assessed MAP and CBF, there is a slight bidirectional interaction between MAP and CBF but MAP mostly had a unidirectional impact on CBF (assessed over ~2 min windows) (Schiatti et al., 2014). Further in patients either in a normal or tilt table posture, there was association between MAP and CBF (Bari et al., 2017). Thus, underlying patient characteristics could influence the directionality of the relationship between cerebrovascular reactivity and autonomic function.

Second, certain ARVs like BRS, HRV_LF_HF and HRV_VLF demonstrated a directional impact on PRx as assessed through Granger causality and a strong impulse response in PRx, in over 10% of patients ($n > 5$). This implies that autonomics may, in fact, be responsible for causing derangements in cerebrovascular reactivity and may be important in achieving optimal CBF. Furthermore, from the associated nature of BRS, HRV_VLF and HRV_LF_HF, the sympathetic autonomic efferent response may be the primary autonomic factor associated with PRx response (Lindvall et al., 1978; Karemaker, 2017). This however is still up for debate as previously stated the direct interpretation of ARV response is limited, with the ARV being most commonly associated with sympathetic response (HRV_LF) remaining similarly connected to PRx as other ARVs. This may be due to the limited number of patients and requires further evaluation.

Despite this, previous literature has linked PRx to HRV_LF (Sykora et al., 2016), with other studies that demonstrated HRV_LF and PRx independently correlated with outcome

(Lavinio et al., 2009; Gao et al., 2016). Moreover, sympathetic nerves modulating resistance vessels tone has been demonstrated (Baumbach and Heistad, 1983). As well the sympathetic tone during spinal cord stimulation confirmed indirectly its role in mediating the CBF (Visocchi, 2006; Visocchi, 2008). Likewise CBF autoregulation is shifted towards higher blood pressure levels during sympathetic activation (Bill and Linder, 1976; Sadoshima et al., 1985). Furthermore, two independent studies suggested that stimulation of the sympathetic nerves can extend the limit of autoregulation (Fitch et al., 1975; MacKenzie et al., 1979). Thus, derangements of sympathetic autonomics may, in turn, interfere with cerebral vascular reactivity in certain patients.

However, previous work has demonstrated that HRV_HF can predict impaired cerebrovascular reactivity (PRx > 0.2) (Lavinio et al., 2009). Likewise, a study performed by Fedriga et al. indicated that CBF is maintained by the baroreflex and the parasympathetic autonomic response (as assessed by the association of BRS and HRV_HF with the upper limits of ICP values) during plateau waves of ICP (Fedriga et al., 2021a). In conjunction with this, HRV_HF and BRS in the past have shown a connection with ICP and cerebral perfusion pressure (Goldstein et al., 1998; Ogoh et al., 2007; Kox et al., 2012; Hasen et al., 2019; Fedriga et al., 2021a). However, the only situations where HRV_HF (representing the parasympathetic autonomic response) and PRx are connected is when cerebrovascular reactivity is already impaired, for example, during extremes of ICP elevation. In such states where ICP is at extreme levels or cerebrovascular reactivity is heavily impaired, the natural homeostasis of cerebral autoregulation is already heavily deranged, and thus any subsequent variation in systemic blood pressure (especially those of a higher frequency nature) would be reciprocated in the ICP response. This may account for why previous studies have linked the parasympathetic response to cerebrovascular reactivity and encourages the idea that the derangement of PRx is linked primarily to the sympathetic response of the autonomic system. However, we must acknowledge that the results found in this manuscript are preliminary and require much further validation.

Finally, nearly all VARIMA IRF plots with a significant response (either PRx on an ARV or vice versa) had a Marshall CT score of 4 or higher. This may indicate intracranial injury burden as a driver of the autonomic/cerebrovascular reactivity relationships, which is in keeping with recent literature supporting the strong association between diffuse acceleration-deceleration injury patterns and the development of cerebral autoregulation impairment/failure (Hiler et al., 2006; Zeiler et al., 2018e; Zeiler et al., 2020e).

In summation, there appears to be a directional impact of PRx on ARVs as assessed through Granger causality and IRF plots in some patients. Despite corroboration through various statistical approaches, the outcome of this study should be interpreted as only a preliminary exploration into the interconnected nature between ARVs and cerebrovascular reactivity. The responses themselves are significantly heterogeneous from patient to patient, with the IRF showing both positive and negative responses in PRx values. Thus, future work continuously analyzing PRx and ARVs would benefit from large cohorts separated into key demographical groups, with the use of clustering methodology to isolate homogeneous physiological factors. Key among these groups would be patients with impaired vs intact cerebrovascular reactivity. The heterogeneity in patient response, coupled with the small cohort size, leaves these statistical models quite limited in their overall assessment.

5 LIMITATIONS

As this was an exploratory analysis of the ARVs and PRx, many overarching limitations could be assessed. First, this is a retrospective analysis of a relatively small prospectively collected dataset. As such, our findings should only be considered exploratory and preliminary. Further, the results here may not be generalizable to other TBI populations and requires validation in larger multi-center high-frequency physiologic datasets. Second, patient injury severity and treatment heterogeneity could have influenced the physiologic signal response and is something that will require more tailored and refined datasets, with the use of clustering methodology to isolate homogeneous physiological factors. Third, the nature of these ARVs and their connection with cerebral autoregulation is severely limited, with the individual variables themselves still up to interpretation as to which aspect of the autonomic nervous system they truly represent (Hayano and Yuda, 2019). Thus, avenues that focus on the more extreme cases of autonomic and PRx change may provide more useful insights as to the effect of such variable responses, as demonstrated in the limited previous literature on plateau waves in moderate/severe TBI cohorts (Hasen et al., 2019; Tymko et al., 2019; Fedriga et al., 2021a; Fedriga et al., 2021b). Likewise there is a known influence of mechanical ventilation on ARV, and thus the evaluation of patient populations outside of intensive care may allow for more conclusive results.

Finally, the statistical methodology employed was computationally tasking and, as such, implementing such methodologies on larger cohorts would benefit from more robust central computing services.

6 CONCLUSION

Using statistical methods like ARIMA, VARIMA IRF, Granger causality and hierarchical clustering, we evaluated the temporal relationship between ARVs and cerebrovascular reactivity (as measured through PRx) in moderate/severe TBI patients. Granger causality testing demonstrated inconclusive results, with bidirectional relationships between PRx and ARVs in most of the cohort studied. However, the ARVs of BRS, HRV_LF_HF and HRV_VLF all demonstrated a stronger connection to PRx than other ARVs, indicating that the sympathetic autonomic response may be connected to cerebrovascular reactivity derangements. Finally, BRS was consistently one of the most responsive ARVs to PRx, possibly demonstrating a unique connection. However, this work is exploratory and preliminary, with further examination required to extract any underlying relationships.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved. The data was collected following a full approval by the University of Manitoba Health Research Ethics Board (H2017: 181, H2017:188, and H2020:118). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

LF and FZ contributed to conception and design of the study. LF, AG, AS, and FZ organized the database. LF performed the statistical analysis. LF wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This study was supported by the Manitoba Public Insurance (MPI) Neuroscience/TBI Research Endowment, the Health Sciences Centre Foundation Winnipeg, and the University of Manitoba Department of Surgery GFT Research Grant program. FZ receives research support from the Manitoba Public Insurance (MPI) Neuroscience/TBI Research Endowment, the Health Sciences Centre Foundation Winnipeg, the United States National Institutes of Health (NIH) through the National Institute of Neurological Disorders and Stroke (NINDS) (Grant #: R03NS114335-01), the Canadian Institutes of Health Research (CIHR) (Grant #: 432061), the Canada Foundation for Innovation (CFI) (Project #: 38583), Research Manitoba (Grant #: 3906), the University of Manitoba VPRI Research Investment Fund (RIF), the University of Manitoba Centre on Aging, and the University of Manitoba Rudy Falk Clinician-Scientist Professorship. AG is supported through the University of Manitoba Clinician Investigator Program. AS is supported through the University of Manitoba, Department of Surgery GFT Research Grant. LF is supported through the University of Manitoba - Department of Surgery GFT Research Grant and the University of Manitoba Office of Research Services (ORS)—University Research Grant Program (URGP). CB is supported through the Centre on Aging at the University of Manitoba.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnetp.2022.837860/full#supplementary-material

REFERENCES

Åkerlund, C. A., Donnelly, J., Zeiler, F. A., Helbok, R., Holst, A., Cabeleira, M., et al. (2020). Impact of Duration and Magnitude of Raised Intracranial Pressure on Outcome after Severe Traumatic Brain Injury: A CENTER-TBI High-Resolution Group Study. *PLoS One* 15 (12), e0243427. doi:10.1371/journal.pone.0243427

Aletti, F., Bassani, T., Lucini, D., Pagani, M., and Baselli, G. (2009). Multivariate Decomposition of Arterial Blood Pressure Variability for the Assessment of Arterial Control of Circulation. *IEEE Trans. Biomed. Eng.* 56 (7), 1781–1790. doi:10.1109/tbme.2009.2016845

Aletti, F., Ferrario, M., Xu, D., Greaves, D. K., Shoemaker, J. K., Arbeille, P., et al. (2012). Short-term Variability of Blood Pressure: Effects of Lower-Body Negative Pressure and Long-Duration Bed Rest. *Am. J. Physiology-Regulatory, Integr. Comp. Physiol.* 303 (1), R77–R85. doi:10.1152/ajpregu.00050.2012

Aletti, F., Hammond, R. L., Sala-Mercado, J. A., Chen, X., O'Leary, D. S., Baselli, G., et al. (2013). Cardiac Output Is Not a Significant Source of Low Frequency Mean Arterial Pressure Variability. *Physiol. Meas.* 34 (9), 1207–1216. doi:10.1088/0967-3334/34/9/1207

auto.arima (2021). auto.arima Function - RDocumentation [Internet].Available from: https://www.rdocumentation.org/packages/forecast/versions/8.15/topics/auto.arima (Accessed December 2, 2021).

Baguley, I. J., Heriseanu, R. E., Cameron, I. D., Nott, M. T., and Slewa-Younan, S. (2008). A Critical Review of the Pathophysiology of Dysautonomia Following Traumatic Brain Injury. *Neurocrit. Care* 8 (2), 293–300. doi:10.1007/s12028-007-9021-3

Baguley, I. J., Nicholls, J. L., Felmingham, K. L., Crooks, J., Gurka, J. A., and Wade, L. D. (1999). Dysautonomia after Traumatic Brain Injury: a Forgotten Syndrome. *J. Neurol. Neurosurg. Psychiatry* 67 (1), 39–43. doi:10.1136/jnnp.67.1.39

Bari, V., De Maria, B., Mazzucco, C. E., Rossato, G., Tonon, D., Nollo, G., et al. (2017). Cerebrovascular and Cardiovascular Variability Interactions Investigated through Conditional Joint Transfer Entropy in Subjects Prone to Postural Syncope. *Physiol. Meas.* 38 (5), 976–991. doi:10.1088/1361-6579/aa638c

Barnett, L., Barrett, A. B., and Seth, A. K. (2009). Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables. *Phys. Rev. Lett.* 103 (23), 238701. doi:10.1103/physrevlett.103.238701

Baumbach, G. L., and Heistad, D. D. (1983). Effects of Sympathetic Stimulation and Changes in Arterial Pressure on Segmental Resistance of Cerebral Vessels in Rabbits and Cats. *Circ. Res.* 52 (5), 527–533. doi:10.1161/01.res.52.5.527

Bennis, F. C., Teeuwen, B., Zeiler, F. A., Elting, J. W., van der Naalt, J., Bonizzi, P., et al. (2020). Improving Prediction of Favourable Outcome after 6 Months in Patients with Severe Traumatic Brain Injury Using Physiological Cerebral Parameters in a Multivariable Logistic Regression Model. *Neurocrit. Care* 33, 542–551. doi:10.1007/s12028-020-00930-6

Berntson, G. G., Thomas Bigger, J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., et al. (1997). Heart Rate Variability: Origins, Methods, and Interpretive Caveats. *Psychophysiology* 34 (6), 623–648. doi:10.1111/j.1469-8986.1997.tb02140.x

Bill, A., and Linder, J. (1976). Sympathetic Control of Cerebral Blood Flow in Acute Arterial Hypertension. *Acta Physiol. Scand.* 96 (1), 114–121. doi:10.1111/j.1748-1716.1976.tb10176.x

Brady, K. M., Easley, R. B., Kibler, K., Kaczka, D. W., Andropoulos, D., Charles, D., et al. (2012). Positive End-Expiratory Pressure Oscillation Facilitates Brain Vascular Reactivity Monitoring. *J. Appl. Physiol.* 113 (9), 1362–1368. doi:10.1152/japplphysiol.00853.2012

Carney, N., Totten, A. M., O'Reilly, C., Ullman, J. S., Hawryluk, G. W. J., Bell, M. J., et al. (2017). Guidelines for the Management of Severe Traumatic Brain Injury, Fourth Edition. *Neurosurgery* 80 (1), 6–15. doi:10.1227/neu.0000000000001432

Chatfield, C., and Xing, H. (2019). *The Analysis of Time Series: An Introduction with R*. New York: CRC Press, 415.

Czosnyka, M., Smielewski, P., Kirkpatrick, P., Laing, R. J., Menon, D., and Pickard, J. D. (1997). Continuous Assessment of the Cerebral Vasomotor Reactivity in Head Injury. *Neurosurgery* 41 (1), 11–19. doi:10.1097/00006123-199707000-00005

Depreitere, B., Citerio, G., Smith, M., Adelson, P. D., Aries, M. J., Bleck, T. P., et al. (2021). Cerebrovascular Autoregulation Monitoring in the Management of Adult Severe Traumatic Brain Injury: A Delphi Consensus of Clinicians. *Neurocrit. Care* 34 (3), 731–738. doi:10.1007/s12028-020-01185-x

Dijkland, S. A., Helmrich, I. R. A. R., Nieboer, D., van der Jagt, M., Dippel, D. W. J., Menon, D. K., et al. (2021). Outcome Prediction after Moderate and Severe Traumatic Brain Injury: External Validation of Two Established Prognostic Models in 1742 European Patients. *J. Neurotrauma* 38 (10), 1377–1388. doi:10.1089/neu.2020.7300

Donnelly, J., Czosnyka, M., Adams, H., Cardim, D., Kolias, A. G., Zeiler, F. A., et al. (2019). Twenty-Five Years of Intracranial Pressure Monitoring after Severe Traumatic Brain Injury: A Retrospective, Single-Center Analysis. *Neurosurgery* 85 (1), E75–E82. doi:10.1093/neuros/nyy468

Efron, B., and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hall, 436. (Monographs on statistics and applied probability).

Electrophysiology (1996). Heart Rate Variability. *Circulation* 93 (5), 1043–1065. doi:10.1161/01.cir.93.5.1043

Elstad, M., Walløe, L., Chon, K. H., and Toska, K. (2011). Low-frequency Fluctuations in Heart Rate, Cardiac Output and Mean Arterial Pressure in Humans: what Are the Physiological Relationships. *J. Hypertens.* 29 (7), 1327–1336. doi:10.1097/hjh.0b013e328347a17a

Fedriga, M., Czigler, A., Nasr, N., Zeiler, F. A., Park, S., Donnelly, J., et al. (2021a). Autonomic Nervous System Activity during Refractory Rise in Intracranial Pressure. *J. Neurotrauma* 38 (12), 1662–1669. doi:10.1089/neu.2020.7091

Fedriga, M., Czigler, A., Nasr, N., Zeiler, F. A., Beqiri, E., Wolf, S., et al. (2021b). Methodological Consideration on Monitoring Refractory Intracranial Hypertension and Autonomic Nervous System Activity. *Acta Neurochir Suppl.* 131, 211–215. doi:10.1007/978-3-030-59436-7_41

Fitch, W., MacKenzie, E. T., and Harper, A. M. (1975). Effects of Decreasing Arterial Blood Pressure on Cerebral Blood Flow in the Baboon. Influence of the Sympathetic Nervous System. *Circ. Res.* 37 (5), 550–557. doi:10.1161/01.res.37.5.550

Fraser, C. D., Brady, K. M., Rhee, C. J., Easley, R. B., Kibler, K., Smielewski, P., et al. (2013). The Frequency Response of Cerebral Autoregulation. *J. Appl. Physiol.* 115 (1), 52–56. doi:10.1152/japplphysiol.00068.2013

Froese, L., Dian, J., Batson, C., Gomez, A., Unger, B., and Zeiler, F. A. (2020c). The Impact of Hypertonic saline on Cerebrovascular Reactivity and Compensatory reserve in Traumatic Brain Injury: an Exploratory Analysis. *Acta Neurochir (Wien)* 162, 2683–2693. doi:10.1007/s00701-020-04579-0

Froese, L., Batson, C., Gomez, A., Dian, J., and Zeiler, F. A. (2020a). The Limited Impact of Current Therapeutic Interventions on Cerebrovascular Reactivity in Traumatic Brain Injury: A Narrative Overview. *Neurocrit. Care* 34, 325–335. doi:10.1007/s12028-020-01003-4

Froese, L., Dian, J., Batson, C., Gomez, A., Alarifi, N., Unger, B., et al. (2020b). The Impact of Vasopressor and Sedative Agents on Cerebrovascular Reactivity and Compensatory Reserve in Traumatic Brain Injury: An Exploratory Analysis. *Neurotrauma Rep.* 1 (1), 157–168. doi:10.1089/neur.2020.0028

Gao, L., Smielewski, P., Czosnyka, M., and Ercole, A. (2016). Cerebrovascular Signal Complexity Six Hours after Intensive Care Unit Admission Correlates with Outcome after Severe Traumatic Brain Injury. *J. Neurotrauma* 33 (22), 2011–2018. doi:10.1089/neu.2015.4228

Goldstein, B., Toweill, D., Lai, S., Sonnenthal, K., and Kimberly, B. (1998). Uncoupling of the Autonomic and Cardiovascular Systems in Acute Brain Injury. *Am. J. Physiology-Regulatory, Integr. Comp. Physiol.* 275 (4), R1287–R1292. doi:10.1152/ajpregu.1998.275.4.r1287

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* 37 (3), 424–438. doi:10.2307/1912791

Hasen, M., Almojuela, A., and Zeiler, F. A. (2019). Autonomic Dysfunction and Associations with Functional and Neurophysiological Outcome in Moderate/Severe Traumatic Brain Injury: A Scoping Review. *J. Neurotrauma* 36 (10), 1491–1504. doi:10.1089/neu.2018.6073

Hayano, J., and Yuda, E. (2019). Pitfalls of Assessment of Autonomic Function by Heart Rate Variability. *J. Physiol. Anthropol.* 38 (1), 3. doi:10.1186/s40101-019-0193-2

Hiler, M., Czosnyka, M., Hutchinson, P., Balestreri, M., Smielewski, P., Matta, B., et al. (2006). Predictive Value of Initial Computerized Tomography Scan, Intracranial Pressure, and State of Autoregulation in Patients with Traumatic Brain Injury. *J. Neurosurg.* 104 (5), 731–737. doi:10.3171/jns.2006.104.5.731

Höcht, C. (2013). Blood Pressure Variability: Prognostic Value and Therapeutic Implications. *ISRN Hypertens.* 2013, e398485. doi:10.5402/2013/398485

Innes, J. A., De Cort, S. C., Kox, W., and Guz, A. (1993). Within-breath Modulation of Left Ventricular Function during normal Breathing and Positive-Pressure Ventilation in Man. *J. Physiol.* 460, 487–502. doi:10.1113/jphysiol.1993.sp019483

Janssen, B. J., Oosting, J., Slaaf, D. W., Persson, P. B., and Struijker-Boudier, H. A. (1995). Hemodynamic Basis of Oscillations in Systemic Arterial Pressure in Conscious Rats. *Am. J. Physiol.* 269 (1 Pt 2), H62–H71. doi:10.1152/ajpheart.1995.269.1.H62

Jordan, J., Shannon, J. R., Diedrich, A., Black, B., Costa, F., Robertson, D., et al. (2000). Interaction of Carbon Dioxide and Sympathetic Nervous System Activity in the Regulation of Cerebral Perfusion in Humans. *Hypertension* 36 (3), 383–388. doi:10.1161/01.hyp.36.3.383

Kardos, A., Watterich, G., de Menezes, R., Csanády, M., Casadei, B., and Rudas, L. (2001). Determinants of Spontaneous Baroreflex Sensitivity in a Healthy Working Population. *Hypertension* 37 (3), 911–916. doi:10.1161/01.hyp.37.3.911

Karemaker, J. M. (2017). An Introduction into Autonomic Nervous Function. *Physiol. Meas.* 38 (5), R89–R118. doi:10.1088/1361-6579/aa6782

Kilian, L., and Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis [Internet]*. Cambridge: Cambridge University Press. (Themes in Modern Econometrics). Available from: https://www.cambridge.org/core/books/structural-vector-autoregressive-analysis/DAF4217439EA585D10902D58A8849E06.

Kox, M., Vrouwenvelder, M. Q., Pompe, J. C., van der Hoeven, J. G., Pickkers, P., and Hoedemaekers, C. W. (2012). The Effects of Brain Injury on Heart Rate Variability and the Innate Immune Response in Critically Ill Patients. *J. Neurotrauma* 29 (5), 747–755. doi:10.1089/neu.2011.2035

La Rovere, M. T., Pinna, G. D., and Raczak, G. (2008). Baroreflex Sensitivity: Measurement and Clinical Implications. *Ann. Noninv Electrocard* 13 (2), 191–207. doi:10.1111/j.1542-474x.2008.00219.x

Lavinio, A., Ene-Iordache, B., Nodari, I., Girardini, A., Cagnazzi, E., Rasulo, F., et al. (2009). "Cerebrovascular Reactivity and Autonomic Drive Following Traumatic Brain Injury," in *Acta Neurochirurgica Supplements*. Editor H-J. Steiger (Vienna: Springer), 3–7. (Acta Neurochirurgica Supplementum).

Levine, B. D., Giller, C. A., Lane, L. D., Buckey, J. C., and Blomqvist, C. G. (1994). Cerebral versus Systemic Hemodynamics during Graded Orthostatic Stress in Humans. *Circulation* 90 (1), 298–306. doi:10.1161/01.cir.90.1.298

Lindvall, M., Edvinsson, L., and Owman, C. (1978). Sympathetic Nervous Control of Cerebrospinal Fluid Production from the Choroid Plexus. *Science* 201 (4351), 176–178. doi:10.1126/science.663649

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin: Springer Science & Business Media, 765.

Maas, A. I. R., Menon, D. K., Adelson, P. D., Andelic, N., Bell, M. J., Belli, A., et al. (2017). Traumatic Brain Injury: Integrated Approaches to Improve Prevention, Clinical Care, and Research. *Lancet Neurol.* 16 (12), 987–1048. doi:10.1016/S1474-4422(17)30371-X

Maas, A. I. R., Menon, D. K., Steyerberg, E. W., Citerio, G., Lecky, F., Manley, G. T., et al. (2015). Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury (CENTER-TBI). *Neurosurgery* 76 (1), 67–80. doi:10.1227/neu.0000000000000575

MacKenzie, E. T., McGeorge, A. P., Graham, D. I., Fitch, W., Edvinsson, L., and Harper, A. M. (1979). Effects of Increasing Arterial Pressure on Cerebral Blood Flow in the Baboon: Influence of the Sympathetic Nervous System. *Pflugers Arch.* 378 (3), 189–195. doi:10.1007/bf00592735

McQuitty, L. L. (1966). Similarity Analysis by Reciprocal Pairs for Discrete and Continuous Data. *Educ. Psychol. Meas.* 26 (4), 825–831. doi:10.1177/001316446602600402

Mena, L., Pintos, S., Queipo, N. V., Aizpúrua, J. A., Maestre, G., and Sulbarán, T. (2005). A Reliable index for the Prognostic Significance of Blood Pressure Variability. *J. Hypertens.* 23 (3), 505–511. doi:10.1097/01.hjh.0000160205.81652.5a

Murtagh, F., and Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion. *J. Classif* 31 (3), 274–295. doi:10.1007/s00357-014-9161-z

Ogoh, S., Brothers, R. M., Barnes, Q., Eubank, W. L., Hawkins, M. N., Purkayastha, S., et al. (2005). The Effect of Changes in Cardiac Output on Middle Cerebral Artery Mean Blood Velocity at Rest and during Exercise. *J. Physiol.* 569 (Pt 2), 697–704. doi:10.1113/jphysiol.2005.095836

Ogoh, S., Dalsgaard, M. K., Secher, N. H., and Raven, P. B. (2007). Dynamic Blood Pressure Control and Middle Cerebral Artery Mean Blood Velocity Variability at Rest and during Exercise in Humans. *Acta Physiol.* 191 (1), 3–14. doi:10.1111/j.1748-1716.2007.01708.x

Pan, J., and Tompkins, W. J. (1985). A Real-Time QRS Detection Algorithm. *IEEE Trans. Biomed. Eng.* 32 (3), 230–236. doi:10.1109/tbme.1985.325532

Parati, G., Ochoa, J. E., Lombardi, C., and Bilo, G. (2013). Assessment and Management of Blood-Pressure Variability. *Nat. Rev. Cardiol.* 10 (3), 143–155. doi:10.1038/nrcardio.2013.1

Pinna, G. D., Maestri, R., and La Rovere, M. T. (2015). Assessment of Baroreflex Sensitivity from Spontaneous Oscillations of Blood Pressure and Heart Rate: Proven Clinical Value. *Physiol. Meas.* 36 (4), 741–753. doi:10.1088/0967-3334/36/4/741

Porta, A., Gelpi, F., Bari, V., Cairo, B., De Maria, B., Tonon, D., et al. (2021). Categorizing the Role of Respiration in Cardiovascular and Cerebrovascular Variability Interactions. *IEEE Trans. Biomed. Eng.* Preprint. doi:10.1109/tbme.2021.3135313

Rokach, L., and Maimon, O. (2005). "Clustering Methods," in *Data Mining and Knowledge Discovery Handbook [Internet]*. Editors O Maimon, and L Rokach (Boston, MA: Springer US), 321–352. doi:10.1007/0-387-25465-X_15

Sadoshima, S., Fujishima, M., Yoshida, F., Ibayashi, S., Shiokawa, O., and Omae, T. (1985). Cerebral Autoregulation in Young Spontaneously Hypertensive Rats. Effect of Sympathetic Denervation. *Hypertension* 7 (3 Pt 1), 392–397. doi:10.1161/01.hyp.7.3.392

Schiatti, L., Nollo, G., Rossato, G., and Faes, L. (2014). "Investigating Cardiovascular and Cerebrovascular Variability in Postural Syncope by Means of Extended Granger Causality," in 2014 8th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO) (Italy, Trento: IEEE), 43–44. doi:10.1109/esgco.2014.6847510

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

Sorrentino, E., Diedler, J., Kasprowicz, M., Budohoski, K. P., Haubrich, C., Smielewski, P., et al. (2012). Critical Thresholds for Cerebrovascular Reactivity after Traumatic Brain Injury. *Neurocrit. Care* 16 (2), 258–266. doi:10.1007/s12028-011-9630-8

Stauss, H. M. (2007). Identification of Blood Pressure Control Mechanisms by Power Spectral Analysis. *Clin. Exp. Pharmacol. Physiol.* 34 (4), 362–368. doi:10.1111/j.1440-1681.2007.04588.x

Steyerberg, E. W., Wiegers, E., Sewalt, C., Buki, A., Citerio, G., De Keyser, V., et al. (2019). Case-mix, Care Pathways, and Outcomes in Patients with Traumatic Brain Injury in CENTER-TBI: a European Prospective, Multicentre, Longitudinal, Cohort Study. *Lancet Neurol.* 18 (10), 923–934. doi:10.1016/S1474-4422(19)30232-7

Sykora, M., Czosnyka, M., Liu, X., Donnelly, J., Nasr, N., Diedler, J., et al. (2016). Autonomic Impairment in Severe Traumatic Brain Injury. *Crit. Care Med.* 44 (6), 1173–1181. doi:10.1097/ccm.0000000000001624

Szabo, J., Smielewski, P., Czosnyka, M., Jakubicek, S., Krebs, S., Siarnik, P., et al. (2018). Heart Rate Variability Is Associated with Outcome in Spontaneous Intracerebral Hemorrhage. *J. Crit. Care* 48, 85–89. doi:10.1016/j.jcrc.2018.08.033

Thelin, E. P., Raj, R., Bellander, B.-M., Nelson, D., Piippo-Karjalainen, A., Siironen, J., et al. (2019). Comparison of High versus Low Frequency Cerebral Physiology for Cerebrovascular Reactivity Assessment in Traumatic Brain Injury: a Multicenter Pilot Study. *J. Clin. Monit. Comput.* 34, 971–994. doi:10.1007/s10877-019-00392-y

Tymko, M. M., Donnelly, J., Smielewski, P., Zeiler, F. A., Sykora, M., Haubrich, C., et al. (2019). Changes in Cardiac Autonomic Activity during Intracranial Pressure Plateau Waves in Patients with Traumatic Brain Injury. *Clin. Auton. Res.* 29 (1), 123–126. doi:10.1007/s10286-018-0579-y

Visocchi, M. (2008). Neuromodulation of Cerebral Blood Flow by Spinal Cord Electrical Stimulation: the Role of the Italian School and State of Art. *J. Neurosurg. Sci.* 52 (2), 41–47.

Visocchi, M. (2006). Spinal Cord Stimulation and Cerebral Haemodynamics. *Acta Neurochir Suppl.* 99, 111–116. doi:10.1007/978-3-211-35205-2_21

Westerhof, B. E., Gisolf, J., Stok, W. J., Wesseling, K. H., and Karemaker, J. M. (2004). Time-domain Cross-Correlation Baroreflex Sensitivity. *J. Hypertens.* 22 (7), 1371–1380. doi:10.1097/01.hjh.0000125439.28861.ed

Zeiler, F. A., Ercole, A., Beqiri, E., Cabeleira, M., Thelin, E. P., Stocchetti, N., et al. (2020b). Association between Cerebrovascular Reactivity Monitoring and Mortality Is Preserved when Adjusting for Baseline Admission Characteristics in Adult Traumatic Brain Injury: A CENTER-TBI Study. *J. Neurotrauma* 37 (10), 1233–1241. doi:10.1089/neu.2019.6808

Zeiler, F. A., Ercole, A., Czosnyka, M., Smielewski, P., Hawryluk, G., Hutchinson, P. J. A., et al. (2020a). Continuous Cerebrovascular Reactivity Monitoring in Moderate/severe Traumatic Brain Injury: a Narrative Review of Advances in Neurocritical Care. *Br. J. Anaesth.* 124 (4), 440–453. doi:10.1016/j.bja.2019.11.031

Zeiler, F. A., Smielewski, P., Stevens, A., Czosnyka, M., Menon, D. K., and Ercole, A. (2018c). Non-Invasive Pressure Reactivity Index Using Doppler Systolic Flow Parameters: A Pilot Analysis. *J. Neurotrauma* 36 (5), 713–720. doi:10.1089/neu.2018.5987

Zeiler, F. A., Aries, M., Cabeleira, M., van Essen, T. A., Stocchetti, N., Menon, D. K., et al. (2020c). Statistical Cerebrovascular Reactivity Signal Properties after Secondary Decompressive Craniectomy in Traumatic Brain Injury: A CENTER-TBI Pilot Analysis. *J. Neurotrauma* 37 (11), 1306–1314. doi:10.1089/neu.2019.6726

Zeiler, F. A., Cabeleira, M., Hutchinson, P. J., Stocchetti, N., Czosnyka, M., Smielewski, P., et al. (2020d). Evaluation of the Relationship between Slow-Waves of Intracranial Pressure, Mean Arterial Pressure and Brain Tissue Oxygen in TBI: a CENTER-TBI Exploratory Analysis. *J. Clin. Monit. Comput.* 35, 711–722. doi:10.1007/s10877-020-00527-6

Zeiler, F. A., Donnelly, J., Calviello, L., Smielewski, P., Menon, D. K., and Czosnyka, M. (2017). Pressure Autoregulation Measurement Techniques in Adult Traumatic Brain Injury, Part II: A Scoping Review of Continuous Methods. *J. Neurotrauma* 34 (23), 3224–3237. doi:10.1089/neu.2017.5086

Zeiler, F. A., Donnelly, J., Nourallah, B., Thelin, E. P., Calviello, L., Smielewski, P., et al. (2018e). Intracranial and Extracranial Injury Burden as Drivers of Impaired Cerebrovascular Reactivity in Traumatic Brain Injury. *J. Neurotrauma* 35 (14), 1569–1577. doi:10.1089/neu.2017.5595

Zeiler, F. A., Donnelly, J., Smielewski, P., Menon, D. K., Hutchinson, P. J., and Czosnyka, M. (2018a). Critical Thresholds of Intracranial Pressure-Derived Continuous Cerebrovascular Reactivity Indices for Outcome Prediction in Noncraniectomized Patients with Traumatic Brain Injury. *J. Neurotrauma* 35 (10), 1107–1115. doi:10.1089/neu.2017.5472

Zeiler, F. A., Ercole, A., Cabeleira, M., Carbonara, M., Stocchetti, N., Menon, D. K., et al. (2019). Comparison of Performance of Different Optimal Cerebral Perfusion Pressure Parameters for Outcome Prediction in Adult Traumatic Brain Injury: A Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury (CENTER-TBI) Study. *J. Neurotrauma* 36 (10), 1505–1517. doi:10.1089/neu.2018.6182

Zeiler, F. A., Lee, J. K., Smielewski, P., Czosnyka, M., and Brady, K. (2018b). Validation of Intracranial Pressure-Derived Cerebrovascular Reactivity Indices against the Lower Limit of Autoregulation, Part II: Experimental Model of Arterial Hypotension. *J. Neurotrauma* 35 (23), 2812–2819. doi:10.1089/neu.2017.5604

Zeiler, F. A., Mathieu, F., Monteiro, M., Glocker, B., Ercole, A., Beqiri, E., et al. (2020e). Diffuse Intracranial Injury Patterns Are Associated with Impaired Cerebrovascular Reactivity in Adult Traumatic Brain Injury: A CENTER-TBI Validation Study. *J. Neurotrauma* 37 (14), 1597–1608. doi:10.1089/neu.2019.6959

Zeiler, F. A., Smielewski, P., Donnelly, J., Czosnyka, M., Menon, D. K., and Ercole, A. (2018d). Estimating Pressure Reactivity Using Noninvasive Doppler-Based Systolic Flow Index. *J. Neurotrauma* 35 (14), 1559–1568. doi:10.1089/neu.2017.5596

Zhang, R., Zuckerman, J. H., Iwasaki, K., Wilson, T. E., Crandall, C. G., and Levine, B. D. (2002). Autonomic Neural Control of Dynamic Cerebral Autoregulation in Humans. *Circulation* 106 (14), 1814–1820. doi:10.1161/01.cir.0000031798.07790.fe

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.

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.

frontiers in Network Physiology

ORIGINAL RESEARCH
published: 15 March 2022
doi: 10.3389/fnetp.2022.834056

Effects of Supplemental Oxygen on Cardiovascular and Respiratory Interactions by Extended Partial Directed Coherence in Idiopathic Pulmonary Fibrosis

Laura M. Santiago-Fuentes[1], Sonia Charleston-Villalobos[1], Ramón González-Camarena[2], Andreas Voss[3], Mayra E. Mejía-Avila[4], Ivette Buendía-Roldan[4], Sina Reulecke[1] and Tomás Aljama-Corrales[1]*

[1]Electrical Engineering Department, Universidad Autónoma Metropolitana, Mexico City, Mexico, [2]Health Science Department, Universidad Autónoma Metropolitana, Mexico City, Mexico, [3]Institute of Biomedical Engineering and Informatics, University of Technology Ilmenau, Ilmenau, Germany, [4]National Institute of Respiratory Diseases, Mexico City, Mexico

OPEN ACCESS

Edited by:
Sebastiano Stramaglia,
University of Bari Aldo Moro, Italy

Reviewed by:
Davide Nuzzi,
University of Bari Aldo Moro, Italy
Luca Faes,
University of Palermo, Italy
Ronny P. Bartsch,
Bar-Ilan University, Israel

***Correspondence:**
Sonia Charleston-Villalobos
schv@xanum.uam.mx

Specialty section:
This article was submitted to Information Theory, a section of the journal Frontiers in Network Physiology

Received: 15 December 2021
Accepted: 07 February 2022
Published: 15 March 2022

Citation:
Santiago-Fuentes LM, Charleston-Villalobos S, González-Camarena R, Voss A, Mejía-Avila ME, Buendía-Roldan I, Reulecke S and Aljama-Corrales T (2022) Effects of Supplemental Oxygen on Cardiovascular and Respiratory Interactions by Extended Partial Directed Coherence in Idiopathic Pulmonary Fibrosis. Front. Netw. Physiol. 2:834056. doi: 10.3389/fnetp.2022.834056

Idiopathic pulmonary fibrosis (IPF) is a chronic and restrictive disease characterized by fibrosis and inflammatory changes in lung tissue producing a reduction in diffusion capacity and leading to exertional chronic arterial hypoxemia and dyspnea. Furthermore, clinically, supplemental oxygen ($SupplO_2$) has been prescribed to IPF patients to improve symptoms. However, the evidence about the benefits or disadvantages of oxygen supplementation is not conclusive. In addition, the impact of $SupplO_2$ on the autonomic nervous system (ANS) regulation in respiratory diseases needs to be evaluated. In this study the interactions between cardiovascular and respiratory systems in IPF patients, during ambient air (AA) and $SupplO_2$ breathing, are compared to those from a matched healthy group. Interactions were estimated by time series of successive beat-to-beat intervals (BBI), respiratory amplitude (RESP) at BBI onset, arterial systolic (SYS) and diastolic (DIA) blood pressures. The paper explores the Granger causality (GC) between systems in the frequency domain by the extended partial directed coherence (ePDC), considering instantaneous effects. Also, traditional linear and nonlinear markers as power in low (LF) and high frequency (HF) bands, symbolic dynamic indices as well as arterial baroreflex, were calculated. The results showed that for IPF during AA phase: 1) mean BBI and power of BBI-HF band, as well as mean respiratory frequency were significantly lower ($p < 0.05$) and higher ($p < 0.001$), respectively, indicating a strong sympathetic influence, and 2) the RESP → SYS interaction was characterized by Mayer waves and diminished RESP → BBI, i.e., decreased respiratory sinus arrhythmia. In contrast, during short-term $SupplO_2$ phase: 1) oxygen might produce a negative influence on the systolic blood pressure variability, 2) the arterial baroreflex reduced significantly ($p < 0.01$) and 3) reduction of RSA reflected by RESP → BBI with simultaneous increase of Traube-Hering waves in RESP → SYS ($p < 0.001$), reflected increased sympathetic modulation to the vessels. The results gathered in this study may be helpful in the management of the administration of $SupplO_2$.

Keywords: idiopathic pulmonary fibrosis, supplemental oxygen, systems interactions, time-frequency cardiorespiratory interactions, extended partial directed coherence

1 INTRODUCTION

Idiopathic pulmonary fibrosis (IPF) is a chronic, restrictive, and progressive disease characterized by fibrosis and inflammatory changes in lung tissue. IPF also affects lung vasculature and prevents an adequate gas exchange, thus leading to exertional chronic arterial hypoxemia and dyspnea. The disease has a bad prognosis since after diagnosis the 50% of the patients die within 3–5 years (King et al., 2019). IPF etiology, diagnosis, treatment, and influences on quality of life, among other, have been investigated, but the focus has always been lungs performance. There is evidence that IPF patients also manifest comorbidities such as cardiovascular diseases, lung cancer, or pulmonary hypertension (Caminati et al., 2019). Furthermore, arterial hypertension has been related to dysfunctional autonomic cardiovascular control (Mancia and Grassi, 2014) and considerable impact on IPF disease progression and patient mortality (Buendía-Roldán et al., 2017).

Recently, IPF has been hypothesized as a systemic disease, but its influence on the autonomic nervous system (ANS) regulation has not been assessed as in other pulmonary disorders. For example, chronic obstructive pulmonary disease (COPD) negatively affects the cardiovascular system and the ANS regulation; the autonomic dysfunction is an important factor in the underlying pathophysiological mechanisms of the disease. The former conclusion has been mainly stated based on the analysis of heart rate variability (Van Gestel and Steier, 2010; Mohammed et al., 2015). However, for IPF disease studies about ANS regulation are scarce. Furthermore, supplemental oxygen ($SupplO_2$) has been prescribed to IPF patients to improve clinical symptoms, but its impact on ANS regulation of cardiovascular and respiratory systems has not been evaluated in these patients.

The effects of oxygen supplementation have been reviewed on healthy subjects and patients with cardiovascular diseases, but on respiratory patients just for COPD. Until now the evidence about the benefits or drawbacks of oxygen supplementation is not conclusive. On the one hand, some authors point out that this type of clinical intervention does not help to increase oxygen delivery, i.e., the rate at which oxygen is transported from lung to microcirculation depending on cardiac output and arterial oxygen contents (Smit et al., 2018a). Furthermore, it is plausible that in critically ill patients, it could be associated with increased hospital mortality (You et al., 2018). Conversely, oxygen therapy in COPD patients is clinically well-accepted and it is recommended to enhance exercise capacity (Stoller et al., 2010). However, some trials have not found enough evidence to support that long-term oxygen therapy improves COPD patients' mortality rate (Khor et al., 2019). It is important to point out that few studies have addressed the efficiency of oxygen administration in COPD and Interstitial Lung Disease (ILD) patients (Khor et al., 2019) and that studies about the ANS regulation in COPD patients were based on classic heart rate variability (HRV) parameters (Mohammed et al., 2017). Regarding IPF disorder, the authors of the present contribution compared the hemodynamic response to $SupplO_2$ between ill and healthy subjects, showing potential detrimental effects of $SupplO_2$ on IPF hemodynamics, particularly on total peripheral resistance (TPR) and cardiac output (Santiago-Fuentes et al., 2021).

Different methods have been employed to assess ANS regulation by the analysis of heart rate variability. In the case of HRV, it reflects the variations in the beat-to-beat interval and the corresponding variability time series is built up by diverse oscillatory modes. To extract the information from the time series, linear and nonlinear schemes have been proposed. For instance, the short-term HRV spectral density representation has been broadly used to analyze the sympathetic and parasympathetic modulation by the power in the low frequency (LF) and high frequency (HF) bands, respectively. Another way to assess ANS regulation of cardiovascular system is the nonlinear approach that allows to incorporate the analysis of complexity of underlying physiological mechanisms. Nonlinear indices by symbolic dynamics (SD) and detrended fluctuation analysis (DFA) have enabled the exploration of the cardiovascular system adaptability, among other, for example in elderly population (Beckers et al., 2006; Voss et al., 2015). Furthermore, the analysis of physiological variability time series has advanced from the univariate to bivariate and, to multivariate type to discover the complex interactions between human body subsystems. Nowadays, an open research area is to assess the cause-effect relationship to elucidate the complex picture of the autonomic control of the cardiovascular system and the complex interplay between cardiovascular and the respiratory systems.

Recently, the use of multivariate autoregressive models and the assessment of the directional interactions among a set of physiological variables (i.e., the so-called Granger causality) has been proposed for the analysis of ANS regulation under pathological and non-pathological conditions (Faes and Nollo, 2010; Faes et al., 2010; Charleston-Villalobos et al., 2019). Particularly, the extended partial directed coherence (ePDC) has gained interest as a tool to estimate the causality in the frequency domain in presence of instantaneous interactions (Faes and Nollo, 2010), i.e., as the effect from respiration to systolic blood pressure and beat to beat interval as well as the systolic blood pressure to beat to beat interval. Consequently, this study aimed to analysis cardiovascular and respiratory times series of variability by linear and nonlinear indices as well as Granger causalities (GC), via a multivariate autoregressive model including instantaneous effects, in IPF patients in comparison with healthy subjects under the effect of short-term $SupplO_2$.

2 MATERIALS AND METHODS

2.1 Subjects, Acquisition Protocol, and Preprocessing

This study includes 19 (8 women and 11 men) healthy subjects (CON) and 20 IPF patients (9 women and 11 men) with 67.79 ± 5.00 and 65.8 ± 6.48 years old, respectively. All subjects were medically evaluated at the National Institute of Respiratory Diseases in Mexico City after they accepted the invitation to participate and signed an informed consent according to the Declaration of Helsinki. **Table 1** depicts different parameters related to clinical measures and respiratory functional tests of both groups. Signals acquisition was performed via a Biopac MP150 system during morning hours including ECG, continuous

TABLE 1 | Clinical and functional measures.

Measure	Control (19 W:8/M:11)	IPF (20 W:9/M:11)
Hematocrit (%)	47.4 ± 4.9	51.6 ± 2.4
Hemoglobin (g/dL)	15.3 ± 1.4	15.0 ± 1.4
Respiratory rate (bpm)	17 ± 4	25 ± 7*
Fibrotic HRCT scan	—	1.99 ± 0.53
FEV_1 (%, predicted)	98.63 ± 12.87	78.15 ± 30.27*
FVC(%, predicted)	94.11 ± 12.72	72.40 ± 26.51*
FEV_1/FVC	77.69 ± 6.37	88.08 ± 9.54*
D_{LCO} (%, predicted)	114.26 ± 20.85	67.20 ± 21.24*
PaO_2(mmHg)	65.9[a]	61.92 ± 8.65
$PaCO_2$(mmHg)	32.7[a]	34.76 ± 5.31

Values expressed as mean plus/minus standard deviation. FEV1, forced expiratory volume in one second; FVC, forced vital capacity; DLCO, diffusing capacity of the lungs for carbon monoxide; PaO2, partial pressure of oxygen in arterial blood; PaCO2, partial pressure of carbon dioxide.

[a]*Estimated values for residents at the altitude of Mexico City.*

**Statistical difference with* $p < 0.05$.

noninvasive arterial blood pressure, and peripheral blood oxygen saturation. Also, a thoracic belt was used to acquire the respiratory signal and all signals were sampled at 1,000 Hz. Raw signals were acquired in supine position continuously during 10 min with the subjects breathing spontaneously ambient air (AA) and an additional 10 min breathing $SupplO_2$ at 3 L/min to ensure an arterial oxygen saturation above 94% (Santiago-Fuentes et al., 2021). Time series of successive beat-to-beat intervals (BBI), respiratory amplitude (RESP) at BBI onset as well as systolic (SYS) and diastolic (DIA) blood pressure (BP) were extracted from recorded signals; all extracted time series were manually reviewed and corrected. For GC analysis, the variability time series were resampled at 2 Hz using spline interpolation and normalized to zero mean and unit variance. For dynamic data analysis, consecutive windows of 5 min shifted by 30 s (90% overlap) were used. Therefore, the influence of $SupplO_2$ on IPF was studied using 31 windows including three phases labelled as ambient-air (AA), transition phase (TPH) and steady supplemental oxygen ($SupplO_2$), as indicated in **Figure 1**. The transition phase (TPH) is characterized by windows sharing AA and starting $SupplO_2$ conditions. In each window, univariate, and bivariate indices, as well as the ePDC were estimated.

2.2 Univariate and Bivariate Analyses

2.2.1 Linear Analysis

Time and frequency domain linear indices were extracted as the mean value, the root mean square of successive differences (rmssd), and power in the very low-frequency range (VLF, 0.003–0.04 Hz), low (LF, 0.04–0.15 Hz), and high frequency (HF, frequencies >0.15 Hz) bands in agreement with the standardization proposed by the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (AuthorAnonymous, 1996). The index rmssd is considered the more precise marker of chronotropic cardiac vagal influence (Minarini, 2020), for example, it provided a good differentiation between healthy subjects and vasovagal syncope patients for systolic and diastolic blood pressure variability (Reulecke et al., 2016). For estimating the power of LF and HF bands a parametric autoregressive model was used in each temporal window along of the variability time series. The model parameters and optimal order were estimated using the Burg method and the Akaike Information Criterion (AIC), respectively.

2.2.2 Nonlinear Analysis

To explore the nonlinear time series properties two techniques were applied: Detrended Fluctuation Analysis (DFA) to quantify the fractal scaling properties of a time series, and Symbolic Dynamics (SD), which is a coarse-grained method based on symbols (Reulecke et al., 2016). For DFA analysis, the short-term fractal scaling exponent (α_1) was calculated over equal and non-overlapping segments with length between 4 and 16 beats while the long-term exponent (α_2) for segments with length of 16–64 beats (Peng et al., 1995). Furthermore, by SD technique, nonlinear indices related to the low or high variability of the signal were determined. An alphabet, consisting only of symbols "0" or "1", was used to create words of length six and for BBI time series, thresholds of 2, 5, 10 and 20 ms were established (Voss et al., 2015). In this sense, a low variability index associated with the word "000000" and a high variability index associated to the word "111111", were counted (Schulz and Voss, 2017). In the case of blood pressure variability (BPV), the 6-length words were created by selecting thresholds of 1, 2, 3 and 4 mmHg (plvar 3 and 4) (Reulecke et al., 2016).

2.2.3 Bivariate Analysis

The Dual Sequence Method (DSM) is a linear technique to estimate the arterial baroreflex sensitivity (BRS) by the analysis of spontaneous fluctuations in systolic BP and BBI time series. Traditionally, a pattern of three consecutive increments or decrements in SYS and BBI are tagged as bradycardic (bslope, increase in SYS that causes an increase in BBI) or tachycardic (tslope, a decrease in SYS that causes a decrease in BBI) sequences, respectively. However, we avoided using the classical thresholds for SYS and BBI due to the criticisms about them (Gouveia et al., 2007) and in consequence, different pattern length was used including one up to three samples (N1-N3). Furthermore, the synchronous responses in the same beat interval (T0) and delayed BBI responses shifted by one up to three beats (T1-T3) were also explored, as suggested by (Malberg et al., 2002; Gouveia et al., 2007).

2.3 Multivariate Autoregressive Modeling, Extended Partial Coherence and Time-Frequency Analysis of Interactions

Establishing the cause and effect, or the driver-response relationship, between physiological systems has been of great interest in diverse biomedical applications. Fundamental to the driver-response relationship is the concept of Granger causality which states that if a signal improves the prediction of a second signal, above and beyond, its prediction in terms of its own past, then the first signal causes the second one. GC has been formalized in terms of multivariate time series analysis and its treatment in the frequency domain leading to the concept of partial directed coherence. Particularly, the extended partial

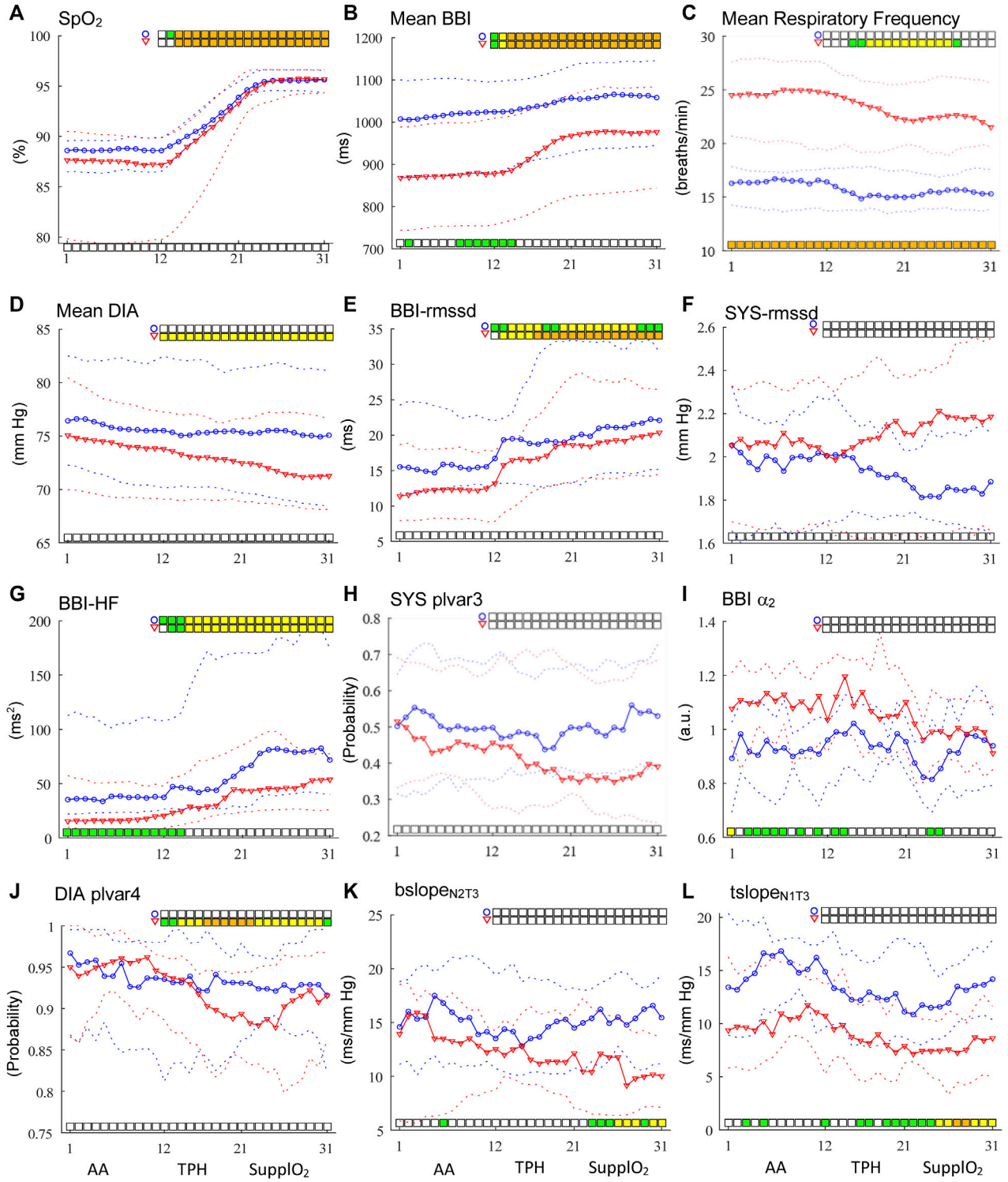


FIGURE 1 | Time courses of linear and nonlinear indices for BBI, SYS and DIA time series **(A–L)**. AA phase occurs from windows 1 to 11, transition phase (TPH) from 12 to 20 while the steady SupplO_2 from windows 21 to 31. Circles in blue represent healthy subjects and red triangles represent IPF patients. The solid line indicates the median values and broken lines indicate the interquartile range from 25 to 75%. Statistically significant differences between groups are shown with bars at the bottom of each graph while differences within-groups with bars at the top ($p < 0.05$ in green, $p < 0.01$ in yellow and $p < 0.001$ in orange).

directed coherence (ePDC) is based on fitting a linear time-invariant parametric model to the observed set of M time series $\boldsymbol{Y}(n)$ including instantaneous effects, i.e.,

$$\boldsymbol{Y}(n) = \sum_{k=0}^{q} \boldsymbol{B}(k)\boldsymbol{Y}(n-k) + \boldsymbol{W}(n), \quad (1)$$

where the model coefficients $\boldsymbol{B}(k)$, $k = 0, \ldots, q$, are related to instantaneous and strictly causal effects, q is the model order and $\boldsymbol{W}(n)$ is the innovation process formed by white and uncorrelated noises with diagonal covariance matrix $\boldsymbol{\Lambda} = \text{diag}(\lambda_i^2)$. The identification of the extended MVAR model can be achieved from a strictly causal MVAR model with coefficient matrix $\hat{\boldsymbol{A}}(k)$, $k = 1, \ldots, q$, as $\hat{\boldsymbol{B}}(k) = [\boldsymbol{I} - \boldsymbol{B}(0)]\hat{\boldsymbol{A}}(k)$ (Faes et al., 2010; Charleston-Villalobos et al., 2019). Furthermore, to calculate $\boldsymbol{B}(0)$ is necessary to perform a Cholesky decomposition of the input covariance matrix $\Sigma = L\Lambda L^T$, with $\Sigma = cov(\boldsymbol{U}(n))$, of the strictly causal model to obtain the matrix $\boldsymbol{L} = (\boldsymbol{I} - \boldsymbol{B}(0))^{-1}$. From this decomposition $\boldsymbol{L}$ is a lower triangular matrix as well as $\boldsymbol{B}(0)$ with null diagonal, but it is required to order the times series in $\boldsymbol{Y}(n)$ to set the direction, but not the strength, of physiological influence among them. Consequently, in the present study, to achieve the former constrain, two models were proposed: i) MVAR model 1 (MVAR_1) where $\boldsymbol{Y}(n)$ was built up as y_1 = RESP, y_2 = SYS and y_3 = BBI and ii) MVAR model 2 (MVAR_2) with y_1 = RESP, y_2 = DIA, and y_3 = BBI. The model order q was obtained by the minimum of the function $AIC(q) = Nlog(det\boldsymbol{\Sigma}) + M^2q$, where $\boldsymbol{\Sigma}$ is the input covariance matrix of the strictly causal model. From the frequency domain representation of **Eq. 1** is possible to obtain the spectral density of $\boldsymbol{Y}(f)$ and its inverse, from which the partial coherence (PC) function $\Pi_{ij}(f)$ between y_i and y_j is defined. However, PC cannot provide information about causality due to its symmetrical nature and, consequently, a factorization of $\Pi_{ij}(f)$ is necessary to produce ePDC, named $\chi_{ij}(f)$, as:

$$\chi_{ij}(f) = \frac{\left(\frac{1}{\lambda_i}\right)\bar{B}_{ij}(f)}{\sqrt{\sum_{m=1}^{M}\left(1/\lambda_m^2\right)\left|\bar{B}_{mj}(f)\right|^2}}. \quad (2)$$

where $\bar{B}_{ij}(f) = 1 - \sum_{k=0}^{q} B_{ij}(k)e^{-2\pi fkT}$ with T equal to the sampling period and λ_i is a diagonal value of the noise covariance matrix of the $\boldsymbol{W}$ process in **Eq. 1**. Therefore, ePDC is a directional frequency domain measure of connectivity quantifying the influence of the process y_j on the process y_i, removing the influence of other processes. The ePDC is normalized with respect to the structure that sends the signal taking values between 0 (absence of causal coupling) and 1 (full causal coupling) at frequency f. In this study, the identification of eMVAR model was performed by the standard least-squares method and the estimated ePDC was corrected by statistical hypothesis testing based on setting a threshold for significance using causal Fourier transform surrogates (Faes et al., 2010; Charleston-Villalobos et al., 2019). In a few words, once the eMVAR model is estimated, if there is a direct causality from the process y_j to y_i, the corresponding coefficients $b_{ij}(k, w_l)$, $k = 0, \ldots, q$, at each temporal window (w_l) are set to zero, and from there the surrogates are obtained; causality is selectively destroyed only over the direction under study, details of the procedure can be found in Faes et al., 2010. To accept or reject the estimated ePDC, its magnitude at each frequency was compared with the threshold obtained from 100 surrogates (95% confidence interval). If the magnitude of ePDC was below the threshold, the null hypothesis was accepted, reflecting the absence of interaction, and the estimated value was set to zero. In the present study, the time series of variability were segmented producing a time-frequency representation of interactions (TFR_i). In TFRi, the x-axis is related to the 31 temporal windows of 5 min and the y-axis depicts the frequency in Hz (0.0–0.50 Hz) using 512 bins. Furthermore, the z-axis represents the magnitude of the ePDC (0–0.4) by a color palette from blue to red, respectively. To count with the same number of samples in each window a resampling was performed at 2 Hz. It is worthy to note that, at least for this research protocol, the TFRi without or with resampling are close to each other and the dynamic changes produced by IPF, and supplemental oxygen remain.

2.4 Statistical Analysis

To evaluate the effect of supplemental oxygen, in this study the statistical analysis was performed with two statistical test, within-groups and between-groups. The within-groups test was carried out on univariate and bivariate indices values comparing window 5 vs. windows 12 to 31 (TPH and steady $\text{Suppl}O_2$) by the Wilcoxon Sign Rank-sum. All p-values were corrected using the Benjamini and Hochberg correction (Benjamini and Hochberg, 1995). In the case of between-groups comparison, the analysis was carried out on the univariate and bivariate indices values, and ePDC magnitude at each tile, in each temporal window by the nonparametric Mann-Whitney-U-test. In both statistical analyses, the significance was set at three levels for descriptive purposes: slightly significant for $p < 0.05$ (green), moderately significant for $p < 0.01$ (yellow), and highly significant for $p < 0.001$ (orange).

3 RESULTS AND DISCUSSION

The section is organized presenting first the results for the AA phase followed by the findings in the TPH and steady $\text{Suppl}O_2$ phases. The $\text{Suppl}O_2$ effect on healthy subjects and IPF patients is obtained looking for the time course of linear and nonlinear indices as well as by the dynamics of cardiovascular and respiratory interactions as compared with those in the AA phase. Statistical analysis between-groups, throughout all the phases, is displayed at the bottom of the graphs while within-group, comparing window 5 vs. windows 12 (start of TPH) to 31, is depicted at the top. For saving the space of article, selected indices are included in **Figure 1**, while others are only discussed. Also, the indices in **Figure 1** were ordered as they are used throughout the manuscript.

3.1 Clinical and Functional Measurements

According to **Table 1**, the IPF group was characterized by moderate reduction in vital capacity and FEV1, mild reduction in diffusion capacity, hypoxemia, mild hypercapnia, and high respiratory rate. In contrast, the CON

group had no signs of pulmonary alterations; although the $PaCO_2$ and PaO_2 were not registered at the time of the study, it is plausible to estimate them at the altitude of Mexico City (Vazquez-Garcia and Perez-Padilla, 2000). Consequently, the estimated normal values of $PaCO_2$ and PaO_2 for Mexico City residents are around 32.7 and 65.9 mmHg, respectively; there were no statistically significant differences in age, anthropometric measures, hematocrit, and hemoglobin values. As can be seen in **Figure 1A**, peripheral oxygen saturation (SpO_2) increased importantly from TPH towards the steady $SupplO_2$ phase for both groups until a saturation of 96% was reached, but no statistically significant differences between-groups were found during $SupplO_2$. In contrast, within-group statistical analysis revealed highly significant differences. Furthermore, Rico et al. observed that S_PO_2 tends to decline during the aging process that is in line with the SpO_2 in our Control group (Rico et al., 2001). On the other hand, the IPF group showed a greater dispersion of SpO_2 during AA phase that may be explained by age, disease stage, and pulmonary condition. Also, it is relevant to point out that the population of this age was under medication for other comorbidities. A comprehensive discussion of this aspect for the groups under study can be found in (Santiago-Fuentes et al., 2021).

3.2 Linear and Nonlinear Univariate Analysis in AA

For mean BBI index, statistically significant differences between CON and IPF groups were found from windows 8 to 12, **Figure 1B**, while for the mean respiratory frequency highly significant differences were found for the whole phase, **Figure 1C**. Also, the mean SYS and DIA BP for CON tended to be higher than for IPF, **Figure 1D**. For linear indices as rmssd, there were no significant differences for BBI, SYS or DIA, **Figures 1E,F**, but for BBI the IPF group showed a tendency to lower values. Also, for BBI power in the LF and HF bands was significantly different ($p < 0.05$); particularly, the CON group showed higher BBI-HF power and consequently, higher cardiac vagal influence than IPF, **Figure 1G**. For SYS-nLF, IPF showed a tendency to higher values, i.e., increased sympathetic influence on the vasculature than in CON. In the case of SD nonlinear analysis, for IPF the indices BBI-phvar2 as well as SYS-plvar3 (**Figure 1H**) showed a tendency to lower probability values than CON, i.e., low BBI variability and higher SYS variability, respectively. Nevertheless, in the DFA analysis, BBI-α_2 was statistically different ($p < 0.05$) between groups, where IPF patients showed higher sympathetic activity than CON, **Figure 1I**. Consequently, by univariate analysis, statistically significant differences between groups were found just for BBI, indicating a significant cardiac sympathetic modulation in IPF patients during AA scenario.

3.3 Linear and Nonlinear Univariate Analysis in TPH and Steady $SupplO_2$

At the firsts windows in TPH, BBI showed statistically significant differences between-groups for mean, BBI-HF, and BBI-α_2 indices, **Figure 1**. Specifically, the mean cardiac frequency decreased in both groups, slowly for CON and more drastically for IPF, whereas for HF and α_2, the CON group showed higher and lower values than IPF, respectively. The former behavior is in line with the expected effect of $SupplO_2$ on the cardiac frequency. The mean respiratory frequency was reduced for both groups however, the between groups differences were kept along both phases, **Figure 1C**. Although between-groups no statistically significant differences were found for mean SYS and DIA in TPH or steady $SupplO_2$, the mean SYS showed a tendency to increase in CON while the mean DIA decreased in IPF.

To evaluate the influence of O_2 within each group, statistical comparisons were achieved between AA and TPH or steady $SupplO_2$ phases. The effect of O_2 was significantly different for CON and IPF. Specifically, significant differences were obtained for BBI, mean respiratory frequency, and DIA. For the IPF group, the BBI-rmssd index showed highly significant differences ($p < 0.001$) from windows 16 to 31, i.e., the cardiac vagal influence in IPF was increased to a greater extent by O_2, **Figure 1E**. The former time course was supported by the nonlinear BBI-phvar2 index ($p < 0.001$). Furthermore, BBI-LF and BBI-α_1 indices provided significant within-group differences. Regarding BBI-LF index, only for the IPF group, the LF power increased from TPH and throughout steady $SupplO_2$ phase ($p < 0.001$) while the BBI-α_1 index decreased. It is worthy to note that for SYS, in none of the groups neither linear nor nonlinear indices provided significant differences. Although, for IPF, SYS-rmssd had a tendency to increase while SYS-plvar3 to decrease, indicating that SYS increased its variability with O_2, **Figures 1F,H**. In the case of DIA, just for IPF, mean DIA decreased due to oxygen from TPH to the end of the steady $SupplO_2$ phase in a moderately significant way ($p < 0.01$), that may be related to the lower cardiac frequency with O_2, i.e., a vagal modulation effect, **Figure 1D**. Also, the DIA-plvar4 index showed moderately and highly significant differences along the two phases, the percentage of words of low variability was reduced, i.e., in the IPF group the DIA variability was increased by $SuppO_2$, **Figure 1J**.

3.4 Bivariate Analysis in AA, TPH and Steady $SupplO_2$

For DSM analysis, different pattern lengths and delays were tested, and the results indicated that statistical differences between groups occurred with lengths of 2–3 samples and shifts between 0 and 3 beats. The bivariate indices during AA, associated with arterial baroreflex sensitivity, did not show statistical differences between groups, only a tendency to lower values for IPF. The former behavior is in line with the literature indicating that hypoxia produces a resetting of arterial baroreflex, without changing the sensitivity, to higher heart rates and systolic blood pressures due to stimulation of peripheral chemoreceptors (Halliwill et al., 2003). In contrast, the IPF patients of the present study were characterized by significantly higher cardiac frequency and cardiac sympathetic modulation but similar systolic pressure. It is worthy to mention that the interaction between baroreflex and peripheral chemoreflexes remains controversial (Kronsbein

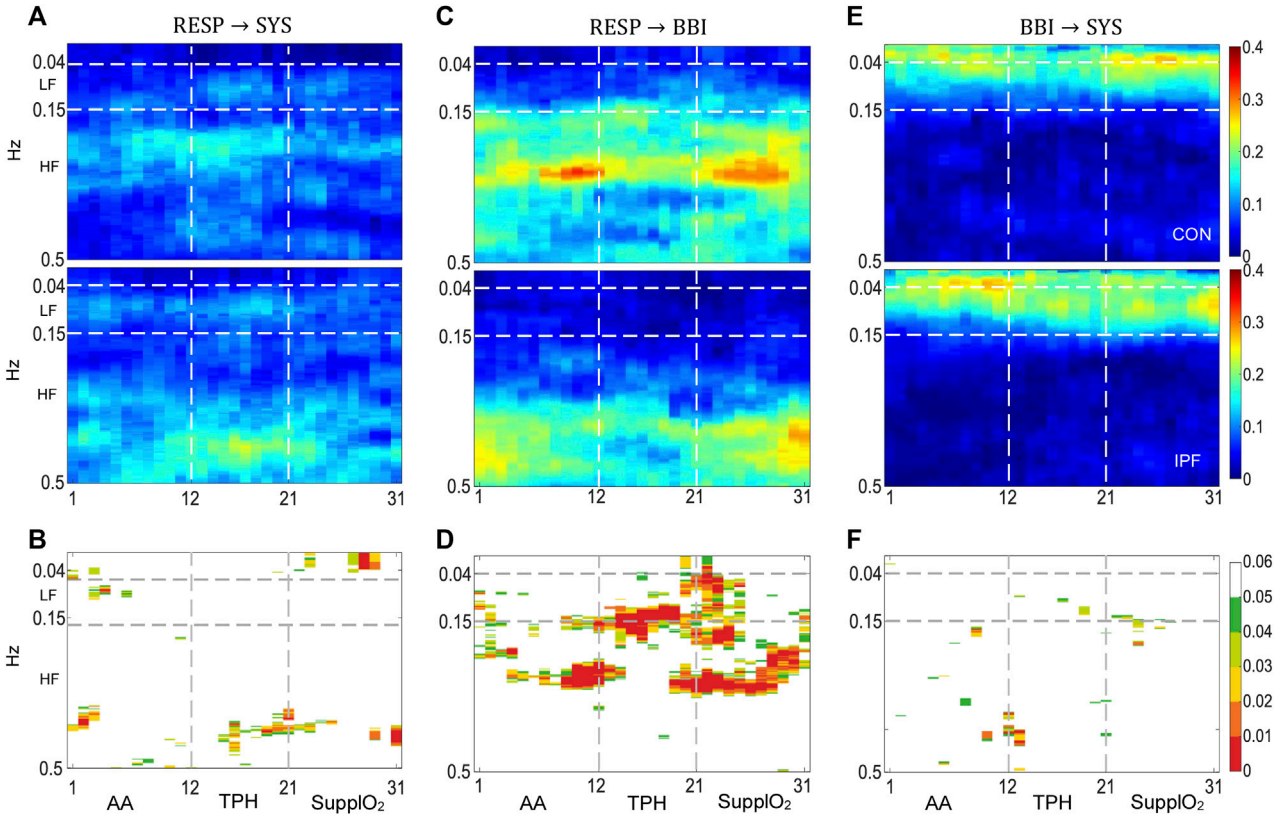


FIGURE 2 | Cadiorespiratory and cardiovascular interactions by the model $MVAR_1$. **(A,C,E)** TFR_i of RESP → SYS, RESP → BBI and BBI → SYS interactions, respectively, **(B,D,F)** the map of corresponding statistical differences between groups [$p < 0.05$ (green), $p < 0.03$ (yellow), $p < 0.01$ (red)]. For each interaction, the TFRi of the control group is at the top while the one for patients is at the bottom of the panel. Horizontal broken lines indicate the LF and HF bands while the vertical lines mark the phases of the protocol.

et al., 2020). During steady $SupplO_2$, from window 21, bslope and tslope indices provided significant differences ($p < 0.05$ and $p < 0.01$) between groups, **Figures 1K, 1**L. For both indices, IPF showed lower values than CON, i.e., IPF decreased their baroreflex in comparison to CON. In the case of within-groups analysis, in TPH and steady $SupplO_2$ phases, bslope tends to increase for CON while for the IPF group tends to decrease. Systolic pressure, being a variable to control, tends to oscillate with greater amplitude when feedback mechanisms are deficient, showing an inverse relationship with variations in heart rate, as a variable to regulate (Mancia et al., 1986; Lanfranchi and Somers, 2002). In fact, it should be noted that the significant reduction in cardiac output in IPF observed during $SupplO_2$ (Santiago-Fuentes et al., 2021) could be a consequence of alteration of the cardiovascular regulation. The former behavior may be explained, on the one hand, by statistically significant changes (attenuation) of the sensitivity of the baroreceptors reflected by bslope and tslope indices (**Figures 1K,L**), increase in BBI (**Figure 1B**), and in total peripheral resistance and, on the other hand, by statistically non-significant trends in systolic blood pressure and its variability. Furthermore, diverse research pointed out that clinicians should be aware of the prognostic implications of increased blood pressure variability, a marker of cardiovascular decompensation, which may lead for example to organ damage (Hoecht, 2013).

3.5 Interactions Between Cardiovascular and Respiratory Systems

3.5.1 TFR_i Magnitude Distribution in AA

An averaged time-frequency representation of RESP → SYS interaction by $MVAR_1$, associated with the information flow from RESP to SYS, is showed in **Figure 2A**, for CON (above) and IPF group (below). The TFR_i magnitude distribution points out differences between groups for the LF and HF bands, the significant differences are displayed in **Figure 2B**. The power in the LF band has been associated with the sympathetically mediated BP vasomotor modulation, the so-called Mayer waves around 0.1 Hz. The origin of Mayer waves has not been elucidated, but some authors agreed that these waves could be related to an oscillatory sympathetic activation that, in the specific case of humans, is independent of factors such as gender, age, or posture and can be induced in BP by hypoxia (Julien, 2006; Ghali and Ghali1, 2020). According to the RESP → SYS interaction in the LF band, the TFRi magnitude for IPF is higher than for CON

that is in line with the SYS univariate results presented above; the IPF group showed a tendency to an augmented sympathetic activity in AA, as revealed by SYS-rmssd (**Figure 1F**). Furthermore, for the CON group the magnitude of the TFR_i in the HF band is concentrated around 0.28 Hz while for IPF group is spread out around 0.42 Hz, i.e., around the respective mean respiratory frequency, **Figure 1C**. It is plausible that different mechanical effect on central blood volumes is produced by higher respiratory frequency and lower FVC in IPF than in CON. Also, in the HF band Traube-Hering waves (THW) have been reported associated with respiratory-related fluctuations in sympathetic outflow that promotes changes in vascular tone (Menuet et al., 2020). In fact, the IPF group showed a significant increased TPR during AA, i.e., increased vascular tone (Santiago-Fuentes et al., 2021). In the case of SYS → BBI interaction there were no statistically significant differences and consequently, the corresponding TFR_i is not shown. A possible explanation could be associated to the fact that subjects were in supine position and then, baroreflex feedback control may be blunted. The RESP → BBI interaction, **Figure 2C**, shows that the magnitude for CON is spread out over the whole HF band and is higher than for IPF that is localized around the mean respiratory frequency, the corresponding statistical differences are displayed in **Figure 2D**. Therefore, the magnitude of the RESP → BBI interaction pointed out that respiratory sinus arrhythmia (RSA), associated with the cardiac vagal influence, occurs at different operating point in CON and IPF. The former result is in line with other studies which suggest that RSA decreases with hypoxia (Yasuma et al., 2001) and in fact, the IPF group of the present study was characterized by hypoxemia (**Table 1**). Furthermore, RSA was more relevant for CON as the magnitude of the RESP → BBI interaction is higher than the corresponding of the RESP → SYS interaction, i.e., the influence of RESP on the heart is higher than on the vascular subsystem. Regarding the BBI → SYS interaction, **Figure 2E**, associated with the mechanical feedforward influence, a relevant magnitude is located towards the VLF and LF bands for both groups; however, there were no statistically significant differences.

Results by $MVAR_2$, using DIA instead of SYS are shown in **Figure 3**. The TFR_i of the RESP → DIA interaction is shown in **Figure 3A** and resembles the TFR_i of RESP → SYS, but for both groups the magnitude spread further in the LF and HF bands, statistically significant differences are shown in **Figure 3B**. For the DIA → BBI interaction, **Figure 3C**, the TFRi displayed a magnitude with a trend to be higher in CON, particularly in the upper part of the HF band, while for the IPF group a more uniform distributed magnitude is observed across the frequency bands. Furthermore, for both groups the DIA → BBI interaction presents a higher magnitude than SYS → BBI interaction. The former behavior may be interpreted in terms of an impaired left ventricular (LV) diastolic filling in IPF in contrast to a preserved LV systolic function, as was shown in a previous study (Papadopoulos et al., 2008). Also, the lower magnitude in the DIA → BBI interaction for the IPF group may reflect peripheral vascular changes due to increased arterial stiffness and in conjunction with a high heart rate may be indicative of the prevalence of sympathetic tone. For the BBI → DIA interaction, the TFR_i magnitude distribution resembles the corresponding to BBI → SYS of $MVAR_1$, **Figure 3E**, but with a tendency to display lower values for both groups. Furthermore, a consistent activity in the LF band was found, the Mayer wave. For the DIA → RESP interaction, the corresponding TFR_i did not show relevant magnitude in any phase and, consequently it was not included in the paper. Therefore during the AA phase, based on the TFR_i magnitude, its spread and relevance of statistically significant differences between groups, IPF during the AA phase mainly impacted the RESP → SYS and RESP → BBI interactions.

3.5.2 TFR_i Magnitude Distribution in TPH and Steady $SupplO_2$

In the case of RESP → SYS interaction, during TPH the $SupplO_2$ increased the TFR_i magnitude in both groups, **Figure 2A**. For the CON group the magnitude spread further in the LF and HF (the Traube-Hering waves) bands as compared with the AA phase. In contrast, for the IPF group the magnitude was increased mainly in the upper part of the HF band, the statistically significant differences are shown in **Figure 2B**. Furthermore, in the steady $SupplO_2$ phase, the TFR_i magnitude decreased in both groups but remains higher in IPF, it seems that the interaction for both groups reached an operation level like the level in AA phase. Presumably, in IPF the heightened arterial peripheral chemoreceptors activity/sensitivity by the hypoxia keeps the drive of RESP on SYS (Stickland et al., 2016). In fact, chronic hypoxia produces among other things, hyperplasia of the carotid body and increases its activity. As can be seen in **Figure 1C**, throughout the $SupplO_2$ phase the mean respiratory frequency decreased in the IPF group, with a moderate *p*-value with respect to AA, but the influence on SYS does not change significantly.

During TPH, the RESP → BBI interaction reveals that RSA was highly significant different between CON and IPF groups, **Figures 2C,D**, i.e., RSA in CON was stronger. The former result is in line with the fact that vagal modulation facilitates RSA at rest while sympathetic activation attenuates its magnitude. Moreover, for the RSA physiological purpose, there are three major hypotheses, i.e., 1) improvement of gas exchange, 2) minimization of the energy consumption for the heart, and 3) reduction of BP fluctuations (Buchner, 2019). One of the premises of hypothesis 3) is that suppression of RSA generally increases BP fluctuations in its HF band. In fact, the former hypothesis may explain the decrease in RSA for the IPF group that may be related to the tendency to increase of SYS-rmssd and SYS-phvar1 indices, through $SupplO_2$ as discussed above. During the steady $SupplO_2$ phase the RESP → BBI interaction showed highly significant differences. For CON, the TFR_i significant magnitude in the HF band, almost in the whole phase, reveals that $SupplO_2$ affects importantly respiratory drive on BBI. In contrast, for the IPF group the RESP → BBI interaction remains at similar operation level as previous phases, in consequence, the RSA in IPF was lower than in CON. It is noteworthy that in $SupplO_2$ phase, for the CON group the RESP → BBI (RSA) interaction increased while RESP → SYS magnitude (THW) decreased (Barnett et al., 2020). In the case of BBI → SYS, for IPF during TPH, the TFR_i magnitude showed a sustained level of driving indicating a higher influence of the cardiac mechanical

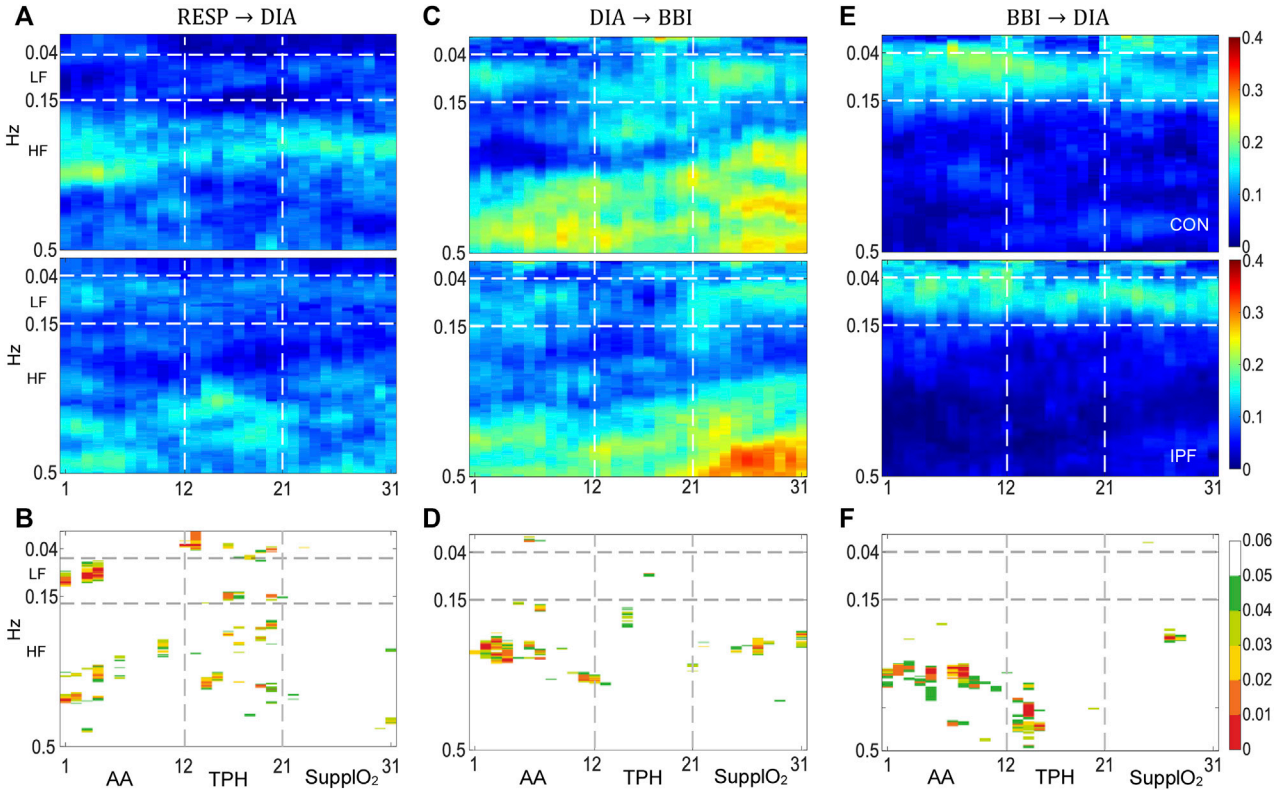


FIGURE 3 | Cadiorespiratory and cardiovascular interactions by the model $MVAR_2$. **(A,C,E)** TFR_i of RESP → DIA, DIA → BBI and BBI → DIA interactions, respectively, **(B,D,F)** the map of corresponding statistical differences between groups [$p < 0.05$ (green), $p < 0.03$ (yellow), $p < 0.01$ (red)]. For each interaction, the TFRi of the control group is at the top while the one for IPF patients is at the bottom of the panel. Horizontal broken lines indicate the LF and HF bands while the vertical lines mark the phases of the protocol.

modulation, **Figure 2E**. In steady $SupplO_2$, in the LF band there were no significant differences, but there was a trend in IPF towards a greater magnitude, which is associated with sympathetic modulation on the vasculature.

By $MVAR_2$ using DIA instead of SYS, the RESP → DIA interaction reveals that the TFRi magnitude for CON and IPF is concentrated around the mean respiratory frequency of each group as occurs for RESP → SYS, **Figures 3A,B**. It seems that the oxygen delivery did not influence significantly the information flow from RESP to DIA. Regarding the DIA → BBI interaction, **Figure 3C**, the TFR_i magnitude in steady $SupplO_2$ phase showed some differences between groups in the HF band, and mainly toward the end of the phase, statistically significant differences are depicted in **Figure 3D**. A plausible explanation is that the diastolic BP is related to the pumping activity of the aorta and large arteries of the circulating system. So factors such as systemic vascular resistance or arterial stiffness modify diastolic BP more than systolic BP, and therefore, the detriment of baroreflex activity at rest, as reflected by the tslope index in Figure 1L, is more evident in DIA → BBI than in SYS → BBI. In the case of the BBI → DIA interaction, **Figure 3E**, the TFR_i resembles the behavior of BBI → SYS interaction in the LF band and it seems to be slightly affected by oxygen but there were no statistically significant differences, **Figure 3F**. It is worthy to note that stationarity is assumed for temporal windows of 5 minutes of cardiovascular and respiratory time-series. However, during TPH nonstationary behavior may be elicited and consequently, a modified time-frequency approach could be applied.

4 CONCLUSION

According to the results of the present study, the IPF group in the AA phase, as compared with a matched healthy group, was characterized by high sympathetic modulation to the heart (supported by BBI-rmssd, BBI-α_2 indices, and increased mean respiratory frequency, among other indices) with a decreased influence of the parasympathetic activity. It is well-known that chemoreceptors sense the partial pressure of oxygen in blood vessels, and they are important modulators of sympathetic activation in response to hypoxemia. The activation is also known as chemoreflex-mediated sympathetic activation and one of the consequences is the hyperventilation that leads to the inhibition of the chemoreflex due to the stretch mechanoreceptors activity in the thoracic cage (Kara et al., 2003). Also, the IPF patients during $SupplO_2$ had blunted baroreflex (confirmed by DSM results and almost inexistent SYS → BBI

interaction), as well as RSA activity (confirmed by BBI-HF and RESP → BBI interaction). The former autonomic behavior could be explained partially by the chronic hypoxemia and hypercapnia in the IPF group. In this study, the patient group suffered mild hypoxia and some evidence of hypercapnia, so the enhanced sympathetic activity and reduced parasympathetic response may be a consequence of alterations in baroreceptors and peripheral chemoreceptors, which reduce the response to supplemental oxygen. In fact, Van Gestel and Steier showed evidence related to the high sympathetic activity in COPD patients due to the impaired baroreflex sensitivity altered by hypoxia but not by hypercapnia (Van Gestel and Steier, 2010). Additionally, an impairment of cardiac control and sympathetic overactivity may be related to age, but these responses seem to be higher for the IPF group. In fact, Porta et al. at supine rest phase, found a possible age-related impairment of the cardiac control and altered response to stressors in conjunction with a gradual decrease in SYS complexity and thus, an increase of the sympathetic activity (Porta et al., 2014).

The effect of oxygen has been studied mainly by its hemodynamic effects and the analysis of heart rate variability (Lund et al., 1999; Gole et al., 2011; Smit et al., 2018a; Smit et al., 2018b). To the best of our knowledge, this is the first study that tackle the effect of oxygen supplementation in IPF with a new perspective from the point of view of linear and non-linear indices as well as the dynamics of cardiovascular and respiratory systems interactions. For IPF patients the results showed that during AA phase: 1) the mean BBI value and power of BBI-HF band, as well as the mean respiratory frequency were significantly lower ($p < 0.05$) and higher ($p < 0.001$), respectively, indicating a strong sympathetic influence, and 2) the RESP → SYS interaction was characterized by Mayer waves and diminished RESP → BBI, i.e., decreased respiratory sinus arrhythmia. In contrast, for IPF during short-term SupplO_2 phase: 1) oxygen might produce a negative influence on the systolic blood pressure variability (SYS-rmssd was increased among other indices), 2) the arterial baroreflex reduced significantly ($p < 0.01$), and 3) reduction of RSA (RESP → BBI) with simultaneous increase of Traube-Hering waves (RESP → SYS), reflected increased sympathetic modulation to the vessels. Our study in patients with IPF, compared with control subjects residing at the same altitude level, suggests that the autonomic alterations induced by the pathology persist or worsen despite the acute administration of oxygen. Based on our previous effort, indicating a relevant increase of TPR (Santiago-Fuentes et al., 2021), current research by the group is directed to analyze interactions including TPR. Finally, the proposed TFRi analysis may be used to better understanding the underlying physiological phenomena of different respiratory diseases.

AUTHOR CONTRIBUTIONS

Conceptualization: LS-F, SC-V, RG-C, and TA-C; Methodology: LS-F, SC-V, SR, RG-C, and TA-C; Software: LS-F, SR, and SC-V; Validation: SC-V, AV, RG-C, and TA-C; Data analysis: LS-F, SC-V, and TA-C, Data acquisition: LS-F, SR, MM-A, and IB-R; Writing—original draft preparation: LS-F, SC-V, and TA-C; writing—review and editing: LS-F, SC-V, AV, and RG-C; Final version approval: LS-F, SC-V, AV, RG-C, MM-A, IB-R, SR, and TA-C.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by National Institute of Respiratory Diseases (INER). The patients/participants provided their written informed consent to participate in this study.

REFERENCES

AuthorAnonymous (1996). Heart Rate Variability: Standards of Measurement, Physiological Interpretation and Clinical Use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation* 93 (5), 1043–1065.

Barnett, W. H., Latash, E. M., Capps, R. A., Dick, T. E., Wehrwein, E. A., and Molkov, Y. I. (2020). Traube-Hering Waves Are Formed by Interaction of Respiratory Sinus Arrhythmia and Pulse Pressure Modulation in Healthy Men. *J. Appl. Physiol.* 129 (5), 1193–1202. doi:10.1152/japplphysiol.00452.2020

Beckers, F., Verheyden, B., and Aubert, A. E. (2006). Aging and Nonlinear Heart Rate Control in a Healthy Population. *Am. J. Physiology-Heart Circulatory Physiol.* 290 (6), H2560–H2570. doi:10.1152/ajpheart.00903.2005

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B (Methodological)* 57 (1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Buchner, T. (2019). A Quantitative Model of Relation between Respiratory-Related Blood Pressure Fluctuations and the Respiratory Sinus Arrhythmia. *Med. Biol. Eng. Comput.* 57 (5), 1069–1078. doi:10.1007/s11517-018-1939-4

Buendía-Roldán, I., Mejía, M., Navarro, C., and Selman, M. (2017). Idiopathic Pulmonary Fibrosis: Clinical Behavior and Aging Associated Comorbidities. *Respir. Med.* 129, 46–52. doi:10.1016/j.rmed.2017.06.001

Caminati, A., Lonati, C., Cassandro, R., Elia, D., Pelosi, G., Torre, O., et al. (2019). Comorbidities in Idiopathic Pulmonary Fibrosis: an Underestimated Issue. *Eur. Respir. Rev.* 28 (153), 190044. doi:10.1183/16000617.0044-2019

Charleston-Villalobos, S., Reulecke, S., Voss, A., Azimi-Sadjadi, M. R., González-Camarena, R., Gaitán-González, M. J., et al. (2019). Time-Frequency Analysis of Cardiovascular and Cardiorespiratory Interactions during Orthostatic Stress by Extended Partial Directed Coherence. *Entropy* 21 (5), 468–482. doi:10.3390/e21050468

Faes, L., and Nollo, G. (2010). Extended Causal Modeling to Assess Partial Directed Coherence in Multiple Time Series with Significant Instantaneous Interactions. *Biol. Cybern* 103 (5), 387–400. doi:10.1007/s00422-010-0406-6

Faes, L., Porta, A., and Nollo, G. (2010). Testing Frequency-Domain Causality in Multivariate Time Series. *IEEE Trans. Biomed. Eng.* 57, 1897–1906. doi:10.1109/TBME.2010.2042715

Ghali, M. G. Z., and Ghali, G. Z. (2020). Mechanisms Contributing to the Generation of Mayer Waves. *Front. Neurosci.* 14, 395. doi:10.3389/fnins.2020.00395

Gole, Y., Gargne, O., Coulange, M., Steinberg, J.-G., Bouhaddi, M., Jammes, Y., et al. (2011). Hyperoxia-induced Alterations in Cardiovascular Function and Autonomic Control during Return to Normoxic Breathing. *Eur. J. Appl. Physiol.* 111 (6), 937–946. doi:10.1007/s00421-010-1711-4

Gouveia, S., Rocha, A. P., Laguna, P., and Lago, P. (2007). Threshold Sensitivity in Time Domain BRS Estimation: Minimum Beat-To-Beat Changes and Minimum Correlation. *Comput. Cardiol.* 34, 557–560. doi:10.1109/CIC.2007.4745546

Halliwill, J. R., Morgan, B. J., and Charkoudian, N. (2003). Peripheral Chemoreflex and Baroreflex Interactions in Cardiovascular Regulation in Humans. *J. Physiol.* 552 (1), 295–302. doi:10.1113/jphysiol.2003.050708

Höcht, C. (2013). Blood Pressure Variability: Prognostic Value and Therapeutic Implications. *ISRN Hypertens.* 2013, 1–16. doi:10.5402/2013/398485

Julien, C. (2006). The enigma of Mayer Waves: Facts and Models. *Cardiovasc. Res.* 70 (1), 12–21. doi:10.1016/j.cardiores.2005.11.008

Kara, T., Narkiewicz, K., and Somers, V. K. (2003). Chemoreflexes - Physiology and Clinical Implications. *Acta Physiol. Scand.* 177 (3), 377–384. doi:10.1046/j.1365-201X.2003.01083.x

Khor, Y. H., Renzoni, E. A., Visca, D., McDonald, C. F., and Goh, N. S. L. (2019). Oxygen Therapy in COPD and Interstitial Lung Disease: Navigating the Knowns and Unknowns. *ERJ Open Res.* 5 (3), 00118–02019. doi:10.1183/23120541.00118-2019

King, T. E., Pardo, A., and Selman, M. (2019). Idiopathic Pulmonary Fibrosis. *Lancet* 378 (9807), 1949–1961. doi:10.1016/S0140-6736(11)60052-4

Kronsbein, H., Gerlach, D. A., Heusser, K., Hoff, A., Hoffmann, F., Diedrich, A., et al. (2020). Testing Individual Baroreflex Responses to Hypoxia-Induced Peripheral Chemoreflex Stimulation. *Clin. Auton. Res.* 30 (6), 531–540. doi:10.1007/s10286-019-00660-6

Lanfranchi, P. A., and Somers, V. K. (2002). Arterial Baroreflex Function and Cardiovascular Variability: Interactions and Implications. *Am. J. Physiology-Regulatory, Integr. Comp. Physiol.* 283 (4), R815–R826. doi:10.1152/ajpregu.00051.2002

Lund, V. E., Kentala, E., Scheinin, H., Klossner, J., Helenius, H., Sariola-Heinonen, K., et al. (1999). Heart Rate Variability in Healthy Volunteers during Normobaric and Hyperbaric Hyperoxia. *Acta Physiol. Scand.* 167 (1), 29–35. doi:10.1046/j.1365-201x.1999.00581.x

Malberg, H., Wessel, N., Hasart, A., Osterziel, K.-J., and Voss, A. (2002). Advanced Analysis of Spontaneous Baroreflex Sensitivity, Blood Pressure and Heart Rate Variability in Patients with Dilated Cardiomyopathy. *Clin. Sci.* 102 (4), 465–473. doi:10.1042/cs1020465

Mancia, G., and Grassi, G. (2014). The Autonomic Nervous System and Hypertension. *Circ. Res.* 114 (11), 1804–1814. doi:10.1161/CIRCRESAHA.114.302524

Mancia, G., Parati, G., Pomidossi, G., Casadei, R., Di Rienzo, M., and Zanchetti, A. (1986). Arterial Baroreflexes and Blood Pressure and Heart Rate Variabilities in Humans. *Hypertension* 8 (2), 147–153. doi:10.1161/01.hyp.8.2.147

Menuet, C., Connelly, A. A., Bassi, J. K., Melo, M. R., Le, S., Kamar, J., et al. (2020). PreBötzinger Complex Neurons Drive Respiratory Modulation of Blood Pressure and Heart Rate. *Elife* 9, e57288. doi:10.7554/eLife.57288

Minarini, G. (2020). "Root Mean Square of the Successive Differences as Marker of the Parasympathetic System and Difference in the Outcome after ANS Stimulation," in *Autonomous Nervous System Monitoring-Heart Rate Variability*. Editor T. Aslanidis (London: IntechOpen Press), 15–28. doi:10.5772/intechopen.89827

Mohammed, J., Da Silva, H., Van Oosterwijck, J., and Calders, P. (2017). Effect of Respiratory Rehabilitation Techniques on the Autonomic Function in Patients with Chronic Obstructive Pulmonary Disease: A Systematic Review. *Chron. Respir. Dis.* 14 (3), 217–230. doi:10.1177/1479972316680844

Mohammed, J., Meeus, M., Derom, E., Da Silva, H., and Calders, P. (2015). Evidence for Autonomic Function and its Influencing Factors in Subjects with COPD: A Systematic Review. *Respir. Care* 60 (12), 1841–1851. doi:10.4187/respcare.04174

Papadopoulos, C. E., Pitsiou, G., Karamitsos, T. D., Karvounis, H. I., Kontakiotis, T., Giannakoulas, G., et al. (2008). Left Ventricular Diastolic Dysfunction in Idiopathic Pulmonary Fibrosis: a Tissue Doppler Echocardiographic Study. *Eur. Respir. J.* 31 (4), 701–706. doi:10.1183/09031936.00102107

Peng, C. K., Havlin, S., Stanley, H. E., and Goldberger, A. L. (1995). Quantification of Scaling Exponents and Crossover Phenomena in Nonstationary Heartbeat Time Series. *Chaos* 5 (1), 82–87. doi:10.1063/1.166141

Porta, A., Faes, L., Bari, V., Marchi, A., Bassani, T., Nollo, G., et al. (2014). Effect of Age on Complexity and Causality of the Cardiovascular Control: Comparison between Model-Based and Model-free Approaches. *Plos One* 9 (2), e89463. doi:10.1371/journal.pone.0089463

Reulecke, S., Charleston-Villalobos, S., Voss, A., González-Camarena, R., González-Hermosillo, J., Gaitán-González, M. J., et al. (2016). Orthostatic Stress Causes Immediately Increased Blood Pressure Variability in Women with Vasovagal Syncope. *Comp. Methods Programs Biomed.* 127, 185–196. doi:10.1016/j.cmpb.2015.12.005

Rico, M. F. G., Urias, A. P., Barquera, C. S., Ochoa, J. L. G., Padilla, N. M. A., Meneses, G. L. C., et al. (2001). Spirometric and Gasometric Values in a Healthy Geriatric Population, at Different Altitudes above Sea Level in the Mexican Republic. Multicentric Study. *Rev. Inst. Nal Enf Resp Mex* 14 (2), 90–98.

Santiago-Fuentes, L. M., González-Camarena, R., Charleston-Villalobos, S., Mejía-Ávila, M. E., Reulecke, S., Buendía-Roldán, I., et al. (2021). Hemodynamic Response to Low-Flow Acute Supplemental Oxygen in Idiopathic Pulmonary Fibrosis and Elderly Healthy Subjects. *Heart & Lung* 50 (1), 197–205. doi:10.1016/j.hrtlng.2020.03.025

Schulz, S., and Voss, A. (2017). "Symbolic Dynamics, Poincaré Plot Analysis and Compression Entropy Estimate Complexity in Biological Time Series," in *Complexity and Nonlinearity in Cardiovascular Signals*. Editors R. Barbieri, E. Scilingo, and G. Valenza (Springer), 45–85. doi:10.1007/978-3-319-58709-7_2

Smit, B., Smulders, Y. M., Eringa, E. C., Oudemans - van Straaten, H. M., Girbes, A. R. J., Wever, K. E., et al. (2018a). Effects of Hyperoxia on Vascular Tone in Animal Models: Systematic Review and Meta-Analysis. *Crit. Care* 22 (1), 189. doi:10.1186/s13054-018-2123-9

Smit, B., Smulders, Y. M., Van der Wouden, J. C., Oudemans-Van Straaten, H. M., and Spoelstra-de Man, A. M. E. (2018b). Hemodynamic Effects of Acute Hyperoxia: Systematic Review and Meta-Analysis. *Crit. Care* 22 (1), 45. doi:10.1186/s13054-018-1968-2

Stickland, M. K., Fuhr, D. P., Edgell, H., Byers, B. W., Bhutani, M., Wong, E. Y. L., et al. (2016). Chemosensitivity, Cardiovascular Risk, and the Ventilatory Response to Exercise in COPD. *PLoS One* 11 (6), e0158341. doi:10.1371/journal.pone.0158341

Stoller, J. K., Panos, R. J., Krachman, S., Doherty, D. E., and Make, B. (2010). Oxygen Therapy for Patients with COPD. *Chest* 138 (1), 179–187. doi:10.1378/chest.09-2555

Van Gestel, A. J., and Steier, J. (2010). Autonomic Dysfunction in Patients with Chronic Obstructive Pulmonary Disease (COPD). *J. Thorac. Dis.* 2 (4), 215–222. doi:10.3978/j.issn.2072-1439.2010.02.04.5

Vazquez-Garcia, J. C., and Perez-Padilla, R. (2000). Valores gasométricos estimados para las principales poblaciones y sitios a mayor altitud en México. *Rev. Inst. Nal Enf Resp Mex* 13 (1), 6–13.

Voss, A., Schroeder, R., Heitmann, A., Peters, A., and Perz, S. (2015). Short-Term Heart Rate Variability-Influence of Gender and Age in Healthy Subjects. *Plos One* 10 (3), e0118308. doi:10.1371/journal.pone.0118308

Yasuma, F., Hirai, M., and Hayano, J.-i. (2001). Differential Effects of Hypoxia and Hypercapnia on Respiratory Sinus Arrhythmia in Conscious Dogs. *Jpn. Circ. J.* 65 (8), 738–742. doi:10.1253/jcj.65.738

You, J., Fan, X., Bi, X., Xian, Y., Xie, D., Fan, M., et al. (2018). Association between Arterial Hyperoxia and Mortality in Critically Ill Patients: a Systematic Review and Meta-Analysis. *J. Crit. Care* 47, 260–268. doi:10.1016/j.jcrc.2018.07.014

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 reviewer LF is currently organizing a Research Topic with one author SC-V.

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.

frontiers | Frontiers in Network Physiology

ORIGINAL RESEARCH
published: 28 April 2022
doi: 10.3389/fnetp.2022.845327

Partial Directed Coherence and the Vector Autoregressive Modelling Myth and a Caveat

Luiz A. Baccalá[1]* *and Koichi Sameshima*[2]

[1]*Laboratório de Comunicações e Sinais, Departamento de Telecomunicações e Controle, Escola Politécnica, Universidade de São Paulo, São Paulo, Brazil,* [2]*Departamento de Radiologia e Oncologia, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil*

Here we dispel the lingering myth that Partial Directed Coherence is a Vector Autoregressive (VAR) Modelling dependent concept. In fact, our examples show that it is *spectral factorization* that lies at its heart, for which VAR modelling is a mere, albeit very efficient and convenient, device. This applies to Granger Causality estimation procedures in general and also includes instantaneous Granger effects. Care, however, must be exercised for connectivity between multivariate data generated through nonminimum phase mechanisms as it may possibly be *incorrectly* captured.

Keywords: partial directed coherence, total partial directed coherence, spectral factorization, Granger causality, time series connectivity modelling, nonminimum phase systems

OPEN ACCESS

Edited by:
Luca Faes,
University of Palermo, Italy

Reviewed by:
Daniele Marinazzo,
Ghent University, Belgium
Yuri Antonacci,
University of Palermo, Italy

***Correspondence:**
Luiz A. Baccalá
baccala@lcs.poli.usp.br

Specialty section:
This article was submitted to Information Theory, a section of the journal Frontiers in Network Physiology

Received: *29 December 2021*
Accepted: *14 February 2022*
Published: *28 April 2022*

Citation:
Baccalá LA and Sameshima K (2022) Partial Directed Coherence and the Vector Autoregressive Modelling Myth and a Caveat. Front. Netw. Physiol. 2:845327. doi: 10.3389/fnetp.2022.845327

1 INTRODUCTION

The aim of Granger time series connectivity modelling is to examine how observations from different simultaneously observed time series may be related in the hope of exposing possible mechanisms behind their generation. This goal is intrinsically limited by a number of factors: chief among them are potential structural artifacts that result from unobserved series (confounders). This plus the fact that Granger analysis rests exclusively on observations rather than active intervention (Baccalá and Sameshima, 2014a) means that one must characterize interactions as "Granger-causal" rather than causal in the strictest sense.

In spite of this, and in connection to situations where intervention is either impossible, such as when impacting phenomena on a geophysical scale as for Solar spot/Melanoma data (Baccalá and Sameshima, 2014b) or undesirable as in physiological data analysis where noninvasive methods, at least in the human case, are always to be preferred, Granger Causality remains of interest in providing clues as to the dynamics behind the observed variables.

In recent years a vast array of methods have been developed; they originated in economics research following Granger's seminal paper (Granger, 1969) who used vector autoregression as a device to model data relationships in the time domain. His "causality" notion rests on how well the knowledge of one time series's past can enhance one's ability to predict another time series, which once vindicated, implies their connectivity. Though initially a strictly bivariate concept, the idea has been extended to the analysis of more than two simultaneously observed time series in an attempt to disentangle the effect of other series that might be acting as interaction confounders to pairwise observations (Baccalá and Sameshima, 2001a). Historically, most developments that followed rest on Geweke (1984)'s work who used Vector Autoregressive (VAR) modelling for more than two time series as a preliminary step to deduct the effect of the other observed series from the time series pair of interest. After that subtraction, the method consists of looking at a power ratio of the prediction errors between when the past of a series is taken into account against when it is not.

Much along the lines of improved estimation and inference of Granger time domain representations has been made since then and can be read in (Lütkepohl, 2005).

As a general rule, much of what followed is patterned on the representation of temporal data in terms of "output-only" systems, i.e., systems where the observed time series, $x_1(n)$, ..., $x_N(n)$, are represented as conveniently filtered versions of white noise—the so called innovations.

Because VAR models can be naturally interpreted in terms of linear filtering, already some aspects of a spectral interpretation to the Granger connectivity scenario were present in (Geweke, 1984)'s work. Further specifics have been developed since (Lütkepohl, 2005; Barrett and Seth, 2010).

The spectral nature of these problems, specially in connection to EEG data processing which are naturally characterized in terms of oscillatory behaviour, was boosted by the introduction of *Directed Transfer Function* (DTF) (Kamiński and Blinowska, 1991) and later by *partial directed coherence* (PDC) (Baccalá and Sameshima, 2001b). Both quantities employed VAR modelling for their definition. Also both have since evolved to more accurate, and thus, more appropriate measures, please see (Baccalá and Sameshima, 2021a) for their development. A *leitmotif* of those improvements was the growing realization of the importance and consequent incorporation of the estimated covariance of the innovations noise driving the observed outputs $x_i(n)$ (Baccalá et al., 2007; Takahashi et al., 2010; Baccalá and Sameshima, 2021a; Baccalá and Sameshima, 2021b).

In fact, explicit consideration of innovations covariance effects are important in connection to the so-called "instantaneous" Granger causality (iGC) and are helpful in unveiling aspects of cardio-hemodynamic behaviour (Faes, 2014). Much as in the case of GC itself, iGC was originally only seen as a time domain aspect. There have been early efforts to portray it in the frequency domain (Faes and Nollo, 2010; Faes, 2014); more general efforts have only recently appeared with Cohen et al. (2019) and Nuzzi et al. (2021) along Geweke's line of description and along PDC/DTF lines (Baccalá and Sameshima, 2021b).

All of the latter developments have relied heavily on VAR modelling. This paper, by contrast, aims to dispel the notion that PDC (Baccalá and Sameshima, 2001b) (and DTF, its dual) or any of its related quantifiers require vector autoregressive (VAR) modelling as a mandatory prerequisite. This notion coupled with limited familiarity with VAR modelling may have been a hindrance to their spread as methods of choice for Granger time series connectivity modelling among non time series specialists. We show here that absolute reliance on VAR modelling is not a must, but rather a matter of convenience, even though PDC and DTF were originally introduced with the help of VAR models.

As we have been alerted in the review process to this paper, an early precursor to the present developments is contained in (Jachan et al., 2009), which undeservedly does not seem to have attracted much following having just 22 citations at the Web of Science at the moment of this writing with only a small fraction of them reflecting actual practical method employment, mostly by its proponents. The present exposition not only confirms those results but provides evidence that they hold for more general PDC/DTF versions as well.

To dispel the VAR reliance misconception we employ a set of examples comprising a variety of methods, parametric and nonparametric, that, as we show next, yield essentially the same results. The methodological equivalence between them holds even for total PDC (tPDC) and total DTF (tDFT) as defined in (Baccalá and Sameshima, 2021b) which represent recently introduced extensions that incorporate the effects of instantaneous Granger causality to connectivity descriptions.

For brevity, we only show results for total PDC since it incorporates a consistent frequency domain description of instantaneous Granger interactions to PDC that automatically extends to total DTF's, given their duality (Baccalá and Sameshima, 2021a; Baccalá and Sameshima, 2021b).

The rest of this paper is organized as follows: **Section 2** reviews the theoretical basis and is followed in **Section 3** with a brief description of the methods employed in the comparative computations which are illustrated in **Section 4** and commented in **Section 5** leading to the conclusion in **Section 6** that tPDC/PDC (tDTF/DTF) representations are essentially canonical factors of the joint power spectral density of the data which portrays the relationship between multivariate data.

A concept that turns out to be key in the present setup is that of *spectral factorization* and the notion of a *minimum phase* spectral factor covered in more detail in **Section 2.1**.

The concept of a *minimum* versus a *nonminimum* phase system is important for our discussion. This is briefly examined in the development that follows as we show it can lead to possibly false connectivity inference when nonminimum phase mechanisms are behind the data generation process.

2 MATHEMATICAL CONSIDERATIONS

2.1 General Linear Models With Rational Spectra

A general class of linear stationary multivariate processes $\mathbf{x}(n) = [x_1(n) \dots x_N(n)]^T$ is represented (Lütkepohl, 2005) by:

$$\mathbf{x}(n) = \sum_{r=1}^{p} \mathfrak{A}_r \boldsymbol{x}(n-r) + \sum_{s=0}^{q} \mathfrak{B}_s \mathbf{w}(n-s), \tag{1}$$

where $\mathbf{w}(n) = [w_1(n) \dots w_N(n)]^T$ is a stationary (zero mean without loss of generality) multivariate innovations process with covariance matrix $\boldsymbol{\Sigma}_{\mathbf{w}}$. The process defined by (1) is termed a Vector Autoregressive Moving Average process, denoted VARMA (p, q), whose structure is defined by the $\mathfrak{A}_r, \mathfrak{B}_s$ matrices (Lütkepohl, 2005). VAR processes and vector moving average (VMA) processes are special cases, respectively when $\mathfrak{B}_s = 0, \forall s > 0$, or $\mathfrak{A}_r = 0, \forall r$. The equivalences between VAR(p) and VMA (∞), and between VMA(q) and VAR(∞) are well known, where p and q refer respectively to the AR and MA orders that make up the model.

We implicitly assume that (1) is stable, i.e., the associated $\mathbf{x}(n)$ is wide sense stationary. For simplicity we consider only the case of finite p and q. This is guaranteed if the magnitude of the roots of

$$det\ \mathfrak{A}(z) = 0 \tag{2}$$

are less than 1 for

$$\mathfrak{A}(z) = \mathbf{I} - \sum_{r=1}^{p} \mathfrak{A}_r z^{-r} \tag{3}$$

where det stands for the determinant.

Definition 1. | The system represented by (1) is minimum phase if the magnitude of the roots of

$$det\ \mathfrak{B}(z) = 0 \tag{4}$$

are less than or equal to 1 for

$$\mathfrak{B}(z) = \sum_{s=0}^{q} \mathfrak{B}_s z^{-s} \tag{5}$$

Definition 1 guarantees that stable $\mathbf{w}(n)$ innovations sequences for $n \geq 0$ may be found that lead to the observations, i.e. the system defined by (1) has a stable inverse.

Remark 1 | Strictly speaking when the roots in (5) are equal to 1, the impulse response of the inverse is merely bounded.

Remark 2 | When used as a data generating mechanism for $\mathbf{x}$(n), (1) does not need to be minimum phase. However, data modelling through (1) always leads to an estimated minimum phase counterpart system. This follows from the fact that only second order statistics are used for estimating (1) coefficients. When the data is Gaussian, this is the only available alternative, as higher order statistics are redundant and offer no additional information that might expose any evidence of possible phase nonminimality.

It is easy to show that the power spectral density matrix of $\mathbf{x}(n)$ (1) is given by:

$$\mathbf{S}_{\mathbf{x}}(\nu) = \mathfrak{A}^{-1}(\nu)\,\mathfrak{B}(\nu)\,\mathbf{\Sigma}_{\mathbf{w}}\,\mathfrak{B}^{H}(\nu)\,\mathfrak{A}^{-H}(\nu), \tag{6}$$

where

$$\mathfrak{A}(\nu) = \mathbf{I} - \sum_{r=1}^{p} \mathfrak{A}_r e^{-\mathbf{j}2\pi r\nu} \tag{7}$$

$$\mathfrak{B}(\nu) = \sum_{s=0}^{q} \mathfrak{B}_s e^{-\mathbf{j}2\pi s\nu}, \tag{8}$$

for $0 \leq |\nu| < 0.5$ which represents the normalized frequency and $\mathbf{j} = \sqrt{-1}$. Naturally (7) and (8) are associated with making $z = e^{\mathbf{j}2\pi r\nu}$ in (3) and (5) respectively.

It is easy to realize that (6) is of the form

$$\mathbf{S}_{\mathbf{x}}(\nu) = \mathfrak{H}(\nu)\,\mathbf{\Sigma}_{\mathbf{w}}\,\mathfrak{H}^{H}(\nu) \tag{9}$$

containing the frequency dependent factor, $\mathfrak{H}(\nu)$, and a frequency independent factor, $\mathbf{\Sigma}_{\mathbf{w}}$.

Remark 3. | Equations (6) and (9) hold regardless of whether (1) is minimum phase or not.

From (9) it is easy to write the coherency matrix $\mathfrak{C}(\nu)$ with entries:

$$C_{ij}(\nu) = \frac{S_{ij}(\nu)}{\sqrt{S_{ii}(\nu)\,S_{jj}(\nu)}} \tag{10}$$

by writing

$$\begin{aligned} \mathfrak{C}(\nu) &= \mathfrak{D}(\mathbf{S}_{\mathbf{x}}(\nu))^{-1/2}\,\mathbf{S}_{\mathbf{x}}(\nu)\,\mathfrak{D}(\mathbf{S}_{\mathbf{x}}(\nu))^{-1/2} \\ &= \mathfrak{D}(\mathbf{S}_{\mathbf{x}}(\nu))^{-1/2}\,\mathfrak{H}(\nu)\,\mathbf{\Sigma}_{\mathbf{w}}\,\mathfrak{H}^{H}(\nu)\,\mathfrak{D}(\mathbf{S}_{\mathbf{x}}(\nu))^{-1/2} \\ &= \mathbf{\Gamma}(\nu)\,\mathfrak{R}\,\mathbf{\Gamma}^{H}(\nu) \end{aligned} \tag{11}$$

where $\mathfrak{D}(\cdot)$ is the diag matrix operator, i.e. one that produces a matrix that is nonzero except for the diagonal elements of the operand so that

$$\mathbf{\Gamma}(\nu) = \mathfrak{D}(\mathbf{S}_{\mathbf{x}}(\nu))^{-1/2}\,\mathfrak{H}(\nu)\,\mathbf{D}^{1/2} \tag{12}$$

and

$$\mathfrak{R} = \mathbf{D}^{-1/2}\mathbf{\Sigma}_{\mathbf{w}}\mathbf{D}^{-1/2} \tag{13}$$

is a correlation matrix with ones along the main diagonal for $\mathbf{D} = \mathfrak{D}(\mathbf{\Sigma}_{\mathbf{w}})$.

Writing (11) as a product of the frequency dependent part $\mathbf{\Gamma}(\nu)$ mediated by a correlation matrix $\mathfrak{R}$ allows one to apply the definition of total DTF matrix (Baccalá and Sameshima, 2021b) as:

$$\widehat{\mathbf{\Gamma}}(\nu) = \mathbf{\Gamma}(\nu) \odot \mathbf{\Gamma}^{*}(\nu) + \mathbf{\Gamma}(\nu)\rho \odot \mathbf{\Gamma}^{*}(\nu) \tag{14}$$

where $\boldsymbol{\rho} = \mathfrak{R} - \mathbf{I}_N$, and $\mathbf{I}_N$ is an $N \times N$ identity matrix with $\odot$ standing for the Hadamard element-wise matrix product.

The entries i, j from $\widehat{\mathbf{\Gamma}}(\nu)$ reduce to the absolute square value of directed coherence from j to i, which is a scale invariant form of DTF (Baccalá et al., 1998), when instantaneous Granger causality is absent. **Eq. 14** describes what we have termed *Total Granger Influentiability* (Baccalá and Sameshima, 2021b).

An entirely parallel development allows defining total partial directed coherence (Baccalá and Sameshima, 2021b), taking advantage of the fact the partial coherence matrix can be shown to equal:

$$\begin{aligned} \mathfrak{K}(\nu) &= \mathfrak{C}^{-1}(\nu) \\ &= \mathbf{\Pi}^{H}(\nu)\,\tilde{\mathfrak{R}}\,\mathbf{\Pi}(\nu) \end{aligned} \tag{15}$$

for

$$\mathbf{\Pi}(\nu) = \mathbf{D}^{1/2}\,\mathfrak{H}^{-H}(\nu)\,\mathfrak{D}(\mathbf{S}_{\mathbf{x}}(\nu))^{1/2}\,\tilde{\mathbf{D}}^{1/2} \tag{16}$$

and

$$\tilde{\mathfrak{R}} = \tilde{\mathbf{D}}^{-1/2}\,\mathbf{\Sigma}_{\mathbf{w}}^{-1}\,\tilde{\mathbf{D}}^{-1/2} \tag{17}$$

which is a partial correlation matrix between the $\boldsymbol{w}_i(n)$ innovations where $\tilde{\mathbf{D}} = \mathfrak{D}(\mathbf{\Sigma}_{\mathbf{w}}^{-1})$.

The form in (15) is what allowed us to define total PDC as:

$$\widehat{\mathbf{\Pi}}(\nu) = \mathbf{\Pi}^{*}(\nu) \odot \mathbf{\Pi}(\nu) + \mathbf{\Pi}^{*}(\nu) \odot \tilde{\rho}\,\mathbf{\Pi}(\nu) \tag{18}$$

where $\tilde{\boldsymbol{\rho}} = \tilde{\mathfrak{R}} - \mathbf{I}_N$. The i, j entries describe what we termed the *Total Granger Connectivity* from j to i (Baccalá and Sameshima, 2021b), which reduce to generalized PDC (Baccalá et al., 2007) when instantaneous Granger causality is absent.

Whenever one can properly write the spectral density matrix as in (9), one may employ the latter quantities to describe multivariate time series within the tPDC-tDTF framework. A

case in point which we describe briefly in **Section 3.3** is provided by Wilson's spectral factorization algorithm (Wilson, 1972), which has been used before in connection with alternative Granger causality characterizations (Dhamala et al., 2008) and is also behind Jachan et al. (2009)'s results.

3 ESTIMATION METHODS

Eq. 1 was used as a general data mechanism for imposing relationships between the time series we examine in **Section 4**. The data generated were analysed via the three main approaches we briefly describe next.

3.1 Vector Autoregressive Modelling

Vector autoregressive modelling is a traditional subject (Lütkepohl, 2005). The version used here was implemented in the AsympPDC package (Sameshima and Baccalá, 2014) and employs Nuttall-Strand's method to obtain the autoregression coefficients (Marple, 1987). One important step in this sort of procedure involves finding the best model order p. Here Hannan-Quinn's method was chosen; it is a variant from the better known Akaike's method (Lütkepohl, 2005).

3.2 Vector Moving Average and Vector Autoregressive Moving Average Modelling

A traditional means of fitting VMA(q) and VARMA (p, q) models is to determine a preliminary VAR model of very large order (p = 50 was adopted here) and use its residuals $\epsilon_i(n)$ to fit the observed data $x_j(n)$ through a mock multi-input/multi-output system via least-squares. An univariate version of this approach can be appreciated in (Stoica and Moses, 2005).

In practical applications, determining p and q can be achieved through minimizing model order choice functions as in Akaike's method. Whereas, minimizing Akaike-type penalization is trivial in the VMA case, bidimensional search of tentative p and q is required in the VARMA case. To simplify matters here, we have employed the theoretical model orders used to get the estimates.

3.3 Wilson's Algorithm

Wilson's method is an iterative method that decomposes (9) into estimates for $\mathfrak{H}(\nu)$ and $\mathbf{\Sigma}_\mathbf{w}$ (Wilson, 1972). It starts by guessing a $\mathfrak{H}(\nu)$ with the restriction of its representing filters to have impulse responses that are identically zero for negative time (the so-called filter causality condition, sometimes referred as nonanticipative filters whose output cannot anticipate the input). The solution essentially amounts to Newton's root finding iterations until a maximum prescribed error is achieved. In the present case, a maximum error of 10^{-6} was adopted.

Wilson's method has been used before in connection with other Granger causality descriptions both related (Jachan et al., 2009) and directly unrelated (Dhamala et al., 2008) to PDC/DTF descriptions. It has the advantage that it can be applied to nonparametric spectral estimates, whether they are obtained by periodogram smoothing (Percival and Walden, 1993) or other means like wavelets (Lima et al., 2020).

The spectral estimates used here (henceforth referred as **WN**, nonparametric Wilson estimates) employed Welch's method as implemented in Matlab's cpsd.m function with von Hann's data window and 50% segment overlap (Percival and Walden, 1993).

The reader may obtain a working Python implementation in (Lima et al., 2020). Here a similar Matlab version was used.

3.4 Brief Comments

The time series modelling methods of **Section 3.1**, **3.2** are essentially least squares approaches. Wilson's algorithm on the other hand is a numerical square-rooting procedure that also achieves the spectral factorization of the power spectral density matrix $\mathbf{S}(\nu)$. In all cases, one obtains the so-called minimum phase spectral factor represented by $\mathfrak{H}(\nu)$ in (9).

All Matlab routines used in this paper have been included as **Supplementary Material**. For convenience, Dhamala's most recent implementation (Henderson et al., 2021) was also included and essentially leads to the same results we report next.

4 NUMERICAL ILLUSTRATIONS

In the following illustrations, the data comprise n_s = 16,384 observed points to minimize misinterpretation due to short time series effects. In all cases, the theoretical models can be used to compute the theoretical total PDC as in (Baccalá and Sameshima, 2021b). In each case, the mean-squared frequency domain approximation error of each estimation method was computed and is presented in **Table 1** after averaging over R = 100 realizations. Here Wilson estimates employed 256-point long data tappers.

Next we present three examples whose allied graphs contain the real and imaginary parts of tPDC plotted against the background of the expected theoretically computed results. These examples share the property of being generated by minimum phase (1) models.

Finally, a fourth example generated by a nonminimum phase (1) is examined. Its numerical results are contrasted to the theoretical tPDC computed with help of the actual generating model parameters.

Example 1. | Vector Moving Average Model (VMA)

We start with conceivably the simplest possible kind of vector moving average example with unidirectional influence and with the clear presence of iGC described by

$$\begin{cases} x_1(n) = w_1(n) + w_2(n-1) \\ x_2(n) = w_2(n) + w_2(n-1) \end{cases}$$

with innovations noise covariance

$$\mathbf{\Sigma}_\mathbf{w} = \begin{bmatrix} 1 & 1 \\ 1 & 5 \end{bmatrix} \tag{19}$$

whose influence of $x_2(n)$ onto $x_1(n)$ is clear due to its lagged dependence on $w_2(n)$ which is the sole input that determines $x_2(n)$. The presence of iGC is clear from (19)'s non diagonal nature.

TABLE 1 | Table containing means squared error to fits of the theoretical tPDC according to estimation method for each Example. Missing values portray when certain estimation approaches were not used.

Example	n_s	VMA	VAR	VARMA	WN
1	16,384	1.27×10^{-5}	6.84×10^{-6}		0.15×10^{-2}
	4,096	5.76×10^{-5}	3.06×10^{-5}		0.61×10^{-2}
	1,024	2.45×10^{-4}	1.57×10^{-5}		3.26×10^{-2}
2	16,384	0.21×10^{-2}	1.72×10^{-4}	2.96×10^{-8}	0.67×10^{-2}
	4,096	0.84×10^{-2}	7.09×10^{-4}	5.02×10^{-7}	2.77×10^{-2}
	1,024	3.10×10^{-2}	2.60×10^{-3}	6.65×10^{-6}	12.36×10^{-2}
3	16,384	6.11×10^{-4}	3.20×10^{-5}		0.13×10^{-2}
	4,096	0.20×10^{-2}	1.37×10^{-4}		0.57×10^{-2}
	1,024	0.90×10^{-2}	5.30×10^{-4}		3.50×10^{-2}

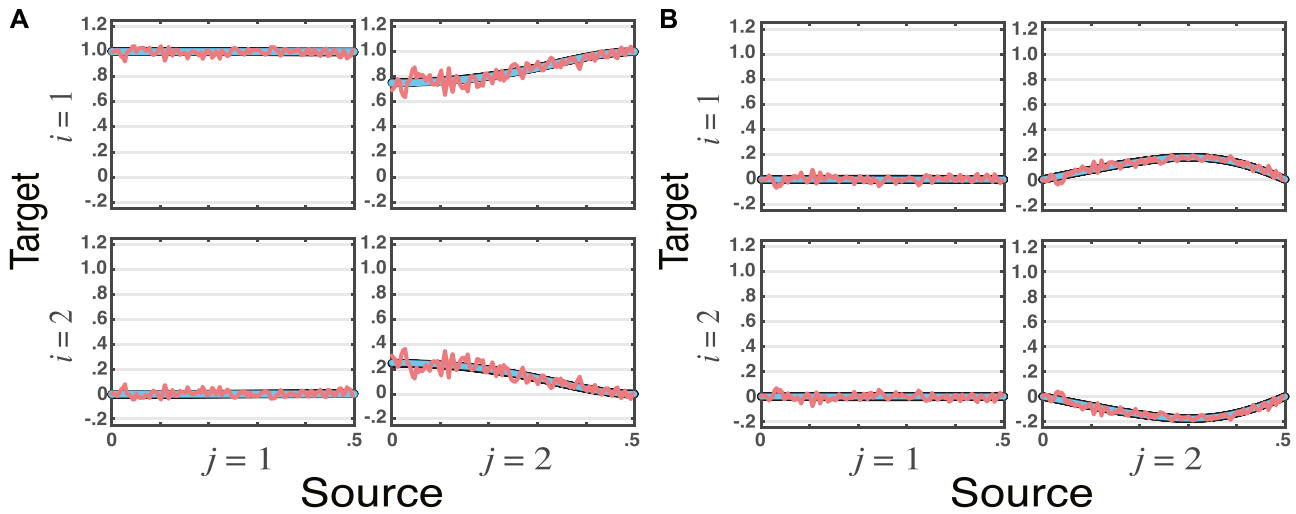


FIGURE 1 | Superimposed graphs of **total partial directed coherence, tPDC**, estimates for the VMA model simulated for n_s = 16, 384 data points (**Example 1**) and three estimation methods (VAR, VMA, WN), where the **real (A)**, and the **imaginary (B)** components are plotted separately. The theoretical tPDCs are depicted in blue lines. WN estimates ripple around theoretical values (topmost red lines), yet they closely resemble that of theoretical values. VAR and VMA estimation methods results—plotted as the two bottommost **black lines**—are visually indistinguishable from the theoretical values (blue lines).

From **Figure 1**, it is clear that for large n_s, all estimates of total PDC agree with the theoretically expected one within the constraints of estimator nature. A case in point is Wilson's factorized version computed from the nonparametric power spectral estimates which is rippled as expected (red lines), following what happens with the original spectral estimates.

Example 2. | Vector Autoregressive Moving Average Model (VARMA)

The next example is a bit more elaborate. It has a VARMA (2, 2) data generating procedure described by

$$\begin{cases} x_1(n) = 2\,r\,\cos(\theta)x_1(n-1) - r^2 x_1(n-2) + w_1(n) + w_3(n) + w_3(n-1) \\ x_2(n) = b\,x_1(n-1) + a\,x_2(n-1) + w_2(n) \\ x_3(n) = c\,x_3(n-1) + w_2(n) + w_2(n-2) + w_3(n) \end{cases}$$

where $r = 0.95$, $\theta = \pi/3$, $b = 0.5$, $a = -0.5$, $c = 0.7$ and $\mathbf{\Sigma}_\mathbf{w}$ equal to the identity matrix.

As in the previous example, total PDC estimates match one another regardless of method, see **Figure 2**.

Albeit at little surprise, it is important to realize that the use of the VARMA modelling scheme (**Section 3.2**) yields substantially better fit. This is confirmed by **Table 1** results.

Example 3. | Vector Autoregressive Model (VAR)

The third toy example covers the one used in (Baccalá and Sameshima, 2021b) and was borrowed from (Faes, 2014) involving three channels whose connectivity is assessed *via* a VAR model taking iGC effects into account through tPDC. One obtains essentially the same results irrespective of the computational approach, see **Figures 3A, B**.

Example 4. | Nonminimum Phase Data

Consider a moving average data generation scheme using (1) with

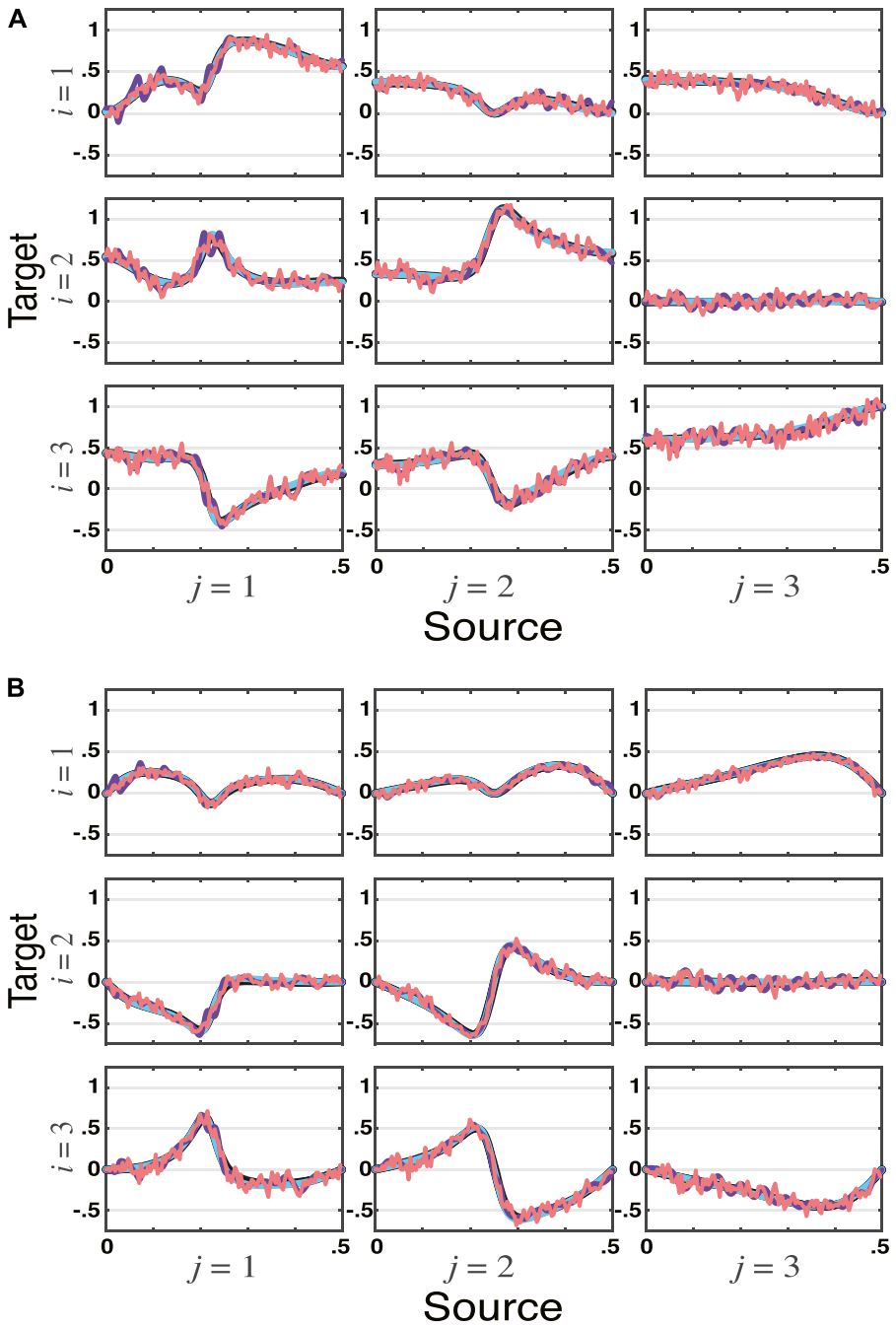


FIGURE 2 | tPDC estimates by all four methods—VAR, VMA, VARMA, and WN—for the VARMA model in **Example 2** simulated for n_s = 16, 384 data points are depicted, with **real (A)**, and **imaginary (B)** components plotted separately. As before, the theoretical tPDCs are also shown (blue lines). Again note that WN estimates (topmost red lines) ripple around theoretical values. In this case, VMA estimates (purple lines) also ripple around theoretical values (blue lines) illustrating estimator accuracy limitations. This is also apparent on **Table 1**. VAR and VARMA results—plotted as the two black bottommost lines just underneath the theoretical values—represent much closer approximations.

$$\mathfrak{B}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \mathfrak{B}_1 = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}, \ \mathfrak{B}_2 = \begin{bmatrix} 4 & 2 \\ 0 & 2 \end{bmatrix}, \qquad (20)$$

whose allied (4) roots $\{-1 + \pm \mathbf{j}\sqrt{3}, \pm \mathbf{j}\sqrt{2}\}$ have magnitudes that are larger than 1, making this a nonminimum phase data generating mechanism as opposed to all previous examples, as computing their (4) easily shows. It is clear from (20) that $x_2(n)$

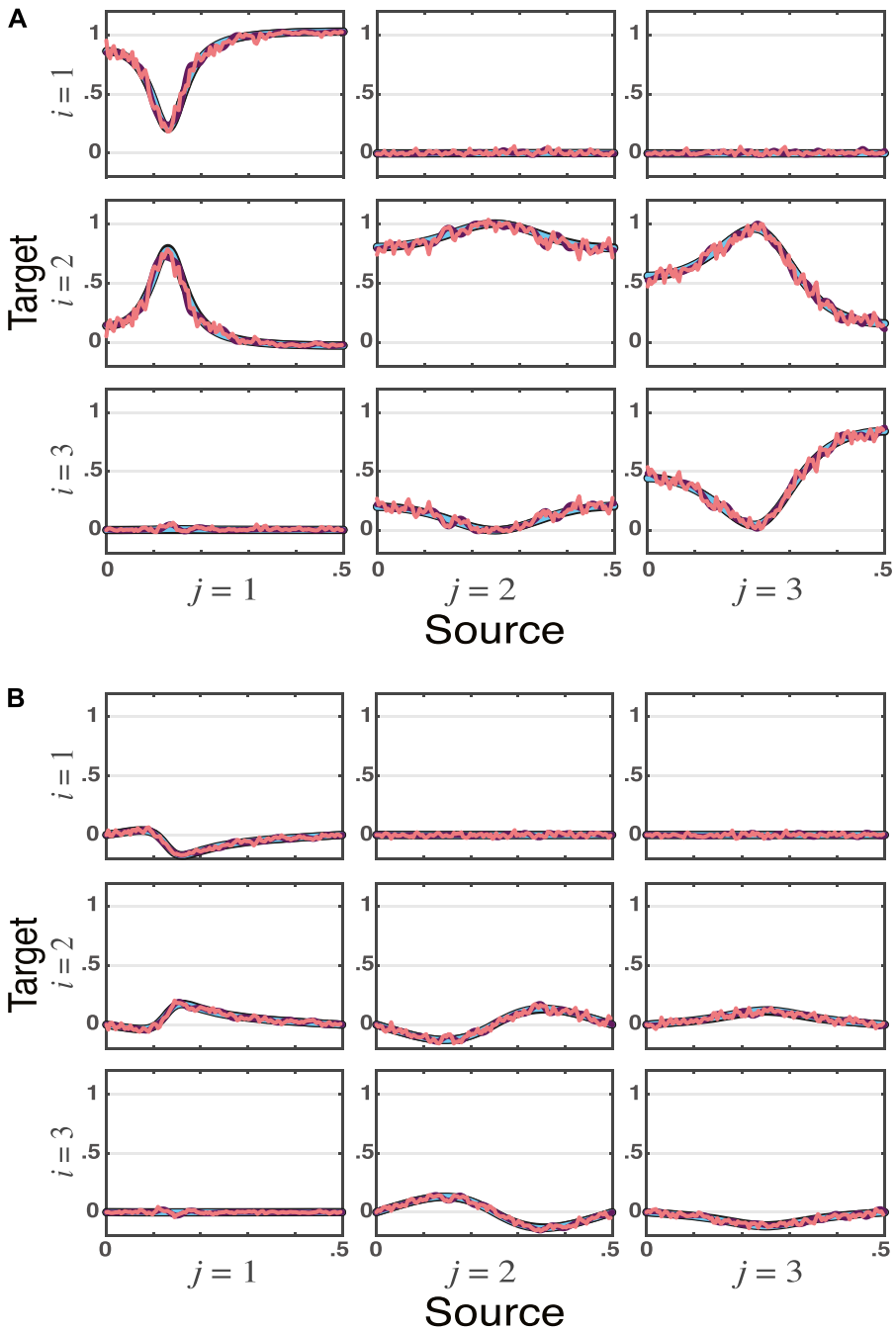


FIGURE 3 | tPDC estimates for **Example 3** are shown for VAR, VMA, and WN methods (n_s = 16, 384) with **real (A)**, and **imaginary (B)** components plotted separately. As before, theoretical tPDCs are also shown (blue lines). Once again, WN estimates (topmost red lines) ripple around theoretical values. Here so do too VMA estimates (purple lines) signalling their poor expected accuracy when fitting VAR data. This is confirmed by results presented on **Table 1**. VAR results are plotted as the two bottommost black lines underneath the theoretical values.

Granger-causes $x_1(n)$ but not otherwise. This is reflected in the computed tPDCT blue lines of **Figure 4**. Here, (19) was adopted as the innovations covariance matrix.

Use of **Section 3** algorithms leads to the results of **Figure 4** where the estimation methods agree among themselves, but are markedly different from the tPDC computed using (20).

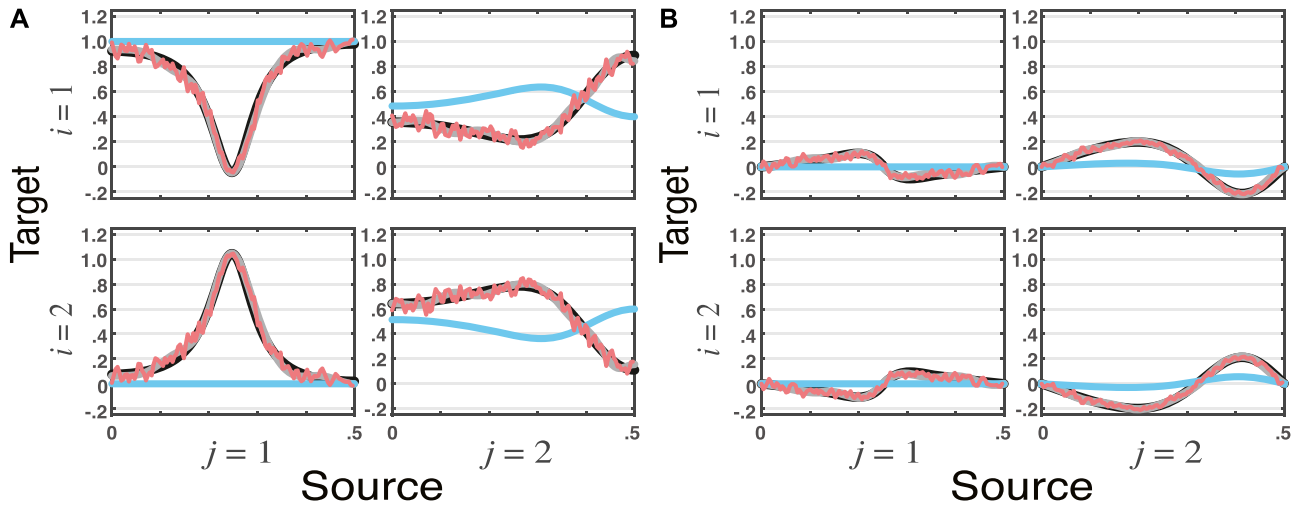


FIGURE 4 | tPDC estimates for VARMA model with nonminium phase data in **Example 4** using the VAR, VMA, and WN methods (n_s = 16, 384) portraying its **(A) real** and **(B) imaginary** components. As before theoretical tPDCs are shown as blue lines. Here, WN estimates (topmost red lines) also ripple and agree with VMA (**black lines**) and VAR estimates (gray lines) are very close to one another but differ significantly from the theoretical values (blue lines). Note the parameters in (20) imply no connection from $x_1(n)$ onto $x_2(n)$, yet all three estimation methods wrongly indicate a non zero **tPDC** real component reflecting strong estimated GC.

The reader may easily verify using the **Supplementary Material** that the estimated solution using VMA modelling leads to (4) roots whose magnitudes are all smaller than 1.

Most importantly, however is that this example shows that GC causal relationships imposed through nonminimum phase systems can be wrongly inferred. The consequences of this are further elaborated in the discussion.

5 DISCUSSION

It is perhaps surprising that PDC/DTF have so long, and unnecessarily so, remained inextricably associated with VAR modelling even in view of early evidence to the contrary (Jachan et al., 2009). Partial explanation may lie in the early virtual exclusive reliance on VAR modelling that also dominated initial approaches to Granger Causality characterization (Granger, 1969; Geweke, 1984). This scenario in connection to time series modelling in the time domain slowly changed as VMA and VARMA approaches have been shown viable and possibly desirable depending on the nature of the data under study (Boudjellaba et al., 1992; Boudjellaba et al., 1994). The latter methods are attractive because they more parsimoniously fit the underlying data as in **Example 2** *via* fewer parameters. This reflects Parzen's Parsimony Principle which formalizes the statistical advantage of describing data via the least possible number of parameters (Yaffee and McGee, 2000) that in the present case leads to lower average estimation error (see **Table 1**). More details on alternative time domain characterization can be appreciated in Lütkepohl (2005).

Because of its prediction improvement ethos, Granger Causality, when originally defined, rested on VAR modelling's predictive ability (Granger, 1969). Moreover, at that time it was the only practical alternative from a computational perspective. It is thus unsurprising that other predictive methods like VMA and VARMA modelling also can fit the purpose.

Given PDC/tPDC's frequency domain ties with Granger causality (Baccalá and Sameshima, 2021a) (with the inclusion of full instantaneous effects) (Baccalá and Sameshima, 2021b), it is therefore no wonder that they too can be carried out *via* other methods like VMA or VARMA modelling.

Thus we have shown that PDC/DTF (total or otherwise) are **not** irrevocably tied to VAR data modelling, though today, VAR remains the best studied and most widely applied option. It has the advantage of having rigorous asymptotic results in the squared PDC/DTF case (Baccalá et al., 2013; Baccalá et al., 2016). Work is in progress to provide the asymptotics to the allied total PDC/DTF quantities.

Further research is needed to pinpoint which of the latter methods is best for what purpose. It is comforting to know that many methods provide equivalent descriptions if used properly.

For example, even though it is possible to combine the response of different trials in event-related experiments while employing VAR models (Rodrigues and Baccalá, 2015), this feat may also, and perhaps more easily in some cases, be achieved through the application of Wilson's method to estimate nonparametric spectra and cross-spectra averaged over trials. Other methods have been proposed to deal with spectral matrix factorization that still need proper practical appraisal (Amblard, 2015).

Though Wilson-type spectral factorization methods seem less effective in practice, it does not mean that they should be discarded. Here we only used Welch's spectral estimator. More research is needed, by employing other spectral estimation procedures like multitappering for instance (Percival and Walden, 1993) that could improve accuracy as they may more appropriately fit certain spectral shapes.

Here we have employed large data sets, but one should expect substantial performance differences for shorter time series. In this case, too, as hinted by **Table 1** results, VAR methods remain quite efficient, except when better approximation can be made through models that portray the data more closely as in the VARMA **Example 2**.

Other approaches have been proposed to obtain Granger-type estimates, namely state space modelling is one such example (Barnett and Seth, 2015); present research is on-going to evaluate them. In fact, as Sayed and Kailath (2001)'s theoretical appraisal of univariate spectral density factorization methods suggests, even state-space models can be seen as spectral factorization providers.

All the above methods, by providing minimum phase spectral factors to the spectral density matrix, ideally portray **identical** Granger relationship representations within the accuracy and characteristic limitations of the employed spectral estimation/factorization techniques.

At this point, before we examine the nonminimum phase data generation issue, and even if the theoretical realization that GC connectivity reduces to a spectral factorization problem were not important, the practice oriented reader might be wondering why so much ado about a VAR 'myth' if in the end VAR remains a reasonable practical compromise? To answer this, please have in mind that spectral fitting is a method of approximation of whatever the real spectra are. According to Parzen's principle the best conceivable statistical inference reliability rests on having the least number of descriptive parameters for a given approximation error (which one can gauge by the residual covariance matrix). Hence, though at present VMA and VARMA methods are not as mature as VAR methods in so far as inference is concerned, they hold the promise of potential higher inference accuracy in appropriate cases as they get to be further developed.

Another question that may be bothering those who are practice oriented is: why use nonparametric methods if their performance is not so good and if they call for much longer data sets to furnish the same level of accuracy? In fact, this may well behind their infrequent use in the past. First remember the issue of ease of use as in the analysis of event related cases we mentioned. Remember too that many investigators remain uneasy about parametric methods because they require model order decisions added to the often glossed over problem of model diagnostic checking (Li, 2003). Despite their wiggly nature, Wilson nonparametric methods dispense with these decisions and can be helpful in providing hints to the approximative quality of contending parametric models. They have issues of their own that also merit further examination. These problems lie in nonparametric spectral estimation shortcomings (Percival and Walden, 1993) that many applied users often overlook.

In short, having more options in one's analysis toolkit is beneficial and should not be discarded.

Now moving on, there is the important caveat we have shown: due to their intrinsic minimum phase limitation, the methods we explored here are unable to properly capture GC-type relations when the underlying data generation mechanism is nonminimum phase as in **Example 4**. This happens because these methods, either through classical time series modelling or direct spectral factorization, employ only second order statistics.

Though we do not show this explicitly here, Geweke-based approaches also suffer from the same limitations. This is easy to realize if one takes into account that they lead to conclusions that are similar to those reached *via* PDC/DTF-type approaches.

This scenario evokes two intertwined questions: 1) whether dynamical (viz. physical, physiological or economical) observations of phenomena actually conform to nonminimum phase generation mechanisms that might obscure their connectivity inference and 2) whether real data using GC methods in the past actually hold in view of this observation.

As an example consider a situation when nonminimum phase signals are a practical reality. It happens in wireless communication, and is due to signal propagation through dispersive multipathway media that leads to serious bit-error rate impairment. As a man made system, this problem is circumvented by the transmission of pre-arranged pseudo random data (training) sequences the receiver uses to estimate channel nonminimality. Use of these sequences maps the receiver "output-only" problem into an equivalent "input-output" problem that can reveal nonminimum phase effects through second order statistics alone. This solution is sometimes unsatisfactory as it imposes a penalty on the transmission rate of useful data. During the 1990's a considerable body of literature appeared to address this problem by dispensing with training sequences and using the received (output) data only (Haykin, 1994). This is possible when the data is nongaussian, i.e., there is information beyond the ordinary second order statistics of the spectrum, something that can be made by design in telecommunication systems. Signal diversity in both time and space, via telecom signal characteristics or through employment of redundant receiver antennas is also an option. This general field has been known as that of "*blind*" identification/equalization (see Chi et al., 2006, for an overview). Whereas real data properties cannot be 'designed' as in man made systems, they are often nongaussian and this could in principle be exploited to overcome the nonminimum phase generation limitation on GC inference we described here.

The answer to 2) must thus await further analysis in what is a matter for further exciting research that may entail the revision of many conclusions regarding formerly analysed real data.

6 CONCLUSION

The first take home lesson is that PDC/DTF-type estimators of Granger connectivity/influentiability (Baccalá and Sameshima, 2014a) even in their latest and most general total form (tPDC/tDTF), incorporating instantaneous Granger effects, do not require vector autoregressive modelling as a mandatory step but can be obtained through any other means of spectral factorization of the spectral density matrix into minimum phase factors. The second lesson is that, though not mandatory, VAR modelling, since it can be used to obtain consistent spectral factors, and because of its practicality and efficiency, remains the method of choice, specially for short data

sets. The third no less important lesson is that care as to conclusions about real data must be exercised as possible unknown nonminimum phase data generating mechanisms may be at play that can confound results as to the actual true underlying connectivity when methods of the present spectral factorization class are used.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

Both authors have equally shared in the conception, writing, and editing the paper.

FUNDING

LB was funded by CNPq, Grant number 308073/2017-7. LB and KS were partially supported by FAPESP Grant 2017/12943-8.

ACKNOWLEDGMENTS

KS is affiliated with LIM 43–HCFMUSP. Both are attached to the Center for Interdisciplinary Research on Applied Neurosciences (NAPNA), Universidade de São Paulo, São Paulo, Brazil.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnetp.2022.845327/full#supplementary-material

REFERENCES

Amblard, P.-O. (2015). A Nonparametric Efficient Evaluation of Partial Directed Coherence. *Biol. Cybern.* 109, 203–214. doi:10.1007/s00422-014-0636-0

Baccalá, L. A., and Sameshima, K. (2014a). in *Causality and Influentiability: The Need for Distinct Neural Connectivity Concepts*. Editors D Ślęzak, AH Tan, JF Peters, and L Schwabe (Cham: Springer International Publishing), 424–435. Brain Informatics and Health. doi:10.1007/978-3-319-09891-3_39

Baccalá, L. A., and Sameshima, K. (2001a). Overcoming the Limitations of Correlation Analysis for many Simultaneously Processed Neural Structures. *Prog. Brain Res.* 130, 33–47. doi:10.1016/s0079-6123(01)30004-3

Baccalá, L. A., Takahashi, D. Y., and Sameshima, K. (2007). "Generalized Partial Directed Coherence," in 2007 15th International Conference on Digital Signal Processing, Cardiff, UK, 1-4 July 2007, 163–166. doi:10.1109/ICDSP.2007.4288544

Baccalá, L. A., De Brito, C. S. N., Takahashi, D. Y., and Sameshima, K. (2013). Unified Asymptotic Theory for All Partial Directed Coherence Forms. *Phil. Trans. R. Soc. A.* 371, 20120158. doi:10.1098/rsta.2012.0158

Baccalá, L. A., Sameshima, K., Ballester, G., Do Valle, A. C., and Timo-Iaria, C. (1998). Studying the Interaction between Brain Structures via Directed Coherence and Granger Causality. *Appl. Sig. Process.* 5, 40–48.

Baccalá, L. A., and Sameshima, K. (2021b). Frequency Domain Repercussions of Instantaneous Granger Causality. *Entropy* 23, 1037. doi:10.3390/e23081037

Baccalá, L. A., and Sameshima, K. (2001b). Partial Directed Coherence: a New Concept in Neural Structure Determination. *Biol. Cybern.* 84, 463–474. doi:10.1007/PL00007990

Baccalá, 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

Baccala, L. A., Takahashi, D. Y., and Sameshima, K. (2016). Directed Transfer Function: Unified Asymptotic Theory and Some of its Implications. *IEEE Trans. Biomed. Eng.* 63, 2450–2460. doi:10.1109/TBME.2016.2550199

Baccalá, L., and Sameshima, K. (2014b). "Partial Directed Coherence," in *Methods in Brain Connectivity Inference through Multivariate Time Series Analysis*. Editors K Sameshima and LA Baccalá (Boca Raton: CRC), 57–73. Frontiers in Neuroengineering Series. doi:10.1201/b16550-11

Barnett, L., and Seth, A. K. (2015). Granger Causality for State-Space Models. *Phys. Rev. E Stat. Nonlin Soft Matter Phys.* 91, 040101. doi:10.1103/PhysRevE.91.040101

Barrett, A. B., Barnett, L., and Seth, A. K. (2010). Multivariate Granger Causality and Generalized Variance. *Phys. Rev. E Stat. Nonlin Soft Matter Phys.* 81, 041907–041914. doi:10.1103/PhysRevE.81.041907

Boudjellaba, H., Dufour, J.-M., and Roy, R. (1994). Simplified Conditions for Noncausality between Vectors in Multivariate ARMA Models. *J. Econom.* 63, 271–287. doi:10.1016/0304-4076(93)01568-7

Boudjellaba, H., Dufour, J.-M., and Roy, R. (1992). Testing Causality between Two Vectors in Multivariate Autoregressive Moving Average Models. *J. Am. Stat. Assoc.* 87, 1082–1090. doi:10.1080/01621459.1992.10476263

Chi, C. Y., Feng, C. C., and Chen, C. H. (2006). *Blind Equalization and System Identification: Batch Processing Algorithms, Performance and Applications*. London: Springer.

Cohen, D., Sasai, S., Tsuchiya, N., and Oizumi, M. (2020). A General Spectral Decomposition of Causal Influences Applied to Integrated Information. *J. Neurosci. Methods* 330, 108443. doi:10.1016/j.jneumeth.2019.108443

Dhamala, M., Rangarajan, G., and Ding, M. (2008). Estimating Granger Causality from Fourier and Wavelet Transforms of Time Series Data. *Phys. Rev. Lett.* 100, 018701. doi:10.1103/PhysRevLett.100.018701

Faes, L. (2014). "Assessing Connectivity in the Presence of Instantaneous Causality," in *Methods in Brain Connectivity Inference through Multivariate Time Series Analysis*. Editors K Sameshima and LA Baccalá (Boca Raton: CRC Press), 87–112. doi:10.1201/b16550-13

Faes, L., and Nollo, G. (2010). Extended Causal Modeling to Assess Partial Directed Coherence in Multiple Time Series with Significant Instantaneous Interactions. *Biol. Cybern.* 103, 387–400. doi:10.1007/s00422-010-0406-6

Geweke, J. F. (1984). Measures of Conditional Linear Dependence and Feedback between Time Series. *J. Am. Stat. Assoc.* 79, 907–915. doi:10.1080/01621459.1984.10477110

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* 37, 424–438. doi:10.2307/1912791

S. Haykin (Editor) (1994). *Blind Deconvolution* (New Jersey: PTR Prentice Hall).

Henderson, J. A., Dhamala, M., and Robinson, P. A. (2021). Brain Dynamics and Structure-Function Relationships via Spectral Factorization and the Transfer Function. *NeuroImage* 235, 117989. doi:10.1016/j.neuroimage.2021.117989

Jachan, M., Henschel, K., Nawrath, J., Schad, A., Timmer, J., and Schelter, B. (2009). Inferring Direct Directed-Information Flow from Multivariate Nonlinear Time Series. *Phys. Rev. E* 80, 011138. doi:10.1103/PhysRevE.80.011138

Kaminski, M. J., and Blinowska, K. J. (1991). A New Method of the Description of the Information Flow in the Brain Structures. *Biol. Cybern.* 65, 203–210. doi:10.1007/BF00198091

Li, W. K. (2003). *Diagnostic Checks in Time Series*. Boca Raton: Taylor & Francis.

Lima, V., Dellajustina, F. J., Shimoura, R. O., Girardi-Schappo, M., Kamiji, N. L., Pena, R. F. O., et al. (2020). Granger Causality in the Frequency Domain: Derivation and Applications. *Rev. Bras. Ensino Fís.* 42, e20200007. doi:10.1590/1806-9126-rbef-2020-0007

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. New York: Springer.

Marple, S. L., Jr (1987). *Digital Spectral Analysis with Applications*. New Jersey: Prentice-Hall.

Nuzzi, D., Stramaglia, S., Javorka, M., Marinazzo, D., Porta, A., and Faes, L. (2021). Extending the Spectral Decomposition of Granger Causality to Include Instantaneous Influences: Application to the Control Mechanisms of Heart Rate Variability. *Phil. Trans. R. Soc. A.* 379, 0263. doi:10.1098/rsta.2020.0263

Percival, D. B., and Walden, A. T. (1993). *Spectral Analysis for Physical Applications*. New York: Cambridge University Press.

Rodrigues, P. L., and Baccalá, L. A. (2015). "A New Algorithm for Neural Connectivity Estimation of EEG Event Related Potentials," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (IEEE), 3787–3790. doi:10.1109/embc.2015.7319218

[Dataset] Sameshima, K., and Baccalá, L. A. (2014). Asymp_PDC Package. Available at: https://www.lcs.poli.usp.br/~baccala/pdc/CRCBrainConnectivity/AsympPDC/index.html (Accessed 12 19, 2021).

Sayed, A. H., and Kailath, T. (2001). A Survey of Spectral Factorization Methods. *Numer. Linear Algebra Appl.* 8, 467–496. doi:10.1002/nla.250

Stoica, P., and Moses, R. L. (2005). *Spectral Analysis of Signals*. Upper Saddle River: Pearson/Prentice Hall.

Takahashi, D. Y., Baccalá, L. A., and Sameshima, K. (2010). Information Theoretic Interpretation of Frequency Domain Connectivity Measures. *Biol. Cybern.* 103, 463–469. doi:10.1007/s00422-010-0410-x

Wilson, G. T. (1972). The Factorization of Matricial Spectral Densities. *SIAM J. Appl. Math.* 23, 420–426. doi:10.1137/0123044

Yaffee, R. A., and McGee, M. (2000). *An Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS*. 1 edn. New York: Academic Press.

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.

frontiers | Frontiers in Network Physiology

ORIGINAL RESEARCH
published: 03 May 2022
doi: 10.3389/fnetp.2022.893743

The Reconstruction of Causal Networks in Physiology

Moritz Günther[1]*, *Jan W. Kantelhardt*[2] and *Ronny P. Bartsch*[3]

[1]Max Planck Institute for Meteorology, Hamburg, Germany, [2]Institute of Physics, Martin-Luther-University Halle-Wittenberg, Halle, Germany, [3]Department of Physics, Bar-Ilan University, Ramat Gan, Israel

We systematically compare strengths and weaknesses of two methods that can be used to quantify causal links between time series: Granger-causality and Bivariate Phase Rectified Signal Averaging (BPRSA). While a statistical test method for Granger-causality has already been established, we show that BPRSA causality can also be probed with existing statistical tests. Our results indicate that more data or stronger interactions are required for the BPRSA method than for the Granger-causality method to detect an existing link. Furthermore, the Granger-causality method can distinguish direct causal links from indirect links as well as links that arise from a common source, while BPRSA cannot. However, in contrast to Granger-causality, BPRSA is suited for the analysis of non-stationary data. We demonstrate the practicability of the Granger-causality method by applying it to polysomnography data from sleep laboratories. An algorithm is presented, which addresses the stationarity condition of Granger-causality by splitting non-stationary data into shorter segments until they pass a stationarity test. We reconstruct causal networks of heart rate, breathing rate, and EEG amplitude from young healthy subjects, elderly healthy subjects, and subjects with obstructive sleep apnea, a condition that leads to disruption of normal respiration during sleep. These networks exhibit differences not only between different sleep stages, but also between young and elderly healthy subjects on the one hand and subjects with sleep apnea on the other hand. Among these differences are 1) weaker interactions in all groups between heart rate, breathing rate and EEG amplitude during deep sleep, compared to light and REM sleep, 2) a stronger causal link from heart rate to breathing rate but disturbances in respiratory sinus arrhythmia (breathing to heart rate coupling) in subjects with sleep apnea, 3) a stronger causal link from EEG amplitude to breathing rate during REM sleep in subjects with sleep apnea. The Granger-causality method, although initially developed for econometric purposes, can provide a quantitative, testable measure for causality in physiological networks.

Keywords: time series analysis, network physiology, Granger causality, bivariate phase rectified signal averaging, sleep apnea, heartbeat, respiration, brain-wave amplitudes

OPEN ACCESS

Edited by:
Sonia Charleston-Villalobos, Metropolitan Autonomous University, Mexico

Reviewed by:
Alexander Caicedo, Universidad del Rosario, Colombia
Andreas Voss, Technische Universität Ilmenau, Germany

***Correspondence:**
Moritz Günther
moritz.guenther@mpimet.mpg.de

Specialty section:
This article was submitted to Systems Interactions and Organ Networks, a section of the journal Frontiers in Network Physiology

Received: *10 March 2022*
Accepted: *06 April 2022*
Published: *03 May 2022*

Citation:
Günther M, Kantelhardt JW and Bartsch RP (2022) The Reconstruction of Causal Networks in Physiology. Front. Netw. Physiol. 2:893743. doi: 10.3389/fnetp.2022.893743

1 INTRODUCTION

Causality is an ambiguous term and there are numerous philosophical, sociological, statistical, physical and information-theoretic approaches to define causality [Granger (1980); Hlaváčková-Schindler et al. (2007); Pearl and Mackenzie (2018)]. Although classified as statistical rather than a causal concept by some authors [see e.g., Hamilton (1994); Pearl (2009) for well-founded arguments], Granger causality ["*G*-causality", Granger (1969)] provides a generally accepted operational framework to investigate causal interactions in time series. Going back to an idea by

Norbert Wiener [Wiener (1956); therefore also "Wiener-Granger Causality"], Clive Granger was first to apply a linear regression model to probe whether a process X has a causal relationship with another process Y {or whether X can forecast Y [Hamilton (1994)]}. Limitations of this linear approach and prominent non-linear extensions of G-causality are discussed in detail by Hlaváčková-Schindler et al. (2007).

In the field of physiological time series analysis and in particular when probing for physiological interactions, G-causality plays a major role along with entropy-based measures [Schulz et al. (2019)], phase synchronization analysis and symbolic dynamics [Müller et al. (2016)]. Indeed, G-causality is frequently used in the emerging field of Network Physiology [Bashan et al. (2012)] to investigate the network interactions between multiple physiological systems involved in cardiovascular/cardiorespiratory control [Porta and Faes (2015); Schulz et al. (2013)] and heart-brain coupling [Faes et al. (2015)]. It is becoming a standard tool in neuroscience to identify directed functional interactions in the brain [Hesse et al. (2003); Bressler and Seth (2011); Seth et al. (2015)]. However, G-causality was initially developed for economic time series [Granger (2004)], which are usually shorter (regarding their number of samples) and sampled at lower frequencies than physiological time series. This must be considered when using G-causality for physiological applications. Notably, the condition of "instantaneous causality" [Granger (1969)] may, in contrast to economic data, not be present in physiological data because typical sampling rates are higher than the delay time of the causal relationships [Lin et al. (2016)]. Furthermore, there are forms of coupling between physiological systems that coexist but operate at different time scales [Bartsch et al. (2014); Bartsch and Ivanov (2014)]. Therefore, in order to identify physiological interactions for fast as well as very slow processes, the original model must be extended and G-causality computed at different temporal resolutions.

An often discussed drawback of G-causality for practical implementations is the necessity of data being stationary, which is usually not the case for physiological recordings. Workarounds range from simply differentiating the data or analyzing shorter ("quasi-stationary") time windows to more complex methods utilizing an adaptive recursive least-square algorithm [Hesse et al. (2003)] or applying spectral density matrix factorization of the Fourier and wavelet transforms [Dhamala et al. (2008)] — with each method having its own pros and cons [Bressler and Seth (2011)].

An alternative, simple yet powerful method, which does not require stationarity to investigate interactions and causal relations between time series, is Bivariate Phase Rectified Signal Averaging (BPRSA) analysis [Schumann et al. (2008)]. While originally developed as a mono-variate method to study quasi-periodic oscillations in non-stationary signals [Bauer et al. (2006b)] and quantify cardiovascular risk [Bauer et al. (2006a)], its bivariate extension has been applied to assess spontaneous baroreflex sensitivity [Müller et al. (2012)], and, more recently, to analyze maternal-fetal heart rate coupling [Montero-Nava et al. (2020)].

Because BPRSA is, in contrast to G-causality, a model-free approach to study inter-relationships and causality in physiological signals, in this paper we aimed for a systematic comparison of both methods. This will be done in Part A after a proper introduction of each method. In particular, we will elaborate on their strengths and weaknesses, and present statistical tests to probe for significant interactions. The analysis is done with a focus on possible applications in physiology. A corresponding example regarding physiological networks during sleep will be presented in Part B of the paper.

2 PART A: METHODS FOR CAUSALITY ANALYSIS

2.1 Differentiating Between Direct and Indirect Links

In physiological networks an important problem is to distinguish direct from indirect links. **Figure 1** depicts direct and indirect links in simple three-node networks. In all three cases, the node corresponding to the source signal z seems to influence the target node, i.e., signal x. While there is a direct link from z to x in subfigure (a), there is only an indirect link in subfigure (b), mediated by signal y. Subfigure (c) shows another form of an indirect link, where the link between z and x is purely due to the common influence of y on both signals. Note that a time lag, indicated by the operator $\mathcal{L}$, is important only in case (c). If the time lag from y to z was longer than the lag from y to x (i.e., $\alpha > \beta$), the indirect link would change its direction and point from x to z instead.

There are more complex and mixed cases e.g., direct and indirect links between the same two nodes, and coexisting links [Bartsch and Ivanov (2014)], but the ones shown in **Figure 1** are the most basic setups, and in the following they will be used to test and compare G-causality and BPRSA. By studying results for modeled data, we show that G-causality is more appropriate to distinguish these setups within a certain range of detection limits.

2.2 Method 1: Granger Causality

A simplified definition of Granger causality is: "Variable z Granger-causes (G-causes) variable x if knowledge about z improves the forecast of x" [Granger (1969)]. This reflects our common understanding of cause and consequence: The cause must precede the consequence in time and if z has no effect on x, we do not call it causal. This idea is formalized in the framework of autoregressive (AR) processes.

Under fairly general conditions a random process can be described by an AR model of order p [see, e.g., Lütkepohl (2005)]. Consider two AR models of the time series x_t, one including and one excluding information on z_t,

$$x_t = \sum_{i=1}^{p} \left(\phi_i^{(1)} x_{t-i} \right) + w_t^{(1)}; \qquad \mathrm{STD}\left(w^{(1)}\right) = \sigma^{(1)}, \quad (1)$$

$$x_t = \sum_{i=1}^{p} \left(\phi_i^{(2)} x_{t-i} + \psi_i^{(2)} z_{t-i} \right) + w_t^{(2)}; \qquad \mathrm{STD}\left(w^{(2)}\right) = \sigma^{(2)}. \quad (2)$$

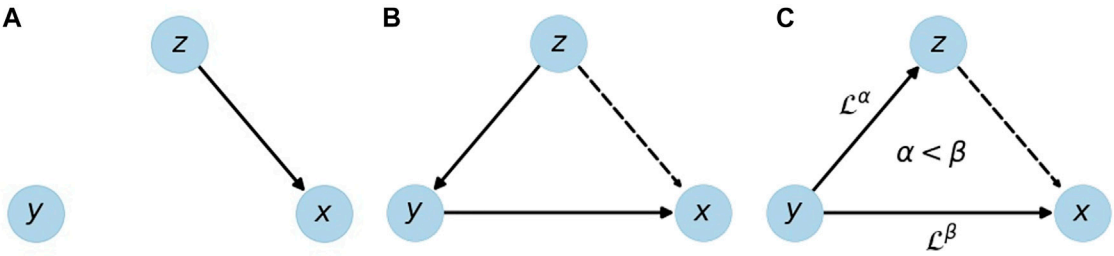


FIGURE 1 | Illustration of a direct link **(A)** and indirect links **(B,C)** from source z to target x. An indirect link can be due to a causal chain $z \to y \to x$ **(B)** or due to time delayed effects of y on x and z with a longer time delay towards x. Solid edges represent direct links, dashed edges the resulting indirect links; $\mathcal{L}$ is the time lag operator.

Equation 1 models the value of x at time t as a weighted average over its own past plus a white noise process w with zero mean and standard deviation (STD) $\sigma^{(1)}$. The weighting factors ϕ_i can be obtained by minimizing the error term $w^{(1)}$. The past of x carries information about its own future, but this information is necessarily incomplete due to the statistical nature of x. The better x can be described (forecast) from its own past, the lower the standard deviation $\sigma^{(1)}$ of the residual $w^{(1)}$.

Equation 2 additionally considers information about z. If this information helps to model (forecast) x, then the standard deviation $\sigma^{(2)}$ will be reduced compared to $\sigma^{(1)}$. The G-value $G_{z\to x}$ is a measure for the improvement of the forecast of x by including z, and therefore a measure of causality,

$$G_{z\to x} = \ln \frac{\sigma^{(1)}}{\sigma^{(2)}}. \tag{3}$$

$G_{z\to x}$ quantifies G-causality, however it remains unclear whether 1) an obtained G-value is significantly different from zero, and 2) the link is direct or indirect.

To resolve issue 1), one can test the null hypothesis that all coefficients $\psi_i^{(2)}$ of **Eq. 2** are practically zero, i. e., G-causality does not exist. This is done by assuming an F or chi-squared distribution, given that the estimators of the least square method coefficients are asymptotically normally distributed [Neusser (2011)], and estimating the probability of a non-zero mean value. Because no other variables (such as y) are taken into account, this is a *pairwise analysis* method, which can identify links from z to x in all three cases of **Figure 1**.

In order to differentiate between direct and indirect links, i.e., to resolve issue 2), a *conditional analysis* is necessary. Consider the two extended AR models,

$$x_t = \sum_{i=1}^{p} \left(\phi_i^{(3)} x_{t-i} + \tau_i^{(3)} y_{t-i} \right) + w_t^{(3)} \qquad \mathrm{STD}\left(w^{(3)}\right) = \sigma^{(3)}, \tag{4}$$

$$x_t = \sum_{i=1}^{p} \left(\phi_i^{(4)} x_{t-i} + \tau_i^{(4)} y_{t-i} + \psi_i^{(4)} z_{t-i} \right) + w_t^{(4)} \quad \mathrm{STD}\left(w^{(4)}\right) = \sigma^{(4)}. \tag{5}$$

Equations 4, **5** are the same as **Eqs. 1**, **2**, except that the past of y is added to both. Hence, $\sigma^{(4)}$ is lower than $\sigma^{(3)}$ if and only if z adds information that is not already provided by y. Therefore, the conditional analysis will not show y-conditional G-causality

$$G_{z\to x}^{(y)} = \ln \frac{\sigma^{(3)}}{\sigma^{(4)}} \tag{6}$$

for the indirect $z \to x$ links shown in **Figures 1B,C**, but only for the direct link in **Figure 1A**. Besides, cases (b) and (c) can be distinguished by a pairwise analysis of z and y (disregarding x). This idea can be extended to a set of arbitrarily many variables, but for the scope of this work three variables (time series x_t, y_t, and z_t) are sufficient.

Stationarity is an important prerequisite for the AR framework, because the process' characteristics (i.e., the coefficients ϕ_i, ψ_i, τ_i) must not change over time. A stationary process is a process with a constant mean and a finite covariance function that is invariant to shifts in time. This requirement is problematic, because many physiological signals are inherently non-stationary [Ivanov et al. (1996); Goldberger et al. (2002)]. Here, we probe stationarity with the Augmented Dickey-Fuller (ADF) test, which is widely accepted [Paparoditis and Politis (2018)] and based on AR modeling, so that it operates in the same framework as G-causality analysis.

In summary: G-causality is based on the improvement of a forecast by including additional data from other signals, and it can only be applied to stationary data. It can be used to distinguish between the three setups shown in **Figure 1**, if pairwise analysis and conditional analysis are applied. With G-causality the existence of a causal link can be decided as yes/no question with a statistical test, and quantified with the G-value.

2.3 Method 2: Bivariate Phase Rectified Signal Averaging

While the G-causality approach is sensitive to non-stationarities, the Phase Rectified Signal Averaging method [PRSA, Bauer et al. (2006b)] and its bivariate version [BPRSA, Schumann et al. (2008)] have been developed to study noisy, non-stationary signals. The methods are especially suitable for quasi-periodic time-series, where perturbations reset the signal phase at random times. The BPRSA approach can easily be extended to an arbitrary number of signals. The idea is to align windows of the target signal x that are in the same phase with respect to one or more trigger signals (y and z) and average over all these windows. The procedure is described in detail in Schumann et al. (2008), see also Bauer et al. (2006b,a, 2009); here we only provide a brief overview.

The easiest and standard way to define trigger events from signal z is to consider all positions, where the signal increases, i. e., $z_{t_\nu} > z_{t_\nu - 1}$ for a trigger event at time t_ν. We denote all m trigger points by t_ν, $\nu = 1, \ldots, m$. Any other criterion that returns a Boolean value (trigger event or no trigger event) is possible, including criteria that are based on multiple signals. Each trigger event at t_ν leads to an anchor point x_{t_ν} of the target signal. Then, windows of width $2L$ are chosen around each anchor point x_{t_ν},

$$\left\{x_{t_\nu - L}, x_{t_\nu - L + 1} \ldots, x_{t_\nu + L - 1}\right\}. \tag{7}$$

The resulting BPRSA function for a potential $z \to x$ link is the point-wise average of x_t in all of these m windows,

$$\mathrm{BPRSA}_j = \frac{1}{m} \sum_{\nu=1}^{m} x_{t_\nu + j}, \quad j = -L, -L + 1, \ldots, L - 1. \tag{8}$$

The choice of anchor points is supposed to guarantee that in each window the index $j = 0$ is at a similar phase of the physiological process. The average over all windows is an in-phase superposition and therefore insensitive to non-stationarities (that are slower than the time scale L) and artifacts.

If there is no relation between z and x, and m is large, the resulting BPRSA time series will be constant everywhere, however with statistical fluctuations. This also happens if the choice of trigger points is not appropriate to reveal a relation between the signals, because it does not reflect the underlying processes. Any significant deviation from a constant value for any BPRSA_j must be interpreted as a relation between z and x, but not necessarily in the sense of G-causality. A positive (negative) peak in the BPRSA time series indicates the positive (negative) influence of a trigger event in z on x, or—in case the peak is at a negative index j—from x on z. Thus, BPRSA yields temporal information on cause and effect, just like G-causality. However, we would like to note that—unlike G-causality analysis—BPRSA does not model the time series data in any way, but relies on an averaging procedure assuming that the selected trigger event criterion is suitable for the relation between the considered signals and that the Central Limit Theorem holds. Therefore, it is expected that longer data are needed for a reliable identification of causality relations with BPRSA.

Since no test for the statistical significance of such a relation has yet been proposed[1], we have studied and compared four tests. The null hypothesis is that there is no causality between z and x. If this holds, the trigger points are randomly distributed and each of the BPRSA_j values, **Eq. 8**, is an average over independent random numbers and thus normally distributed due to the central limit theorem.

The following statistical tests are considered and compared:

1) The one-sided Kolmogorov-Smirnov test [Massey (1951)] measures the difference between the probability density functions of the BPRSA and a normal distribution, providing a p-value for the null hypothesis.
2) The two-sided Kolmogorov-Smirnov test [Massey (1951)] compares the distribution of the real BPRSA values with the distribution of BPRSA values for random trigger points (i.e., disregarding the trigger signal z), also providing a p-value for the null hypothesis.
3) The Anderson-Darling test [Anderson and Darling (1952)] works similar to this idea, but introduces a weight function that increases the importance of the tails of the distribution. This is particularly useful if deviations from normality appear as abnormally high or low values instead of deviations in the middle of the bell-shaped curve.
4) The Shapiro-Wilk test [Shapiro and Wilk (1965)] is the most powerful of these tests according to Razali and Wah (2011). It is based on variance analysis and compares the variance of a normal distribution with the estimated variance of the sample.

In summary: The existence of causality in the sense of BPRSA can be tested by checking whether the BPRSA_j values are normally distributed. In addition, the peak height can provide quantitative information on the link strength. Compared to G-causality, this method is less sensitive to non-stationarities, and it is a model-free approach. However, BPRSA cannot distinguish direct from indirect links unless more evolved trigger criteria could be established for a conditional analysis.

2.4 Results and Discussion: Comparison of *G*-Causality and BPRSA Causality

2.4.1 Pairwise Analysis

With the tools presented above, both G-causality and BPRSA offer ways to test pairwise causality of two signals. In the first step, we quantified the detection limits. To this end, two signals of $1/f^\alpha$ noise with $\alpha = 0.5$ were created by the Fourier filtering method [Makse et al. (1996); Bashan et al. (2008)]. Starting with white noise, the power spectrum was rescaled to follow $1/f^{0.5}$ behavior, and—back in the time domain—the values were rescaled to have unit variance. These original noise signals are called o_1 and o_2. The signals z and x are defined as

$$z = o_1, \qquad x = \mathcal{L}^3 z \times q + o_2 \times (1 - q), \quad 0 < q < 1. \tag{9}$$

Here, $\mathcal{L}$ shifts the series by one time unit, so that $\mathcal{L}^3$ shifts it by three units; the number three was arbitrarily chosen. The unitless number q quantifies how much x is influenced by z. This setup was designed to test what influence is necessary for causality to be detected by the different methods. The length of the time series was varied from $2^6 = 64$ to $2^{16} = 65{,}536$ samples. The length of the BPRSA time series was chosen to be $2L = 30$, triggering on a rising signal [Bauer et al. (2006a,b); Schumann et al. (2008)]. Especially for the short time series large parts of the data had to be discarded due to overlap of window and boundary. While choosing a shorter window length might result in less discarded data, the averaged BPRSA time series would become shorter and therefore impair the quality of the statistical tests. The boundary effects are negligible for longer time series. For all tests the null hypothesis was no causality. In our setup z causes x with varying strength q. Therefore, a test yields the correct result if the null hypothesis is rejected. We rejected the null hypothesis for p-values lower than 0.05.

[1]Previous publications focus on the analysis of definitely existing relations [Bauer et al. (2006b,a); Schumann et al. (2008); Bauer et al. (2009, 2010)].

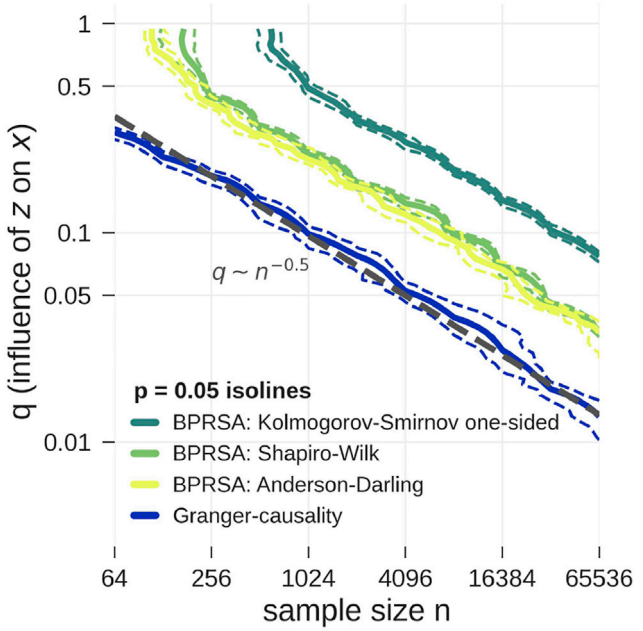


FIGURE 2 | Model data according to **Eq. 9** is tested for pairwise causality with statistical tests applied to *G*-values (blue) and BPRSA values (yellow and green). The plot shows the $p = 0.05$ isolines for each test as function of the influence strength q and the length of the time series (x_t and z_t). The isolines' 5 and 95% confidence interval are marked by dashed lines, computed from a bootstrap procedure. The tests correctly reject the null hypothesis *above* the shown isolines towards the top right corner of the plot. Clearly, the lowest detection limit is achieved by *G*-causality (applying the F-test). BPRSA causality with the Anderson-Darling test (3) and the Shapiro-Wilk test (4) is less sensitive by a factor of two to three in the influence strength q as compared to *G*-causality. BPRSA causality with the one-sided Kolmogorov-Smirnov test (1) is even less sensitive by another factor of 2, while the two-sided Kolmogorov-Smirnov test (2) only yielded the correct result in the top right corner of the plot and is therefore not shown. There seems to be a power-law relationship between the critical q-value as function of the sample size, $q \sim N^{-0.5}$.

Figure 2 shows the dependency of the threshold for the identification of existing causality on sample size and link strength for all proposed tests. All experiments were averaged over 20 realizations in order to get a statistically reliable result. The confidence intervals are given by the 5th and 95th percentile of a bootstrap distribution, which is obtained from 100 random samples. For each sample, we drew 20 out of the 20 realizations, allowing individual realizations to be picked multiple times.

In the limit of large sample sizes and strong influence q, all tests correctly identified causality (upper right corner of **Figure 2**). In the limit of small sample size and weak influence q, all tests failed (lower left corner of **Figure 2**). The detection limit is given by the $p = 0.05$ isoline. For small sample sizes (64 points or less) BPRSA is unable to detect causality, no matter how strong the link, possibly because a non-negligible amount of data near the boundaries is discarded. The two-sided Kolmogorov-Smirnov test failed to detect causality in all cases and is therefore not shown in **Figure 2**. The one-sided version turned out more powerful, but remained weaker than the Anderson-Darling test and the Shapiro-Wilk test. These results are in line with the findings of Razali and Wah (2011). The *G*-causality test detects causality for smaller sample sizes and weaker influences than all BPRSA tests.

In the calculations for **Figure 2**, BPRSA causality methods turned out to need about two times more computational effort than *G*-causality, and all BPRSA tests failed for short time series. As mentioned above, it is expected that BPRSA needs longer data for a significant result, since it relies on the Central Limit Theorem. For nonstationary data, BPRSA can still be applied, but even longer data would be needed, since the non-stationarities must cancel out in its averaging procedure. Furthermore, BPRSA is dependent on the choice of the trigger criterion. Only if a trigger criterion that fits to the physiological process is chosen, relationships can be established. Without prior knowledge, several trials with all kinds of criteria need to be done in order to probe for BPRSA causality. *G*-causality does not require such procedure, but also cannot be used to test for different kinds of time-varying relations between the considered signals in a system with non-stationary dynamics. Therefore, BPRSA can still be advantageous if specific hypotheses on the nature of the signals' relations exist or if the signals' relations change in time.

Both methods provide information on the direction of coupling, while methods such as cross correlation analysis or cross-spectral analysis are symmetric in the sense that a coupling of signal x with signal z is always also a coupling of z with x. BPRSA and *G*-causality both overcome this problem. We note that in **Figure 2** a power-law relationship between the critical q-value as function of the sample size can be seen, however, a detailed study of this scaling is beyond the scope of the present work.

In summary, *G*-causality is a more powerful method than BPRSA causality. Still, the BPRSA method is more likely to be able to handle nonstationarities because of its natural strength in phase-aligning the signal parts. However, methods have been developed to overcome the stationary constraint and apply time-dependent AR-models [Ding et al. (2000)], which allow the definition of time-dependent *G*-causality [Hesse et al. (2003)]. We concluded that *G*-causality is better suited to detect causality in multivariate, stationary setups and continued our analysis with this method.

2.4.2 Conditional Analysis

Consider the setup shown in **Figure 1C**. There is a relation between z and x, but it only exists because of the common influence of y. While BPRSA cannot distinguish between a direct influence of z on x or a common driver y on z and x, with *G*-causality these cases can be separated. This is a real strength of *G*-causality, and in this subsection we are testing the limits of such detection. One way to formalize the setup is by extending **Eq. 9** to three variables, with o_1, o_2 and o_3 again being independent $1/f^{0.5}$ noise signals,

$$\begin{aligned} y &= o_2; \\ z &= \mathcal{L}^2 y \times q_{y\to z} + o_1 \times \left(1 - q_{y\to z}\right); \quad 0 < q_{y\to z} < 1 \\ x &= \mathcal{L}^4 y \times q_{y\to x} + o_3 \times \left(1 - q_{y\to x}\right); \quad 0 < q_{y\to x} < 1. \end{aligned} \tag{10}$$

We performed pairwise and conditional analysis, see **Eqs. 1–5**, and tested for *G*-causality with an F-test for signals of sample size

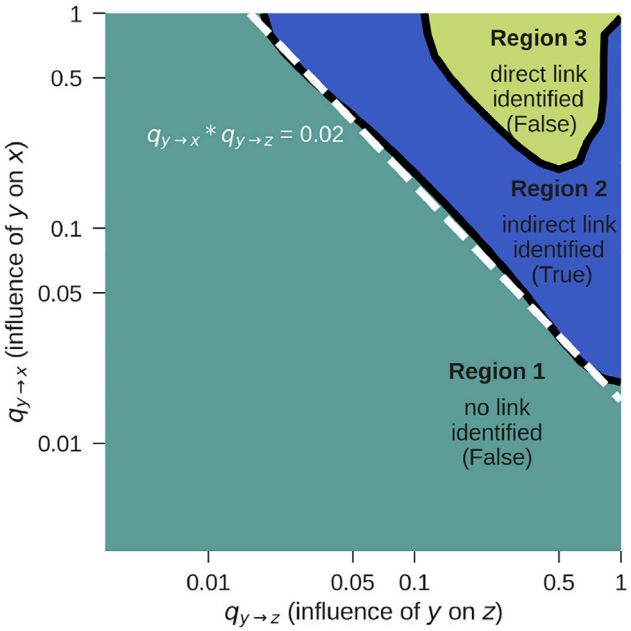


FIGURE 3 | *G*-causality F-test for $z \to x$ for the setup from **Figure 1C**, according to **Eq. 10**. The correct result is obtained in Region 2. For further details see text.

2^{15} = 32,768 and varying influences $q_{y\to z}$, $q_{y\to x}$. The results, averaged over 20 realizations, are shown in **Figure 3**.

We identified three regions:

- **Region 1**: Both, the pairwise and the conditional test do not reject the null hypothesis that there is no *G*-causality. This would lead to the conclusion that there is neither a direct nor an indirect link, which is false. The reason for the false negative result is that the influence coefficients $q_{y\to z}$ and $q_{y\to x}$ are too weak. Longer time series will shift this detection limit towards the lower left corner of the figure.
- **Region 2**: While the pairwise test rejects the null hypothesis, the conditional analysis does not. This leads to the conclusion that there is a $z \to x$ link, which can only be indirect. This is indeed true and corresponds to the case shown in **Figure 1C**.
- **Region 3**: Both the pairwise as well as the conditional test reject the null hypothesis, indicating that there is a direct $z \to x$ link, which is false. If $q_{y\to z}$ and $q_{y\to x}$ become so strong that x and z are both very tightly coupled to y, *G*-causality is mistakenly detected. In practice, this limitation rarely applies, because signals are usually not free of noise and not coupled so strongly.

If y is known, the method works for $\log\, q_{y\to z} + \log\, q_{y\to x} = \log\,(q_{y\to z} \cdot q_{y\to x}) \geq \log\,(0.02) = -1.7$ (dashed line, Region 2), as long as the link is not extremely strong (Region 3). If the influence is too weak, no causality will be detected at all. If it is extremely strong, the indirect link will be mistaken for a direct link.

In practice, measurements will always be limited to certain variables, and there is no way to exclude the possibility of external variables y that constitute a common cause to the measured signals. For this reason, there can practically never be certainty if a detected link is direct or due to a common source. One approach to overcome this is to include as many variables in the model as possible, which is, however, problematic because an increasing number of parameters must be estimated in this procedure. Prior knowledge on the modeled process can improve the interpretation of the results.

3 PART B: RECONSTRUCTION OF CAUSAL PHYSIOLOGICAL NETWORKS

In this part, conditional *G*-causality is applied to detect direct physiological couplings between heart rate, breathing rate, and EEG amplitude during sleep. We will show in the following that the couplings among these physiological systems differ between groups of young healthy subjects, elderly healthy subjects and patients with obstructive sleep apnea (OSA). OSA is the temporary, complete or partial disruption of normal respiration during sleep, caused by a reduced tonus of upper airways muscles [Dempsey et al. (2010)]. The increased negative intrathoraic pressure upon inspiration causes the upper airways to collapse, which results in a drop of blood oxygen and increase in blood carbon dioxide levels. This leads to an arousal from sleep, followed by recovery of normal respiration [Penzel et al. (2003b)].

3.1 Methods

Physiological time series were derived from polysomnography (PSG) measurements that were recorded in several European sleep laboratories between September 1997 and April 2000 as part of the EU-project SIESTA [Klosch et al. (2001)]. Before any analysis, we chose 36 young, healthy subjects with excellent signal quality (young control group—YC, aged 29 ± 6), 36 elderly, healthy subjects (elderly control group—EC, aged 51 ± 10) and 43 age-matched, elderly subjects with an apnea-hypopnea index (AHI) of at least 10 per hour (OSA group, aged 51 ± 9). AHI is the mean number of apnea and hypopnea events per hour when considering a full-night sleep. Genders are distributed approximately equally in the YC (17 male, 19 female) and EC group (18 male, 18 female), but the OSA group consists mostly of male participants (38 male, five female). We address this in the results section. For each subject, we derived:

- **Instantaneous heart rate H** as the inverse RR-interval, i.e., the time between two successive heart beats.
- **Instantaneous breathing rate B** from the inverse interval between two extrema of the raw respiration signal. The raw respiration signal was chosen for each subject individually as the best-quality signal out of effective oronasal airflow and stretch belts placed around abdomen and thorax.
- **EEG** α instantaneous amplitude by applying a bandpass filter on the EEG signal using the α frequency band 7.8–15.6 Hz.

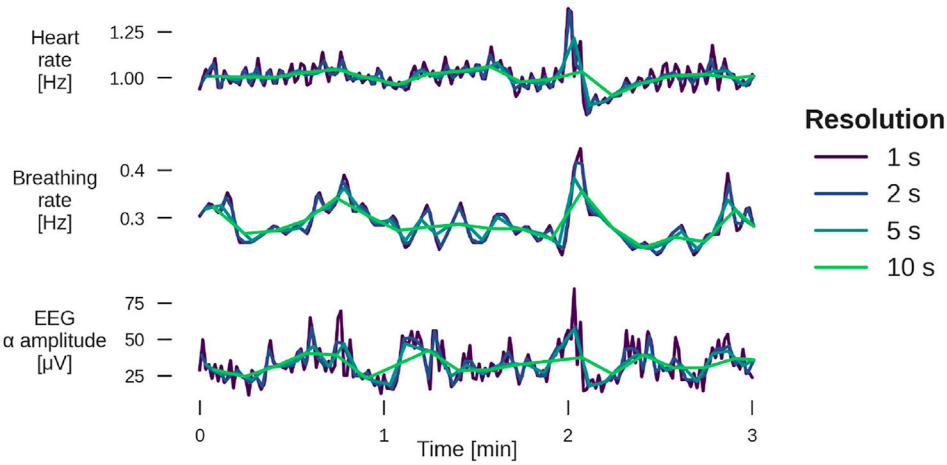


FIGURE 4 | Example of 3 min segments from a YC subject for heart rate (top panel), breathing rate (middle), and EEG α amplitude (bottom) for four different temporal resolutions. The sleep stage is light sleep. For examples from other sleep stages, see **Supplementary Figures S1, S2**.

Subsequently, signals H and B were interpolated to 1 Hz resolution, for EEG α, averages over non-overlapping windows of one second length were taken. The resulting three time series were all sampled at 1 Hz, and were averaged to resolutions of 2, 5, 10, 15, 30, and 60 s (i.e., 0.5, 0.2, 0.1, 0.067, 0.033, and 0.017 Hz) for further analysis (see **Figure 4**).

For each block of 30 seconds, sleep stages were scored by a sleep technician on the basis of the PSG signal following the rules by Rechtschaffen and Kales [Hobson (1969)]: light sleep (stages 1 and 2—LS), deep sleep (stages 3 and 4—DS) and REM sleep. Each triple of time series (H, B and EEG α) was partitioned into patches of continuous sleep of the same sleep stage, typically several minutes, and then normalized to zero mean and unity variance.

Since stationarity is a crucial precondition for *G*-causality analysis, a stationarity test must precede further analysis. We based our algorithm on the widely accepted Augmented Dickey-Fuller Test [ADF test, Paparoditis and Politis (2018)]. On the one hand, only stationary patches can be used for *G*-causality analysis, on the other hand the percentage of usable patches should be maximized in order to obtain the best possible statistics. A trade-off can be achieved through variation of the model order. For each continuous patch of the same sleep stage, all three signals of the node triple (i.e., heart rate, respiration rate and EEG amplitude) are tested for stationarity on a model of order 5. If all of them are stationary, the *G*-value is calculated according to **Eq. 6**. If at least one is nonstationary, the same procedure is repeated with model order 4, then with model order 3. If the process can still not be modeled as stationary at model order 3, the patch is split in half and the same procedure is applied to both shorter patches, motivated by the fact that shorter time series are more likely to be sufficiently free of trends and variability in variance and autocorrelations. A stopping condition is set when the patches reach a length of less than six times the time series resolution (i.e., 90 s length is the lower limit for time series of 15 s resolution), because this is the minimum length required for an AR model of order five. If this condition is met, the data is considered nonstationary and discarded.

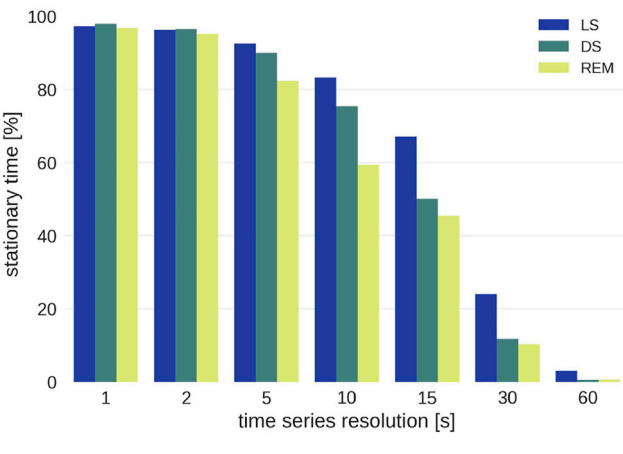


FIGURE 5 | Percentage of stationary sleep episodes as function of the time scale for the three different sleep stages. Note that wake epochs were excluded from the analysis because of insufficient statistics.

Of all the data that are potentially available, the percentage of stationary patches is shown in **Figure 5**. It does not make sense to analyze data for resolutions broader than 15 s as most of these patches do not contain stationary data. This is mainly because of two reasons: Firstly, there are too few sufficiently long sleep stage epochs for resolutions >15 s, and secondly, windows with an equal amount of data points but lower temporal resolutions cover longer sleep episodes, which are less likely to be stationary (i.e., compare high resolution 1 s time series with low resolution 60 s data). In our datasets, the second reason is dominant.

Even though resolutions above 15 s are of interest for research on long-term correlations, they cannot be included because there is not enough stationary data. For each group the conditional *G*-values (**Eq. 6**) for each time series resolution and each sleep stage were averaged and weighted by the length of the patch that they

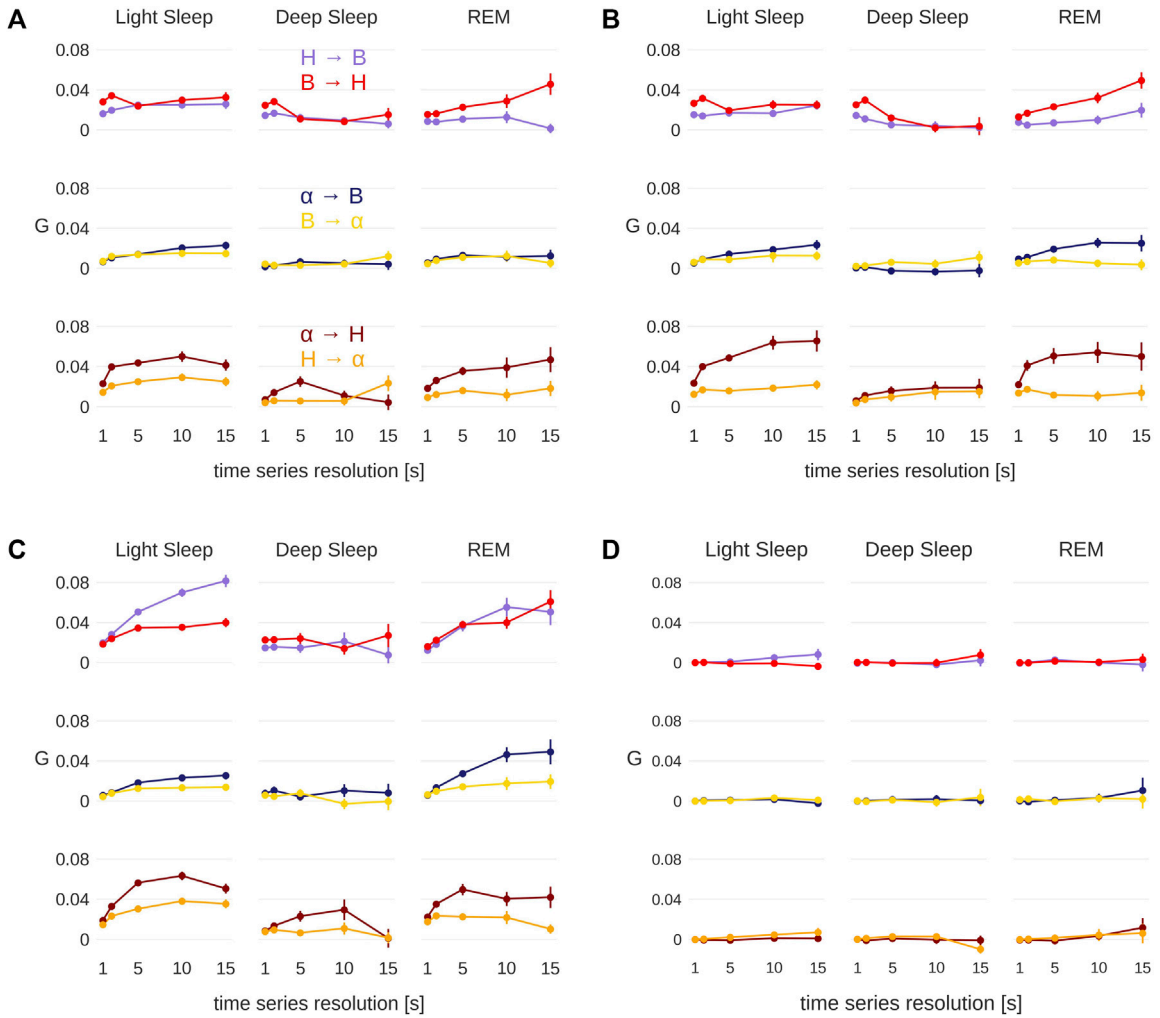


FIGURE 6 | Pairwise conditional *G*-causality for the three groups [YC—**(A)**, EC—**(B)**, OSA—**(C)**] and surrogate data **(D)**. Error bars represent the standard error and were calculated using a bootstrap method [Efron and Tibshirani (1993), see text for more details].

were calculated from in order to account for the varying length of sleep stages.[2]

Error bars were calculated using a bootstrap method [Efron and Tibshirani (1993)]: Out of all *G*-values that comprise a data point, a new set of *G*-values was randomly drawn. This set is as large as the original one, but the same *G*-value can be picked several times, while others are not included in a particular sample. Overall, 100 such samples were drawn and for each sample the (unweighted) mean was calculated. The standard deviation of these means is an estimate for the standard deviation of the actual mean, the standard error, and is plotted as error bar in **Figure 6**.

Each analyzed data patch consists of only tens to hundreds of data points, which is not enough to reliably detect weak links with the *G*-causality test (see **Figure 2**). Therefore, we compare our results with surrogate data that allow for an alternative way to validate *G*-causality: If the average *G*-value obtained for a particular group and pair of signals is different from the *G*-value of the surrogate data, *G*-causality can be assumed. For our surrogate analysis, the respiration rate, heart rate and EEG amplitude data were taken from three different subjects, so that no *G*-causality can be expected between any of the three signals [Toledo et al. (2002); Bartsch et al. (2007)]. An alternative method would be the adjusted-amplitude Fourier transform (AAFT) method [Theiler et al. (1992); Lavanga et al. (2020)].

3.2 Results

The results for all groups and the surrogate data are shown in **Figure 6**. The figure shows two of the six combinations in each

[2]We note that theoretically *G*-values cannot be negative, because this would imply a reduction in predictive power caused by adding new information, which is impossible. In the worst case the new information could just be neglected, which should lead to a *G*-value of zero. However, because of numerical reasons the minimization algorithm used for the parameter calculation of the models can lead to negative *G*-values in few cases. Instead of artificially removing those cases they were kept because in other cases the algorithm slightly overestimates the *G*-value. Removing the negative values would introduce a bias towards higher *G*-values.

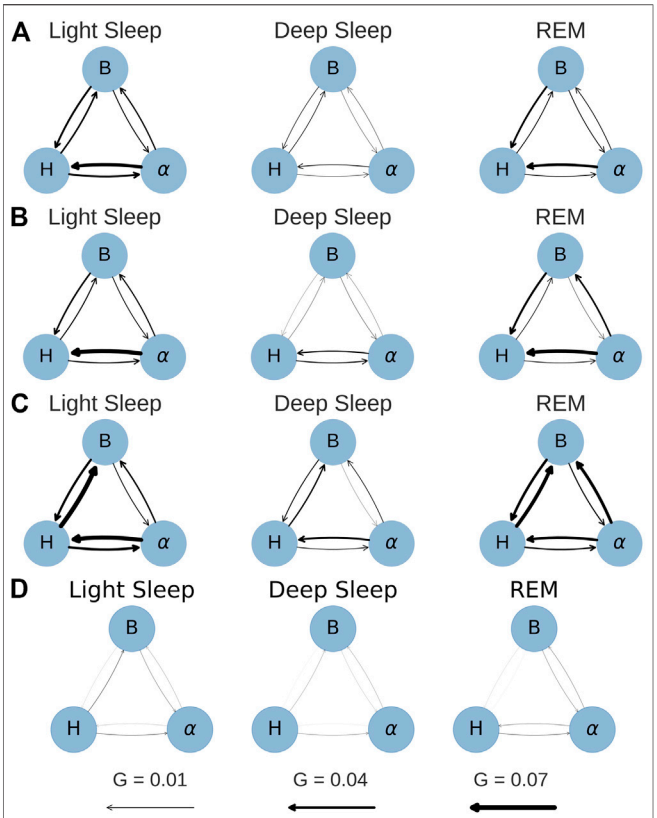


FIGURE 7 | Network plots for conditional *G*-causality at a time series resolution of 10 s. The arrow width is directly proportional to the link strength (*G*-value). **(A)** YC. **(B)** EC. **(C)** OSA. **(D)** Surrogate.

panel, grouped into pairs of signals for each sleep stage. The error bars represent the standard error calculated from the bootstrap procedure described above. Particularly for LS and REM, the difference in directionality increases for larger time windows (i.e., lower resolutions at 10 and 15 s), however, for 15 s the error bars are quite large in many cases. For this reason, time series with 10 s resolution were chosen to generate the physiological networks presented in **Figure 7**. Thus, the strength of the connection line between two nodes ("network link") is proportional to the *G*-value at a resolution of 10 s. While **Figure 7** presents less information than **Figure 6**, it is more helpful to find patterns and recognize important differences between the groups.

It is important to note that a time series resolution of 10 s does not mean that the relevant processes happen within 10 s. The model that is used to calculate the *G*-values comprises three to five past terms, which means that the processes occur in time windows of 30–50 s, but by default neglect causalities from processes with dynamics on time scales shorter than 10 s.

Subfigures (a), (b) and (c) in **Figures 6**, **7** show the results for the YC, EC and OSA group, respectively. The results for YC and EC are qualitatively very similar: The interaction between heart and respiration rate is symmetrical in light and deep sleep across all time scales and takes low to medium values ($G < 0.04$ at all times). This changes in REM sleep, where the causality from respiration to heart rate clearly dominates over the opposite direction, especially towards larger time scales. The *G*-value for the causality from respiration to heart rate at time scale 2 seconds is slightly increased (see **Figures 6A,B** only), which is an indicator of respiratory sinus arrhythmia (RSA), a well-known effect of modulation of the heart rate within the breathing cycle. During inspiration, the heart rate accelerates and it slows down during expiration [Angelone and Coulter (1964)]. This RSA peak disappears for REM sleep, which is in agreement with Bond et al. (1973), who describe "a total dissociation between respiration and rhythmic heart rate variability" during REM sleep. However, while RSA (which acts on relatively short time scales on the order of a breathing cycle) disappears in REM sleep, the causal relation from respiration to heart rate during REM sleep is shifted to longer time scales. The RSA peak is also absent in OSA subjects during non-REM sleep (see **Figure 6C**), possibly indicating reduced RSA for OSA subjects.

The coupling between EEG α amplitude and respiration rate remains, for all groups, constantly weak and symmetrical throughout all sleep stages and time series resolutions. The only deviation from this behavior is a slight increase in the causality from EEG α to respiration rate at larger time scales for EC and OSA subjects in REM sleep, which could be related to aging. Regarding the coupling between heart rate and EEG α, there is a clear dominance from EEG α to heart rate during light and REM sleep. This coupling almost completely vanishes during deep sleep, causing EEG α to heart rate coupling to become more symmetric.

The network plots 7(a) and (b) show the small differences between the YC and EC group. Furthermore, throughout all groups there is only little *G*-causality during deep sleep. This is in accordance with results of Bashan et al. (2012) and Bartsch et al. (2015) who show low network connectivity during deep sleep using a time delay stability approach to quantify interactions. At the same time, during deep sleep there is also a loss of long-term correlations in heart beats [Bunde et al. (2000); Penzel et al. (2003b)] as well as in respiratory inter-breath intervals [Schumann et al. (2010)]. Such long-term correlations, however, exist during light and REM sleep and are assumed to be due to influences from the sympathetic nervous system on cardiac and respiratory dynamics [Schmitt et al. (2009); Bashan et al. (2012)]. In contrast, deep sleep, which is considered the most restorative sleep stage [Dijk (2009)], is characterized by sympathetic withdrawal and greatly reduced influence of the autonomic nervous system on heart and respiratory dynamics. In our results, this is reflected by a more autonomous behavior of all three nodes for all groups. A careful comparison between the YC and EC group reveals that EEG α to heart rate coupling may slightly increase with age. For OSA patients this effect is even more pronounced and accompanied by an overall increase of coupling strength also between heart and breathing as well as EEG α to breathing, possibly indicating a decreased deep sleep quality because of sleep apnea.

In general, the OSA group follows similar sleep-stage patterns as the YC and EC groups. Distinct differences, however, can be seen in the overall strength of coupling: First and most strikingly, the *G*-causality from heart rate to respiration rate is much stronger during light and REM sleep than in the other two groups. This can be attributed to different relaxation speeds of heart and respiratory rate after apnea events. Following an episode of sleep apnea, heart and respiration rate are increased [Penzel et al. (2003a; 2016)], but the relaxation of the heart rate happens faster than the relaxation of the respiratory rate. This leads to a situation where changes in heart

rate precede changes in respiratory rate and thus lead to a detection of increased *G*-causality, as indicated by higher *G*-values. Secondly, there is an increased *G*-causality from EEG α amplitude to breathing during REM sleep.

Figures 6D, **7D** show results for the surrogate data. As expected, the *G*-value is very close to zero in all cases at all resolutions, with small deviations for the 15 s data point, which could be due to the large error bars at this resolution. These results can be used as a baseline showing that randomly assembled physiological data yield negligible *G*-values, and that any deviation from zero, as seen in all other sub-figures, are physiologically meaningful results.

Performing the same analysis without testing for stationarity yields slightly different results (see **Supplementary Figures S4, S5**). Values at short timescales are mostly unchanged, but it seems that the G-value is overestimated in some cases when ignoring the stationarity test. Additionally, the standard errors are larger, especially for longer time scales. Therefore, we conclude that only stationary data yield non-spurious network links when applying the *G*-causality method. The fact that the OSA group consists of mostly male subjects does not influence our results. We repeated the analyses of **Figures 6**, **7** excluding all female subjects from YC and EC, and obtained similar results (see **Supplementary Figures S6, S7**).

4 CONCLUSION AND OUTLOOK

In Part A, the causality methods BPRSA and *G*-causality were analyzed. Both methods are strong analysis tools to detect interrelations and causality in time series, but also have limitations. While BPRSA has rather weak testing methods and fails to distinguish direct from indirect links, the application of *G*-causality is limited due to the fact that it requires time series to be stationary (which is often not the case in physiology). For the setups investigated in this work, *G*-causality yielded better results than BPRSA causality tests. The *G*-value provides a measure of the strength of causality, and it can be computed for pairwise (bivariate) and conditional (multivariate) setups, where the latter includes additional information beyond the causing and the caused signal. Important for Network Physiology, this enables the distinction between direct and indirect links as well as links that arise from a common source signal. While (in contrast to BPRSA) the *G*-causality method is constrained by the stationarity condition, there are extensions to the method that circumvent the problem [Hesse et al. (2003); Dhamala et al. (2008); Bressler and Seth (2011)]. A particular, simple way to overcome this restriction is to split non-stationary data into shorter, stationary patches. Further investigations and alternative approaches to overcome the stationarity condition could be promising future research pathways.

In part B, the *G*-causality method was applied to the node triple consisting of heart rate, respiration rate, and EEG α amplitude, recorded from subjects with and without OSA. The *G*-value was calculated for time series resolutions between 1 and 15 s. Causal physiological networks were constructed based on these *G*-values. In all groups, strong coupling between respiration and heart rate and from EEG α to heart rate can be observed in light and REM sleep. In contrast, during deep sleep, the three nodes are practically "decoupled", especially for the young group. This result supports earlier findings and the understanding of deep sleep as the sleep stage with lowest sympathetic tone. Because aging changes sympathovagal balance due to a reduced parasympathetic tone [Schmitt et al. (2009)], leading effectively to higher sympathetic activity, the *G*-causality coupling in deep sleep is slightly increased for the elderly and OSA groups. Apart from deep sleep, results are very similar for young and elderly healthy subjects, however, OSA subjects show some distinct differences. Compared to the other two groups, the most prominent difference is an increase in *G*-causality from heart rate to respiration rate in light and REM sleep due to different relaxation times of heart rate and respiration rate, which are both increased at the end of an apnea event. Disturbances in respiratory sinus arrhythmia during light and deep sleep and a stronger causal link from EEG α to breathing rate during REM sleep can also be observed in OSA subjects.

Our findings point to the conclusion that sleep of persons with sleep apnea is not only different with respect to breathing behavior, but also with respect to coupling mechanisms like respiratory sinus arrhythmia and deep sleep decoupling. Comparisons with surrogate data prove the significance of the obtained results. Overall, the application to causal networks in subjects with and without sleep apnea demonstrates the usefulness of *G*-causality as a measure for physiological coupling.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: All code is available at the following GitHub repository: https://github.com/moritz-g/causal_networks_physiology. Data can be obtained upon request from the SIESTA Group. Requests to access these datasets should be directed to www.thesiestagroup.com.

AUTHOR CONTRIBUTIONS

MG performed the data analysis, prepared the figures, and wrote the initial draft of the paper. JK and RB conceived the project and supervised the research. All authors discussed the results and wrote the paper.

FUNDING

The study was supported by the German-Israeli Foundation (GIF) Grant I-1372-303.7/2016. MG acknowledges support from a Minerva Research Grant.

ACKNOWLEDGMENTS

We would like to thank Prof. G. Dorffner and the SIESTA group for providing the data for this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnetp.2022.893743/full#supplementary-material

REFERENCES

Anderson, T. W., and Darling, D. A. (1952). Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *Ann. Math. Statist.* 23, 193–212. doi:10.1214/aoms/1177729437

Angelone, A., and Coulter, N. A. (1964). Respiratory Sinus Arrhythmia: a Frequency Dependent Phenomenon. *J. Appl. Physiol.* 19, 479–482. doi:10.1152/jappl.1964.19.3.479

Bartsch, R., Kantelhardt, J. W., Penzel, T., and Havlin, S. (2007). Experimental Evidence for Phase Synchronization Transitions in the Human Cardiorespiratory System. *Phys. Rev. Lett.* 98, 054102. doi:10.1103/PhysRevLett.98.054102

Bartsch, R. P., Liu, K. K., Ma, Q. D., and Ivanov, P. C. (2014). Three Independent Forms of Cardio-Respiratory Coupling: Transitions across Sleep Stages. *Comput. Cardiol. (2010)* 41, 781–784.

Bartsch, R. P., and Ivanov, P. C. (2014). Coexisting Forms of Coupling and Phase-Transitions in Physiological Networks. *Commun. Comp. Inf. Sci.* 438, 270–287. doi:10.1007/978-3-319-08672-9_33

Bartsch, R. P., Liu, K. K. L., Bashan, A., and Ivanov, P. C. (2015). Network Physiology: How Organ Systems Dynamically Interact. *PLoS One* 10, e0142143. doi:10.1371/journal.pone.0142143

Bashan, A., Bartsch, R. P., Kantelhardt, J. W., Havlin, S., and Ivanov, P. Ch. (2012). Network Physiology Reveals Relations between Network Topology and Physiological Function. *Nat. Commun.* 3, 702–709. doi:10.1038/ncomms1705

Bashan, A., Bartsch, R., Kantelhardt, J. W., and Havlin, S. (2008). Comparison of Detrending Methods for Fluctuation Analysis. *Physica A: Stat. Mech. its Appl.* 387, 5080–5090. doi:10.1016/j.physa.2008.04.023

Bauer, A., Barthel, P., Müller, A., Kantelhardt, J., and Schmidt, G. (2009). Bivariate Phase-Rectified Signal Averaging-A Novel Technique for Cross-Correlation Analysis in Noisy Nonstationary Signals. *J. Electrocardiol.* 42, 602–606. doi:10.1016/j.jelectrocard.2009.06.023

Bauer, A., Kantelhardt, J. W., Barthel, P., Schneider, R., Mäkikallio, T., Ulm, K., et al. (2006a). Deceleration Capacity of Heart Rate as a Predictor of Mortality after Myocardial Infarction: Cohort Study. *The Lancet* 367, 1674–1681. doi:10.1016/S0140-6736(06)68735-7

Bauer, A., Kantelhardt, J. W., Bunde, A., Barthel, P., Schneider, R., Malik, M., et al. (2006b). Phase-rectified Signal Averaging Detects Quasi-Periodicities in Non-stationary Data. *Physica A: Stat. Mech. its Appl.* 364, 423–434. doi:10.1016/j.physa.2005.08.080

Bauer, A., Morley-Davies, A., Barthel, P., Müller, A., Ulm, K., Malik, M., et al. (2010). Bivariate Phase-Rectified Signal Averaging for Assessment of Spontaneous Baroreflex Sensitivity: Pilot Study of the Technology. *J. Electrocardiol.* 43, 649–653. doi:10.1016/j.jelectrocard.2010.05.012

Bond, W. C., Bohs, C., Ebey, J., and Wolf, S. (1973). Rhythmic Heart Rate Variability (Sinus Arrhythmia) Related to Stages of Sleep. *Conditional Reflex* 8, 98–107. doi:10.1007/bf03000310

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

Bunde, A., Havlin, S., Kantelhardt, J. W., Penzel, T., Peter, J.-H., and Voigt, K. (2000). Correlated and Uncorrelated Regions in Heart-Rate Fluctuations during Sleep. *Phys. Rev. Lett.* 85, 3736–3739. doi:10.1103/physrevlett.85.3736

Dempsey, J. A., Veasey, S. C., Morgan, B. J., and O'Donnell, C. P. (2010). Pathophysiology of Sleep Apnea. *Physiol. Rev.* 90, 47–112. doi:10.1152/physrev.00043.2008

Dhamala, M., Rangarajan, G., and Ding, M. (2008). Estimating granger Causality from Fourier and Wavelet Transforms of Time Series Data. *Phys. Rev. Lett.* 100, 018701. doi:10.1103/PhysRevLett.100.018701

Dijk, D. J. (2009). Regulation and Functional Correlates of Slow Wave Sleep. *J. Clin. Sleep Med.* 5, S6–S15. doi:10.5664/jcsm.5.2s.s6

Ding, M., Bressler, S. L., Yang, W., and Liang, H. (2000). Short-window Spectral Analysis of Cortical Event-Related Potentials by Adaptive Multivariate Autoregressive Modeling: Data Preprocessing, Model Validation, and Variability Assessment. *Biol. Cybernetics* 83, 35–45. doi:10.1007/s004229900137

Efron, B., and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hall/CRC Monographs on Statistics and Applied Probability.

Faes, L., Marinazzo, D., Jurysta, F., and Nollo, G. (2015). Linear and Non-linear Brain-Heart and Brain-Brain Interactions during Sleep. *Physiol. Meas.* 36, 683–698. doi:10.1088/0967-3334/36/4/683

Goldberger, A. L., Amaral, L. A. N., Hausdorff, J. M., Ivanov, P. C., Peng, C.-K., and Stanley, H. E. (2002). Fractal Dynamics in Physiology: Alterations with Disease and Aging. *Proc. Natl. Acad. Sci. U.S.A.* 99, 2466–2472. doi:10.1073/pnas.012579499

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* 37, 424–438. doi:10.2307/1912791

Granger, C. W. J. (1980). Testing for Causality. *J. Econ. Dyn. Control* 2, 329–352. doi:10.1016/0165-1889(80)90069-x

Granger, C. W. J. (2004). Time Series Analysis, Cointegration, and Applications. *Am. Econ. Rev.* 94, 421–425. doi:10.1257/0002828041464669

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Hesse, W., Möller, E., Arnold, M., and Schack, B. (2003). The Use of Time-Variant EEG Granger Causality for Inspecting Directed Interdependencies of Neural Assemblies. *J. Neurosci. Methods* 124, 27–44. doi:10.1016/s0165-0270(02)00366-7

Hlavackovaschindler, K., Paluš, M., Vejmelka, M., and Bhattacharya, J. (2007). Causality Detection Based on Information-Theoretic Approaches in Time Series Analysis. *Phys. Rep.* 441, 1–46. doi:10.1016/j.physrep.2006.12.004

Hobson, J. A. (1969). "A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects," in *Public Health Service, US Government Printing Office*. Editors A. Rechtschaffen and A. Kales (Washington, DC: Elsevier). doi:10.1016/0013-4694(69)90021-2

Ivanov, P. C., Peng, C.-K., Mietus, J., Havlin, S., Stanley, H. E., Goldberger, A. L., et al. (1996). Scaling Behaviour of Heartbeat Intervals Obtained by Wavelet-Based Time-Series Analysis. *Nature* 383, 323–327. doi:10.1038/383323a0

Klösch, G., Kemp, B., Penzel, T., Schlögl, A., Rappelsberger, P., Trenker, E., et al. (2001). The SIESTA Project Polygraphic and Clinical Database. *IEEE Eng. Med. Biol. Mag.* 20, 51–57. doi:10.1109/51.932725

Lavanga, M., Bollen, B., Jansen, K., Ortibus, E., Naulaers, G., Van Huffel, S., et al. (2020). A Bradycardia-Based Stress Calculator for the Neonatal Intensive Care Unit: A Multisystem Approach. *Front. Physiol.* 11. doi:10.3389/fphys.2020.00741

Lin, A., Liu, K. K. L., Bartsch, R. P., and Ivanov, P. C. (2016). Delay-correlation Landscape Reveals Characteristic Time Delays of Brain Rhythms and Heart Interactions. *Phil. Trans. R. Soc. A.* 374, 20150182. doi:10.1098/rsta.2015.0182

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin Heidelberg: Springer.

Makse, H. A., Havlin, S., Schwartz, M., and Stanley, H. E. (1996). Method for Generating Long-Range Correlations for Large Systems. *Phys. Rev. E* 53, 5445–5449. doi:10.1103/physreve.53.5445

Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *J. Am. Stat. Assoc.* 46, 68–78. doi:10.1080/01621459.1951.10500769

Montero-Nava, J. E., Pliego-Carrillo, A. C., Ledesma-Ramírez, C. I., Peña-Castillo, M. Á., Echeverría, J. C., Pacheco-López, G., et al. (2020). Analysis of the Fetal Cardio-Electrohysterographic Coupling at the Third Trimester of Gestation in Healthy Women by Bivariate Phase-Rectified Signal Averaging. *Plos one* 15, e0236123. doi:10.1371/journal.pone.0236123

Müller, A., Kraemer, J. F., Penzel, T., Bonnemeier, H., Kurths, J., and Wessel, N. (2016). Causality in Physiological Signals. *Physiol. Meas.* 37, R46–R72. doi:10.1088/0967-3334/37/5/r46

Müller, A., Morley-Davies, A., Barthel, P., Hnatkova, K., Bauer, A., Ulm, K., et al. (2012). Bivariate Phase-Rectified Signal Averaging for Assessment of Spontaneous Baroreflex Sensitivity: Normalization of the Results. *J. Electrocardiol.* 45, 77–81. doi:10.1016/j.jelectrocard.2011.07.010

Neusser, K. (2011). *Zeitreihenanalyse in Den Wirtschaftswissenschaften*. Wiesbaden: Vieweg.

Paparoditis, E., and Politis, D. N. (2018). The Asymptotic Size and Power of the Augmented Dickey-Fuller Test for a Unit Root. *Econometric Rev.* 37, 955–973. doi:10.1080/00927872.2016.1178887

Pearl, J. (2009). *Causality*. Cambridge: Cambridge University Press.

Pearl, J., and Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books.

Penzel, T., Kantelhardt, J. W., Bartsch, R. P., Riedl, M., Kraemer, J. F., Wessel, N., et al. (2016). Modulations of Heart Rate, Ecg, and Cardio-Respiratory Coupling

Observed in Polysomnography. *Front. Physiol.* 7, 460. doi:10.3389/fphys.2016.00460

Penzel, T., Kantelhardt, J. W., Grote, L., Peter, J., and Bunde, A. (2003a). Comparison of Detrended Fluctuation Analysis and Spectral Analysis for Heart Rate Variability in Sleep and Sleep Apnea. *IEEE Trans. Biomed. Eng.* 50, 1143–1151. doi:10.1109/tbme.2003.817636

Penzel, T., Kantelhardt, J. W., Lo, C.-C., Voigt, K., and Vogelmeier, C. (2003b). Dynamics of Heart Rate and Sleep Stages in Normals and Patients with Sleep Apnea. *Neuropsychopharmacol* 28, S48–S53. doi:10.1038/sj.npp.1300146

Porta, A., and Faes, L. (2015). Wiener–granger Causality in Network Physiology with Applications to Cardiovascular Control and Neuroscience. *Proc. IEEE* 104, 282–309.

Razali, N. M., and Wah, Y. B. (2011). Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *J. Stat. Model. analytics* 2, 21–33.

Schmitt, D. T., Stein, P. K., and Ivanov, P. C. (2009). Stratification Pattern of Static and Scale-Invariant Dynamic Measures of Heartbeat Fluctuations across Sleep Stages in Young and Elderly. *IEEE Trans. Biomed. Eng.* 56, 1564–1573. doi:10.1109/TBME.2009.2014819

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. *Phil. Trans. R. Soc. A.* 371, 20120191. doi:10.1098/rsta.2012.0191

Schulz, S., Haueisen, J., Bär, K.-J., and Voss, A. (2019). Altered Causal Coupling Pathways within the central-autonomic-network in Patients Suffering from Schizophrenia. *Entropy* 21, 733. doi:10.3390/e21080733

Schumann, A. Y., Bartsch, R. P., Penzel, T., Ivanov, P. C., and Kantelhardt, J. W. (2010). Aging Effects on Cardiac and Respiratory Dynamics in Healthy Subjects across Sleep Stages. *Sleep* 33, 943–955. doi:10.1093/sleep/33.7.943

Schumann, A. Y., Kantelhardt, J. W., Bauer, A., and Schmidt, G. (2008). Bivariate Phase-Rectified Signal Averaging. *Physica A: Stat. Mech. its Appl.* 387, 5091–5100. doi:10.1016/j.physa.2008.05.002

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

Shapiro, S. S., and Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* 52, 591–611. doi:10.1093/biomet/52.3-4.591

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Doyne Farmer, J. (1992). Testing for Nonlinearity in Time Series: the Method of Surrogate Data. *Physica D: Nonlinear Phenomena* 58, 77–94. doi:10.1016/0167-2789(92)90102-S

Toledo, E., Akselrod, S., Pinhas, I., and Aravot, D. (2002). Does Synchronization Reflect a True Interaction in the Cardiorespiratory System? *Med. Eng. Phys.* 24, 45–52. doi:10.1016/s1350-4533(01)00114-x

Wiener, N. (1956). *The Theory of Prediction*. New York: Modern mathematics for engineers.

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.

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.

frontiers | Frontiers in Network Physiology

TYPE Original Research
PUBLISHED 29 August 2022
DOI 10.3389/fnetp.2022.912733

OPEN ACCESS

EDITED BY
Luca Faes,
University of Palermo, Italy

REVIEWED BY
Gorana Mijatovic,
University of Novi Sad, Serbia
Karin Schiecke,
Friedrich Schiller University Jena, Germany

*CORRESPONDENCE
Carmen González,
cgz@quantiummedical.com

SPECIALTY SECTION
This article was submitted to Networks of Dynamical Systems, a section of the journal Frontiers in Network Physiology

RECEIVED 04 April 2022
ACCEPTED 08 July 2022
PUBLISHED 29 August 2022

CITATION
González C, Garcia-Hernando G, Jensen EW and Vallverdú-Ferrer M (2022), Assessing rheoencephalography dynamics through analysis of the interactions among brain and cardiac networks during general anesthesia. *Front. Netw. Physiol.* 2:912733. doi: 10.3389/fnetp.2022.912733

Assessing rheoencephalography dynamics through analysis of the interactions among brain and cardiac networks during general anesthesia

Carmen González[1,2]*, Gabriel Garcia-Hernando[1,2], Erik W. Jensen[2] and Montserrat Vallverdú-Ferrer[1]

[1]Biomedical Engineering Research Centre, CIBER of Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Universitat Politècnica de Catalunya, Barcelona, Spain, [2]Research and Development Department, Quantium Medical, Mataró, Spain

Cerebral blood flow (CBF) reflects the rate of delivery of arterial blood to the brain. Since no nutrients, oxygen or water can be stored in the cranial cavity due to space and pressure restrictions, a continuous perfusion of the brain is critical for survival. Anesthetic procedures are known to affect cerebral hemodynamics, but CBF is only monitored in critical patients due, among others, to the lack of a continuous and affordable bedside monitor for this purpose. A potential solution through bioelectrical impedance technology, also known as rheoencephalography (REG), is proposed, that could fill the existing gap for a low-cost and effective CBF monitoring tool. The underlying hypothesis is that REG signals carry information on CBF that might be recovered by means of the application of advanced signal processing techniques, allowing to track CBF alterations during anesthetic procedures. The analysis of REG signals was based on geometric features extracted from the time domain in the first place, since this is the standard processing strategy for this type of physiological data. Geometric features were tested to distinguish between different anesthetic depths, and they proved to be capable of tracking cerebral hemodynamic changes during anesthesia. Furthermore, an approach based on Poincaré plot features was proposed, where the reconstructed attractors form REG signals showed significant differences between different anesthetic states. This was a key finding, providing an alternative to standard processing of REG signals and supporting the hypothesis that REG signals do carry CBF information. Furthermore, the analysis of cerebral hemodynamics during anesthetic procedures was performed by means of studying causal relationships between global hemodynamics, cerebral hemodynamics and electroencephalogram (EEG) based-parameters. Interactions were detected during anesthetic drug infusion and patient positioning (Trendelenburg positioning and passive leg raise), providing evidence of the causal coupling between hemodynamics and brain activity. The provided alternative of REG signal processing confirmed the hypothesis that REG signals carry information on CBF. The simplicity of the technology, together with its low cost and easily interpretable outcomes, should provide a new opportunity for REG to reach

standard clinical practice. Moreover, causal relationships among the hemodynamic physiological signals and brain activity were assessed, suggesting that the inclusion of REG information in depth of anesthesia monitors could be of valuable use to prevent unwanted CBF alterations during anesthetic procedures.

KEYWORDS

general anesthesia, cerebral blood flow, electroencephalography, rheoencephalography, Poincaré plot descriptors, Granger causality

1 Introduction

In the last decades, medical devices have flooded operating theaters to provide healthcare professionals updated and reliable information on patient vital signs, as well as advanced algorithms aiming at improving patient care. Nonetheless, certain clinical signs are not included in standard patient monitoring during surgeries under general anesthesia, such as cerebral blood flow (CBF). Even though CBF is monitored in critical patients, it is not part of the standard of care since it is invasive or extremely unwieldy and expensive.

General anesthetics are known to affect brain hemodynamics, provoking changes in CBF that might interfere in the transit times of the anesthetics towards the target organ, the brain. The main research hypothesis of this research (Slupe and Kirsch, 2018; Tauber et al., 2021; Porta et al., 2022) suggests that CBF plays an important role in anesthesia and might be useful to enhance current algorithms used for depth of anesthesia monitoring. Moreover, to be accepted for standard clinical practice, a CBF monitor to be used for anesthesia titration should be easy to use, non-invasive and cost-effective, provide real time information and guarantee that it does not cause alterations in blood flow during its use.

Rheoencephalography (REG) is an explorative method of cerebral circulation that measures electrical impedance which allows a continuous observation of the blood flow in different cerebral regions. The principle of this method is that blood is a good electrical conductor, therefore any increase in blood volume will lead to a reduction of the brain electrical resistance, and this will be reflected in a decrease of REG pulse amplitude given a constant current. Therefore, REG would comply with the requirements of low-cost and effective CBF monitoring tool (Bodo, 2010; Moskalenko, 2015). REG signals have traditionally been analyzed by assessing the geometrical properties of the blood pulse waves in the time domain (Montgomery et al., 1995; Bodo et al., 2004), such as the duration of the anacrotic phase of the pulse, the maximum and minimum amplitudes, the slope and the area under the curve.

The closest technology to REG is impedance cardiography (ICG), since both share the same working principle based on the electrical bioimpedance. ICG measures the electrical impedance of the thoracic cavity and allows the assessment of several hemodynamic variables, such as cardiac output (CO), stroke volume (SV), left ventricular ejection time (LVET) and systemic vascular resistance (SVR), among others (Siedlecka et al., 2015). Due to the similarities between REG and ICG, and the positive clinical outcome of the use of ICG, the rationale behind the analysis of ICG waves will be applied to REG recordings for CBF estimation.

Within the field of time series nonlinear analysis, many features have been developed for signals characterization, such as the Lyapunov exponents, fractal dimension, Poincaré plot analysis or entropy. Even though none of those algorithms has been applied to REG signals, some authors have studied their performance in processing similar data, such as intracranial pressure (ICP) recordings. For instance (Lu et al., 2012), used multiscale entropy applied to ICP recordings to study their complexity in brain injured patients, concluding that multiscale entropy was a good predictor of mortality and favorable outcome in those patients. Another metric entropy, approximate entropy (ApEn) was selected by (Hornero et al., 2006) to analyze ICP signals in the pediatric population, providing evidence that decreased complexity in ICP was related to events of intracranial hypertension. However, entropy calculations are often cumbersome for real time applications; they could be a powerful tool for post hoc analysis but are not the optimal solution for patient bedside monitoring. In contrast, within the set of nonlinear algorithms applied to biomedical signals, Poincaré plot analysis has shorter computation times and has also been extensively used in physiological signal processing, namely in heart rate variability (HRV) analysis (Khandoker et al., 2013), hence being a suitable tool for REG analysis.

Related to Poincaré plot analysis (Dimitriev et al., 2016), analyzed by means of nonlinear dynamics based on Poincaré plots how the state of anxiety affected heart rate variability. Voss et al. have previously published on the effects of age and gender in short-term heart rate variability analyzed with Poincaré plots among other features (Voss et al., 2015). Other biological signals have been studied by means of Poincaré plots. Hayashi et al. related the delayed coordinates map to changes provoked by anesthesia in the electroencephalogram (EEG) (Hayashi et al., 2015). Hoshi et al. used standard features of Poincaré plot analysis to distinguish between healthy subjects and patients suffering coronary disease, concluding that the SD1/SD2 index

provided useful information for that purpose (Hoshi et al., 2013). Even though some features extracted from Poincaré plots are known to be highly correlated to linear time domain information, some others reflect nonlinear behaviors, complementing the diagnosis capabilities of heart rate variability signals, such as the SD1/SD2 parameter or the Complex Correlation Measure (Karmakar et al., 2011).

The causality analysis of different physiological signals has gained popularity in the last decade. Most of the causality studies of biomedical signals have been published on the analyses of causal relationships between heart period, systolic arterial pressure (SAP) and respiration (Faes et al., 2011; Porta et al., 2014). Relevant clinical results have arisen such as the work from (Rield et al., 2010) exploring short-term couplings between respiration, systolic and diastolic blood pressure and heart rate, in order to have a deeper understanding on pre-eclampsia, which is responsible for significant neonatal and maternal mortality. Additionally, Porta et al. studied the causal interactions between heart period, respiration and systolic arterial pressure at rest and after the administration of different drugs, concluding that Granger causality is a suitable tool to describe cardiovascular control and the effects of the administered drugs (Porta et al., 2013). Schulz et al. studied cardiorespiratory causality couplings involved in the processes of an autonomic dysfunction present in patients suffering from schizophrenia (Schulz et al., 2020).

Besides interactions in the hemodynamics system, several publications have focused on the application of Granger Causality on EEG signals. For instance (Juan et al., 2017), studied the connectivity across EEG frequency bands in patients with Alzheimer's disease, detecting increments of connectivity in the δ band, together with decrement connectivity in other EEG frequency bands. They concluded that Granger Causality (GC) was suitable for Alzheimer diagnosis, since the disconnection among different brain regions is a well-known effect of the disease. Another application of GC in EEG signals was presented by (Coben and Mohammad-Rezazadeh. 2015), who analyzed pre and post-ictal periods of epileptic seizures to study the connectivity between brain regions in epileptic patients.

The GC principles have also been applied to EEG signals during anesthesia. Nicolaou et al., developed a system capable of classifying EEG signals as belonging to awake or anesthetized patients with a 96% accuracy, using as inputs the interactions between EEG signals from different brain areas (Nicolaou and Georgiou, 2014). Moreover, in (Nicolaou et al., 2012) an accuracy of 98% was obtained for loss of consciousness detection, suggesting that GC could be used as an awareness detection system. Barrett et al. analyzed steady state EEG signals during propofol induced anesthesia recorded from the anterior and posterior brain areas, detecting a bilateral increase in GC for the power spectral density in the β and γ bands during loss of consciousness (Barrett et al., 2012).

The interactions between the brain and the hemodynamic system have also been the target of many research projects. Duggento et al. analyzed functional magnetic resonance imaging data, respiration and heartbeat recordings, concluding that GC is a suitable tool to assess causality among brain and heart activity (Duggento et al., 2016). In (Greco et al., 2019), it was studied the causality between hemodynamics and EEG activity during the exposure to pleasant or unpleasant visual stimulation, to relate the reaction to emotions with the changes at the cardiovascular and brain level. Pleasant images increased the coupling from the left hemisphere to the heart, while unpleasant images increased the coupling with the right one, when compared to GC indices at rest. An analysis of brain, cardiovascular and respiratory dynamics was conducted by (Zanetti et al., 2019) combining information-theoretic measures with network physiology during different levels of mental stress. Results indicated that a characterization of these networks is possible in terms of the amount of information transferred within and between the brain and peripheral subnetworks. Faes et al. analyzed causal relationships brain-heart and brain-brain during sleep and concluded that both kinds of interactions were effectively taking place (Faes et al., 2015). Moreover, brain-heart interactions were also studied by (Won et al., 2019) for different sedation levels in anesthetic procedures. EEG spectral power and heart rate signals were analyzed, showing a higher connectivity from brain to heart when compared with the opposite direction for all sedation levels, finding as well a higher coupling in deeper sedation states.

Therefore, the aim of this study is to track cerebral blood flow (CBF) changes during anesthesia by means of rheoencephalography (REG) signals. Thus, REG signals are analyzed using a traditional approach based on the extraction of geometrical properties in the time domain as well as non-linear features extraction by Poincaré plot analysis. Those analyses are applied to different anesthesia scenarios. Moreover, interactions between depth of anesthesia monitoring, REG signal based features and other clinical variables recorded during general anesthesia procedures, such as EEG, infused drugs, heart rate and mean arterial pressure are analyzed by means of causal Granger analysis. This last step aims at detecting cause-effect relationships taking place during general anesthesia procedures, involving interactions between different physiological systems to better characterize the effect of anesthetics on brain hemodynamics.

2 Materials and methods

2.1 Data acquisition

The analyzed database is composed of 88 female patients enrolled for elective gynecological surgeries under total intravenous anesthesia (TIVA) with propofol and

TABLE 1 General anesthesia dataset.

Patients demographic data	
Age (years)	49.5 ± 16.4
Height (cm)	161.3 ± 7.0
Weight (kg)	68.1 ± 13.9
BMI (kg/m^2)	26.2 ± 5.2
Drugs administered during surgical procedures	
Propofol	88/88 (100%)
Remifentanil	88/88 (100%)
Rocuronium	43/88 (48.9%)
Atropine	16/88 (18.2%)
Ephedrine	7/88 (7.9%)
Methadone	16/88 (18.2%)

Demographic data values are presented by mean value ± standard deviation.
BMI: body mass index.

remifentanil. Exclusion criteria considered were cardiac or neurosurgeries, as well as traumatic brain injuries. Summarized demographic data, the set of drugs used for anesthesia management, control of hemodynamics and the occurrence of administration of each drug are specified in Table 1.

The initial dosage of propofol at anesthesia induction was 5.8 μg/ml (ranging from 4.8 to 7 μg/ml) and was administered together with remifentanil targeted at 3.8 ng/ml (ranging from 2 to 6.2 ng/ml). After induction, those target dosages for propofol and remifentanil were reduced to 3.4 μg/ml (from 2.5 to 4.3 μg/ml) and to 3.4 ng/ml (from 2.3 to 4.5 ng/ml), respectively. From the 88 patients suitable for analysis, 22 were intubated through laryngoscopy while in the remaining 66 a Laryngeal Mask Airway (LMA) was used. Patient positioning was also considered in this work, two different positions were assessed besides the standard supine position in steady state anesthesia: 24 patients (27.3%) were kept in the horizontal plane for the whole procedure, while passive leg raising took place in 51 cases (57.9%) and 13 participants (14.8%) were placed in Trendelenburg position (surgical position where the subject lies supine, or flat on their back, with their feet raised higher than their head).

All patients were monitored from 3 min prior to the anesthesia induction until 3 min after extubating. Patient monitoring consisted on the use of a Depth of Anesthesia device, the Conox (Fresenius Kabi, Bad Homburg, Germany) providing the electroencephalogram (EEG) signal and the qCON index that evaluates the hypnotic effects in the brain, as well the qCO (Quantium Medical, Barcelona, Spain) device, an electrical bioimpedance monitor for rheoencephalography (REG) data collection, and a Dräger (Dräger, Lübeck, Germany) hemodynamic monitor for the heart rate (HR in bpm, 1 value/s), systolic blood pressure (SBP, mmHg), diastolic blood pressure (DBP, mmHg) and mean arterial pressure (MAP in mmHg, 1 value/s). Data from the qCO monitor were continuously collected at a sampling frequency of 250 Hz and EEG from Conox with a sampling rate of 1024 Hz, and a resolution of 3 bytes in the range of ± 374 mV.

Data from those monitors, as well as data from the TCI pumps were recorded through the RugloopII software (Demed, Belgium). Moreover, annotation of relevant events during the surgical procedure was performed through the same software, to make sure the occurrence of those events was synchronized with all other clinical data.

The clinical trial followed the principles of the Declaration of Helsinki for human subjects. All participants were informed about the study and gave their written consent prior to participation.

Recorded signals were classified in 5 different categories depending on the clinical state of the patients during general anesthesia:

– Awake, corresponding to the data recorded prior to anesthesia induction.
– Loss of consciousness (LOC), data recorded right after LOC is detected and while intubation is being prepared.
– Steady state anesthesia (Anes), data recorded during anesthesia, without burst suppression episodes (EEG pattern with continuous alternation between high-voltage slow waves or even sharp waves and depressed or even suppressed electrographic activity) and after intubation and patient positioned for surgery.
– Burst suppression rate (BSR), data belonging to periods in which the Conox BSR index provides values higher than 10. The burst suppression rate (BSR), is defined as the fraction of EEG spent in suppression per epoch, is the standard quantitative measure used to characterize burst suppression.
– Recovery of consciousness (ROC), data belonging to the end of the procedure, once drug infusion has been stopped and patient is ready to be extubated.

2.2 Signal preprocessing

An automatic artefact rejection algorithm was applied to the recorded EEG signals, in order to avoid processing noisy data resulting from patient movements or the use of other devices, mainly the surgical knife. The traditional frequency band analysis (δ, θ, α, ß) was performed on EEG signals filtered between 0.1 and 50 Hz with a second-order Butterworth filter resampled at 256 Hz. Subsequently, time series were processed in moving time windows of 2 s with 1 s overlap, thus providing new results every second.

REG data were screened for artefact rejection and processed with linear filters. A high-pass filter was applied to REG signals using a fourth-order Chebyshev type II, with 0.1 Hz stop band frequency to eliminate DC fluctuations, followed by a

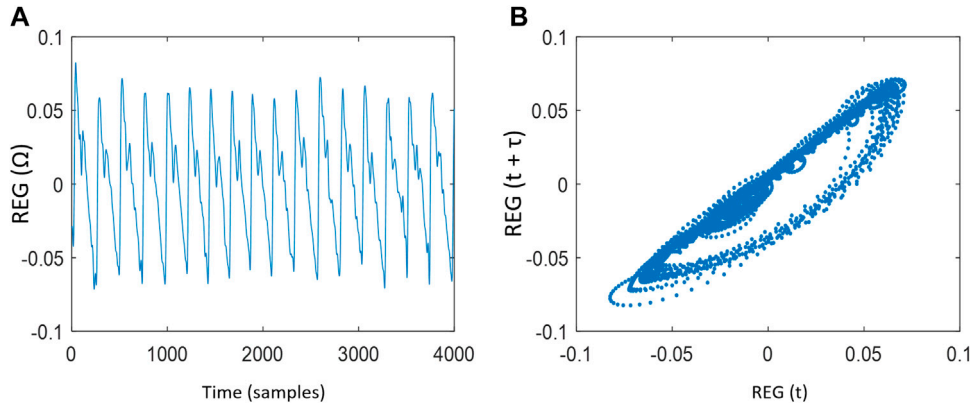


FIGURE 1
(A) Rheoencephalography (REG) signal trend and **(B)** Poincaré plot reconstruction of a REG signal with time lag τ = 5 samples.

Butterworth second-order low-pass filter with a cut-off frequency at 4 Hz. Subsequent calculations of REG data were applied to sliding windows of 8 s, resulting in a new value every second.

Finally, once time series were preprocessed, they were synchronized with all other data collected during the surgical procedures, such as hemodynamic variables, drug infusion dosages and events recorded during surgery.

2.2.1 Geometric features of rheoencephalography signals

The classical methods used to assess REG signals rely on the analysis of the geometry of the pulse waves (González et al., 2018). In this way, for REG recordings, the minimum and maximum values of each pulse wave and their respective derivatives were automatically detected, and the following features were calculated: amplitude range of the pulse (Range), time between two consecutive maximums (Δtmax, samples), time between two consecutive minimums (Δtmin, samples), time between each minimum and the following maximum (Δtmin-max, samples), the slope (α, a.u.) of the pulse during Δtmin-max interval, the area under the pulse wave (Area, Ω s), the systolic area (AreaSyst, Ω s) that is calculated as the area of a pulse wave delimited by a minimum and its consecutive maximum, the maximum derivative (δmax, $\Omega\,s^{-1}$) and the range of the derivative (δrange, $\Omega\,s^{-1}$). In addition, blood volume and blood flow estimations were also considered. The relative cerebral blood volume (CBVrel, Ω) was calculated as:

$$\mathrm{CBVrel} = \delta\mathrm{max}\,\mathrm{LVET} \tag{1}$$

where the left ventricular ejection time (LVET, ms) was considered as a function of HR_{REG} (bpm), LVET = 416–1.56 HR_{REG} (Willems et al., 1970), computing HR_{REG} from the difference between two consecutive maximums of the REG curve. The cerebral blood estimation (CBFest, $\Omega\,s^{-1}$) was calculated as:

$$\mathrm{CBFest} = \mathrm{CBVrel}\ HR_{REG}/60 \tag{2}$$

2.2.2 Poincaré plot analysis of rheoencephalography signals

Two-dimensional Poincaré plot was constructed from REG sequences, with REG(t) at *x*-axis and REG (t+τ) at *y*-axis, where t moves from 1 to N-τ samples, being N the length of the series. The choice of the time lag τ is critical, since very low values would not allow the attractor to expand, with a majority of points laying on the diagonal line (Chen at al., 2007), while very large values of τ would cause deformations of the attractor due to the fact that pairs of samples would be uncorrelated (Fraser and Swinney, 1986; King et al., 1987). Since no previous work has been done on the analysis of REG attractors during general anesthesia, a wide range of τ values was analyzed (from 1 to 20 samples) in order to provide a τ value able to give the maximum possible information related to the dynamics hidden in REG signals. Figure 1 shows a rheoencephalography (REG) signal trend and its related Poincaré plot reconstruction of a patient.

To generate quantitative information on the distribution of REG signals in the Poincaré plots, several features were extracted from the reconstructed attractor:

– SD1 and SD2 which respectively are the standard deviation (SD) of REG(t) dispersion perpendicular to the diagonal line (the identity line) and the SD of the REG(t) dispersion along the diagonal line. They are computed following Eq. 3 where var is the variance.

$$\begin{aligned} \mathrm{SD1} &= \sqrt{\mathrm{var}\left(\frac{\mathrm{REG(t)} - \mathrm{REG(t+\tau)}}{\sqrt{2}}\right)} \\ \mathrm{SD2} &= \sqrt{\mathrm{var}\left(\frac{\mathrm{REG(t)} + \mathrm{REG(t+\tau)}}{\sqrt{2}}\right)} \end{aligned} \tag{3}$$

– Area of the ellipse (SDarea), calculated as SDarea = π SD1 SD2
– Ratio SDratio, defined as SD1/SD2
– Correlation R, measured between REG(t) and REG (t+τ) signals.
– Complex correlation measure (CCM) (Karmakar et al., 2009), identifies all possible sets of three consecutive attractor points of the Poincaré plot and the area of the triangle they define is calculated (González, et al., 2018). In cases where all three points are aligned, the area is considered to be zero. CCM is computed as indicated in Eq. 4:

$$CCM(\tau) = \frac{1}{SDarea\ (N-2)} \sum_{i=1}^{N-2} \|M(i)\| \tag{4}$$

where M(i) is the matrix including the coordinates of the three points from each subset whose determinant is the area of the triangle formed by them and SDarea is the normalizing constant and represents the area of the fitted ellipse over Poincaré plot.

2.2.3 Granger causality analysis

Granger Causality (GC) has been applied to the collected signals to assess causality between pairs of time series. It relies on a hypothesis test in which the null hypothesis is that, given two time series $x_1(t)$ and $x_2(t)$, $x_2(t)$ does not cause $x_1(t)$. In order to assess the causality between the signals, autoregressive models (AR) are built, the restricted and the unrestricted model. The restricted model (univariate AR model) uses only past values from the signal $x_1(t)$ to predict its future values, while the unrestricted model (bivariate ARX model) uses past values from both $x_1(t)$ and $x_2(t)$ to predict values of $x_1(t)$. The restricted model is defined as:

$$x_1(t) = a_1 + \sum_{i=1}^{L} a_{1,i}\, x_1(t-i) + \varepsilon_{1r}(t) \tag{5}$$

$$x_2(t) = a_2 + \sum_{i=1}^{L} a_{2,i}\, x_2(t-i) + \varepsilon_{2r}(t) \tag{6}$$

while the unrestricted model are represented by:

$$x_1(t) = b_1 + \sum_{i=1}^{L} b_{1,i}\, x_1(t-i) + \sum_{i=1}^{L} c_{1,i}\, x_2(t-i) + \varepsilon_{1u}(t) \tag{7}$$

$$x_2(t) = b_2 + \sum_{i=1}^{L} b_{2,i}\, x_2(t-i) + \sum_{i=1}^{L} c_{2,i}\, x_1(t-i) + \varepsilon_{2u}(t) \tag{8}$$

where a_{ji}, b_{ji} and c_{ji} are the estimated coefficients of the models of order L, being j = {1,2} and the residuals (prediction errors) of the models are $\varepsilon_{jr}(t)$ and $\varepsilon_{ju}(t)$.

The Schwartz's Bayesian Information Criterion (BIC) (Schwartz 1978) was selected to fit the order L of the model, since it has been published to be more consistent (Zhang, 1993) and demonstrated (Nicolau and Georgiou, 2013) to provide reliable values for EEG models under general anesthesia. The optimal order L was *a priori* tested from 1 to 10 samples (i.e. 10 s).

To decide if the null hypothesis is rejected, an analysis of variance test was carried out. In this context, F-statistic is computed as:

$$F = \frac{SSR/d_n}{SSE/d_d} \tag{9}$$

where SSR is the sum of squares explained by the regression, SSE is the sum of squares errors, d_n equals the number of independent variables and the degrees of freedom of the SSE are $d_d = N - d_n - 1$. If the statistic is found significant at level p-value<0.05, the null hypothesis is rejected and causality from the time series $x_2(t)$ to $x_1(t)$ is considered to take place. Following a similar procedure, the causality from times series $x_1(t)$ to $x_2(t)$ is evaluated. The magnitude of the causality from $x_1(t)$ to $x_2(t)$ and $x_2(t)$ to $x_1(t)$ was measured respectively as function of the model error variances:

$$\begin{aligned} C_{x1\to x2} &= \ln \frac{var(\varepsilon_{2r})}{var(\varepsilon_{2u})} \\ C_{x2\to x1} &= \ln \frac{var(\varepsilon_{1r})}{var(\varepsilon_{1u})} \end{aligned} \tag{10}$$

2.2.4 Data analysis and statistical analysis

The features extracted from each constructed two-dimensional Poincaré plot on REG(t) signals were SD1, SD2, SDratio, SDarea, R and CCM. A statistical analysis was performed to select the τ value from 1 to 20 samples (based on González et al., 2018) that allows those features to statistically distinguish the clinical states of a general anesthesia (Awake, LOC, Anes, BSR, ROC).

Furthermore, a statistical analysis was applied on REG geometric features (Range, Δtmax, Δtmin, Δtmin-max, Slope α, Area, AreaSyst, δmax, δrange, CBVrel and CBFest), REG Poincaré plot descriptors (SD1, SD2, SDratio, SDarea, CCM and R), global hemodynamics (HR, MAP), the effect site concentrations of propofol and remifentanil (CePropo, CeRemi) and EEG based-parameters related to depth of anesthesia (qCON and the EEGδ, EEGθ, EEGα and EEGß energy bands) in order to statistically study their performance in discriminating between the consecutive clinical states of the patients during general anesthesia.

For each patient, all these descriptors were calculated on each time-varying 8-s sliding-window at each state and their averaged value calculated. ANOVA for repeated measures for normal distributions and Friedman test for non-normal distributions verified by the Kolmogorov–Smirnov test were applied. This analysis was followed by the post hoc non-parametric paired samples Wilcoxon test. A significant level p-value<0.01 (Bonferroni correction) was considered.

Causality analysis was applied among pairs of simultaneous feature values, calculated on each time-varying 8-s sliding-window obtained from each patient undergoing general anesthesia. All windows of all data were synchronized. The features taken into account were those REG(t) geometric descriptors (named as CBF lin), REG(t) Poincaré plot descriptors (named as CBF PP), global hemodynamics features (HR and MAP), the effect site concentrations of propofol (CePropo) and remifentanil (CeRemi) and EEG based-parameters. The coupling strength between those families of features were analyzed under different general anesthesia events:

a) Steady state anesthesia (n = 84 segments): 400s periods in which effect site concentrations of propofol and remifentanil were constant, and no surgical events took place.
b) Propofol infusion (n = 29 segments): periods from 200s before to 200s after the change of the target effect site concentration of propofol, while remifentanil was kept constant.
c) Remifentanil infusion (n = 16 segments) periods from 200s before to 200s after the change of the target effect site concentration of remifentanil, while propofol was kept constant.
d) Atropine infusion (n = 16 segments): periods from 200s before to 200s after the administration of atropine.
e) Ephedrine infusion (n = 7 segments): periods from 200s before to 200s after the administration of ephedrine.
f) Trendelenburg position (n = 12 segments): periods from 200s before to 200s after the positioning of the patient from the horizontal supine position to the Trendelenburg position.
g) Passive leg raising (n = 48 segments) periods from 200s before to 200s after the elevation of patient legs in preparation for surgery.

Among the set of clinical events in which causality was studied, the periods of steady state anesthesia were used as reference, thus the results from the other events, such as atropine infusion or Threndelenburg positioning were compared to those obtained during stable anesthesia.

Given a pair of variables $x_1(t)$ and $x_2(t)$, causality indices $C_{x1\to x2}$ and $C_{x2\to x1}$ were compared through statistical hypothesis testing. Normality of the data was assessed by means of a Kolmogorov-Smirnov test and subsequently, t-student test was applied. This analysis was followed by the post hoc non-parametric U Mann-Whitney test. Statistical significance level p-value < 0.005 was considered.

Causality diagrams are drawn for each general anesthesia event. Whenever causality indices were higher in one-way, with statistical significance, this direction of causality is considered and represented in the causality diagrams with a single arrow. The occurrence of the interactions between two groups was computed as the number of patients presenting at least one statistically significant causal relationship between any pair of features belonging to the two groups under analysis.

Moreover, for each event, Spearman correlations (ρ) between the causality indices and patient demographics were calculated and considered as confounding factors for p-value <0.01, due to the large number of correlations being analyzed simultaneously. Only correlations reaching absolute values above 0.5 were included for analysis. Those relations that presented a correlation higher than 0.5, a regression analysis based on one variable was constructed for analyzing the influence of patient demographics on causality indices.

3 Results

3.1 Estimating Poincaré plot time-lag on rheoencephalography signals

To determine the time lag τ of the Poincaré plot of REG sequences able to provide the maximum possible information related to the dynamics hidden in REG signals, a wide range of τ values is studied (τ = {1, . . ., 20} samples).

The evolution of each Poincaré plot descriptor as a function of the time lag τ for each anesthesia phase is depicted in Figure 2. SD1 increases as τ increases in all states, reaching higher values for awake and LOC, which are also characterized by a wider interquartile range. In contrast, SD2 remains stable for all τ values, providing a higher score during Awake and LOC states. Subsequently, their ratio (SDratio) increases as τ increases, with similar interquartile ranges among the various anesthesia stages, while the ellipse area (SDarea) shows higher values for Awake and LOC, with a higher dispersion in those two states as τ increases compared with dispersions of Anes, BSR and ROC states. The behavior of the correlation R decreases for increasing τ values in all anesthesia phases, and showing similar interquartile ranges across states. Finally, CCM is the only feature showing a local maximum, identified in low τ values ($\tau \leq 5$) and providing its highest values in Anes state.

All the extracted features (SD1, SD2, SDratio, SDarea, CCM and R) present differences between the targeted set of anesthetic states. The statistical significance of those differences is assessed in Figure 3, SD1 and SD2 showed the ability to differentiate between LOC and Anes (p-value <0.01) for all τ values, while they failed in reflecting differences among all other transitions between consecutive states. Regarding SDratio, significant differences were detected in both transitions Awake vs LOC and LOC vs Anes. Nonetheless, the τ range in which p-values were under the significance threshold (p-value <0.01) was reduced to the intervals 8 to 20 samples and 12 to 20 samples, respectively.

The correlation R provided a similar performance but with narrower τ ranges for significance: 10 to 20 samples for the transition between Awake and LOC and 17 to 20 samples for LOC and Anes. The ability of SDarea to distinguish between consecutive states was limited to the transition between LOC and

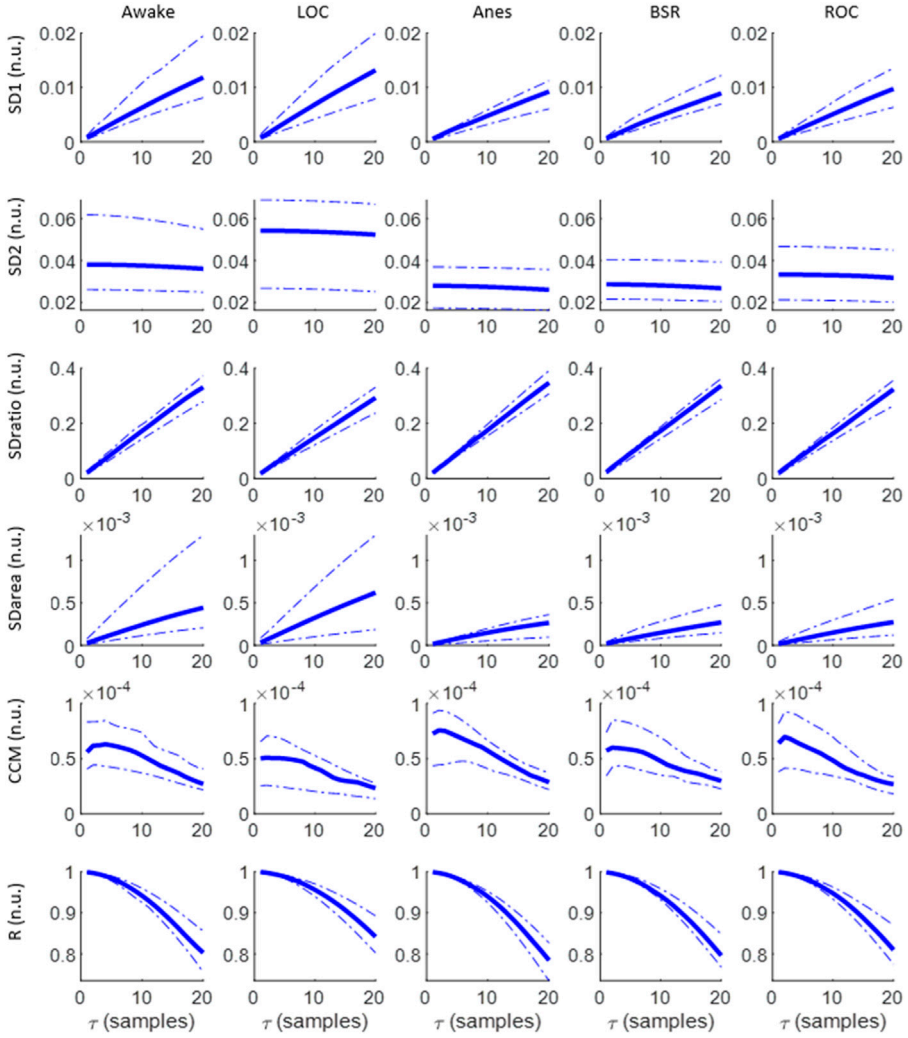


FIGURE 2
Evolution of SD1, SD2, SDratio, SDarea, CCM and R as a function of τ for the set of anesthesia states under analysis: Awake, LOC, Anes, BSR and LOC. Median values are graphed, together with the 25th and 75th quartiles represented with dashed lines.

Anes, preserving the statistical significance for all τ values tested. CCM is the only feature that does not distinguish between LOC and Anes, but it provides statistical significant results (p-value <0.01) in the transition between Anes and BSR for τ from 18 to 20 samples. Moreover, it also reflects differences between Awake and LOC states for $\tau > 10$ samples.

Considering the selection of the optimal τ values to assess differences between consecutive anesthesia states (Awake vs LOC, LOC vs Anes, Anes vs BSR and BSR vs ROC) using the set of features extracted from the Poincaré plot, low τ values have proved to fail in reflecting changes while the highest range of the tested interval showed a better performance considering all features and anesthesia states (Figure 3). Therefore, the value $\tau = 20$ samples was chosen to be appropriate to detect changes in anesthesia states.

3.2 Analysis of the behaviors of the anesthesia state descriptors

Figure 4 includes a set of data recorded from one subject participating in the clinical trial. The anesthesia induction started at t = 500s approximately, with the infusion of remifentanil and propofol (Figure 4C). A decrease in qCON (Figure 4A) took place as a consequence of the effect of the drugs, resulting in the transition from the awake state to anesthesia around t = 700s. Different events can be observed, steady state anesthesia (Figure 4A) begins right after the drug concentrations of propofol and remifentanil are lowered and stabilized at t = 1000s and lasts for 1000s (Figure 4C). Immediately afterwards, the remifentanil effect site concentration was increased, originating the new clinical event that seems to be followed by

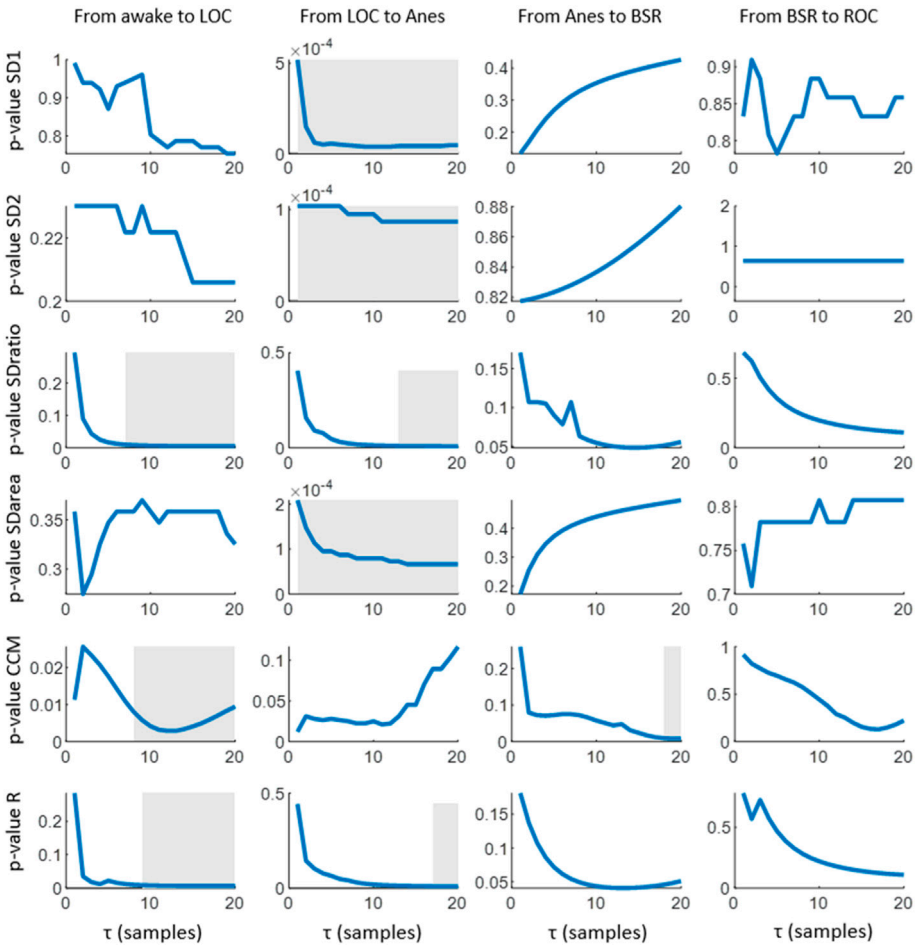


FIGURE 3
Statistical significance (*p*-values) obtained for the comparison of the median values of each Poincaré feature (SD1, SD2, SDratio, SDarea, CCM and R) among consecutive anesthesia states. The post hoc non-parametric paired samples Wilcoxon test was applied. Grey areas represent intervals in which the graphed parameter shows statistical significance of *p*-value<0.01.

EEGδ(t), EEGθ(t), EEGα(t), EEGß(t) energies (Figure 4B), HR (Figure 4D) and δmax (Figure 4E).

The mean and standard deviation values of REG geometric features calculated on each anesthesia state are depicted in Table 2. Statistical significance level obtained comparing each two consecutive states (Awake vs LOC, LOC vs Anes, Anes vs BSR and BSR vs ROC) are also indicated. Values of REG Range, δmax, δrange, CBVrel and CBFest were higher at LOC state and minima at Anes, *p*-value < 0.005. Regarding the systolic area (AreaSyst), its value increased from BSR to ROC state, *p*-value = 0.006. No statistical differences were presented between the remainder consecutive states.

Results concerning to Poincaré plot features of REG segments for τ = 20 samples are reported in Table 2. SD1 presented similar values, in average, in Awake vs LOC, decreasing during anesthesia (LOC, m ± std = 0.015 ± 0.009; Anes, m ± std = 0.009 ± 0.003; *p*-value = 0.00005) and slightly increasing for BSR and ROC, but without recovering the initial values at Awake and LOC. Descriptor SD2 had a similar performance, except for the transition between Awake and LOC, where SD1 showed similar values while SD2 increased. Since SDarea is proportional to the product of SD1 and SD2, it followed similar behavior described for those two features comparing LOC vs Anes (*p*-value = 0.00007). Regarding SDratio and R descriptors, they presented opposite trends, as expected, with similar variances for all patient states: SDratio showed an absolute minimum for LOC, while R had its maximum in this same state. Both descriptors were able to statistically differentiate Awake vs LOC (*p*-value = 0.0053 and *p*-value = 0.0065, respectively) and LOC vs Anes (*p*-value = 0.0076 and *p*-value = 0.00873, respectively), as it is indicated in Table 2. Finally, CCM presented a maximum for the Awake state and a minimum for LOC state (*p*-value = 0.0095), presenting similar values for all states except for LOC.

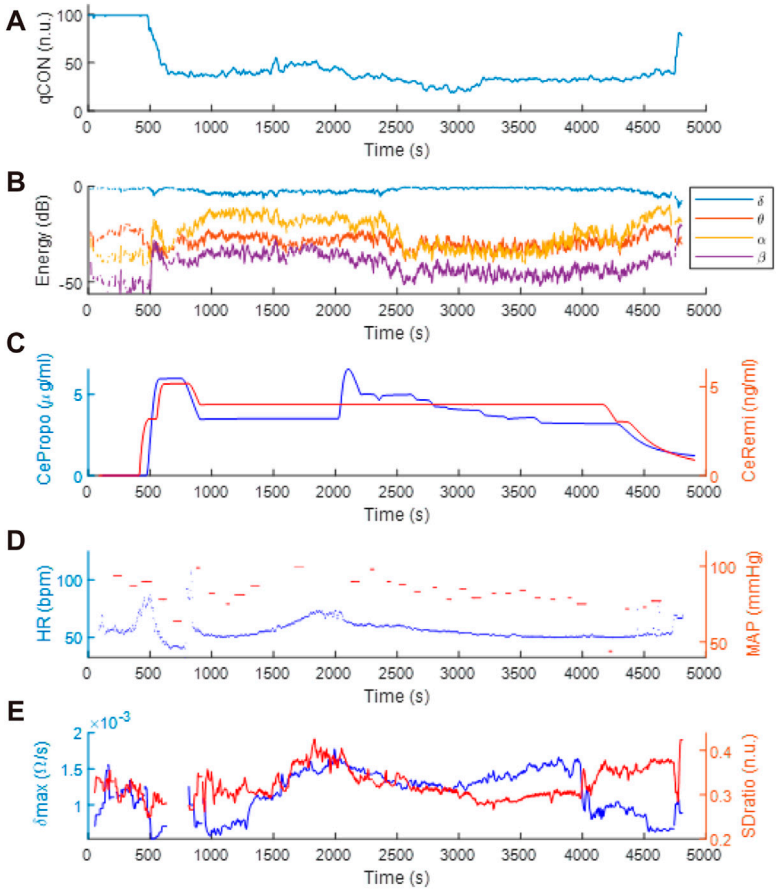


FIGURE 4
Clinical data recorded during anesthetic procedure: **(A)** qCON index, **(B)** EEG frequency bands, **(C)** propofol and remifentanil effect site concentrations (CePropo and CeRemi, respectively), **(D)** heart rate (HR) and mean arterial pressure (MAP) and **(E)** δmax and SDratio REG features.

The trends of all extracted features present changes along the anesthetic procedure for τ = 20 samples, mainly during LOC. However, only some of those changes are statistically significant. All the features except CCM were able to detect changes in the transition between LOC and Anes states. Regarding the differences between the Awake and LOC states, SDratio, CCM and R provided statistically significant results while none of the features led to positive results for the transitions among other anesthesia states.

The evolution of the EEG energy, qCON, HR and MAP across the identified anesthesia states is presented in Table 2. It is observed that qCON index statistically differentiates (p-value < 0.01) Awake to LOC, Anes to BSR and BSR to ROC state transitions, showing decreasing values from Awake to BSR but increasing at the recovery of consciousness (ROC) state (qCON (m ± std): Awake, 95.9 ± 8.04; LOC, 47.0 ± 13.6; Anes, 40.1 ± 9.4; BSR, 24.2 ± 8,7; ROC, 71.1 ± 15.0). The MAP could statistically differentiate LOC vs Anes (MAP (m ± std): LOC, 89.2 ± 20.6; Anes, 76.7 ± 15.4; p-value = 0.008), while HR was not able to differentiate any transition. Regarding to EEG frequency bands, a similar trend was observed in EEGα and EEGß with high statistical differences when comparing Awake vs LOC and BSR vs ROC. Both EEGδ and EEGθ energies only statistically differentiated BSR vs ROC states.

3.3 Causality analysis at different anesthesia events

For every general anesthesia event (steady state anesthesia, propofol infusion, remifentanil infusion, atropine infusion, ephedrine infusion, Trendelenburg position and passive leg raising) all couplings between pairs of variables (CBF lin, CBF PP, HR, MAP, EEG based-parameters, CePropo, CeRemi) are presented through Granger causality, with the aim of studying

TABLE 2 Averaged values of the rheoencephalography REG(t) signal features, electroencephalogram EEG(t) features, and the clinical variables such as heart rate and mean arterial pressure, recorded during general anesthesia. The differences between consecutive anesthetic states (Awake vs LOC, LOC vs Anes, Anes vs BSR, BSR vs ROC) are indicated by means of the statistical significance level.

Index	Awake	LOC	Anes	BSR	ROC	Awake	LOC	Anes	BSR
						LOC	Anes	BSR	ROC
Cerebral hemodynamic features									
CBF lin									
Range	0.103 ± 0.061	0.130 ± 0.094	0.062 ± 0.025	0.071 ± 0.031	0.082 ± 0.049	n.s	***	n.s	n.s
Δtmax	379.2 ± 222.2	361.4 ± 154.4	277.0 ± 93.3	272.9 ± 45.8	357.5 ± 179.9	n.s	●	n.s	n.s
Δtmin	383.0 ± 227.0	352.1 ± 150.2	275.4 ± 80.5	285.8 ± 69.4	359.0 ± 186.0	n.s	n.s	n.s	n.s
$\Delta t_{min\text{-}max}$	182.7 ± 145.0	162.7 ± 107.1	197.4 ± 74.8	114.6 ± 36.3	179.5 ± 133.0	n.s	●	n.s	n.s
Slope(α)	9E-4 ± 6E-4	11E-4 ± 9E-4	8E-4 ± 3E-4	8E-4 ± 4E-4	7E-4 ± 4E-4	n.s	●	n.s	n.s
Area	417.6 ± 257.8	393.9 ± 184.4	287.8 ± 83.2	301.0 ± 72.0	387.6 ± 210.1	n.s	●	n.s	●
AreaSyst	198.7 ± 159.8	181.5 ± 123.3	112.2 ± 77.7	120.7 ± 37.6	194.8 ± 150.1	n.s	n.s	n.s	*
δmax	19E-4 ± 12E-4	23E-4 ± 17E-4	12E-4 ± 4E-4	13E-4 ± 6E-4	14E-4 ± 8E-4	n.s	***	n.s	n.s
δrange	30E-4 ± 20E-4	33E-4 ± 22E-4	18E-4 ± 6E-4	20E-4 ± 9E-4	22E-4 ± 12E-4	n.s	***	n.s	n.s
CBVrel	1.015 ± 0.669	1.128 ± 0.822	0.593 ± 0.194	0.660 ± 0.295	0.747 ± 0.412	n.s	***	n.s	n.s
CBFest	51.91 ± 40.34	48.76 ± 31.14	33.99 ± 12.42	37.51 ± 18.86	34.87 ± 19.90	n.s	**	n.s	n.s
CBF PP									
SD1	0.015 ± 0.009	0.015 ± 0.009	0.009 ± 0.003	0.010 ± 0.004	0.011 ± 0.006	n.s	***	n.s	n.s
SD2	0.045 ± 0.027	0.058 ± 0.043	0.027 ± 0.011	0.031 ± 0.014	0.036 ± 0.022	n.s	***	n.s	n.s
SDratio	0.339 ± 0.074	0.295 ± 0.074	0.345 ± 0.054	0.329 ± 0.053	0.316 ± 0.060	*	*	n.s	n.s
SDarea	9E-4 ± 11E-4	12E-4 ± 16E-4	3E-4 ± 2E-4	4E-4 ± 3E-4	5E-4 ± 6E-4	n.s	***	n.s	n.s
CCM	3E-5 ± 2E-5	3E-5 ± 3E-5	3E-5 ± 1E-5	3E-5 ± 1E-5	3E-5 ± 2E-5	*	n.s	n.s	n.s
R	0.789 ± 0.084	0.834 ± 0.073	0.785 ± 0.059	0.802 ± 0.056	0.814 ± 0.063	*	*	n.s	n.s
Global hemodynamics									
HR	69.8 ± 13.8	64.9 ± 10.3	61.4 ± 9.6	60.1 ± 10.2	66.1 ± 14.0	n.s	n.s	n.s	n.s
MAP	99.1 ± 12.6	89.2 ± 20.6	76.7 ± 15.4	76.2 ± 16.0	81.4 ± 16.3	●	*	n.s	n.s
EEG based-parameters									
qCON	95.9 ± 8.04	47.0 ± 13.6	40.1 ± 9.4	24.2 ± 8.7	71.1 ± 15.0	***	●	***	***
EEGδ	−0.153 ± 0.095	−0.173 ± 0.100	−0.238 ± 0.152	−0.207 ± 0.117	−0.595 ± 0.298	n.s	n.s	n.s	***
EEGθ	−3.374 ± 0.627	−3.616 ± 0.616	−3.561 ± 0.618	−3.459 ± 0.670	−2.943 ± 0.540	n.s	n.s	n.s	***
EEGα	−4.387 ± 0.714	−3.360 ± 0.801	−2.954 ± 0.771	−3.245 ± 0.584	−2.278 ± 0.849	***	n.s	n.s	***
EEGß	−5.250 ± 0.905	−4.411 ± 0.654	−4.368 ± 0.640	−4.320 ± 0.634	−2.951 ± 0.800	***	n.s	n.s	***

Significant levels (p-value): n.s not significant; ● < 0.05; * < 0.01; ** < 0.005; *** < 0.0005.
Value of the descriptor is expressed by mean ± standard deviation.

the causality between different physiological systems. For this reason, causalities among pairs of REG features are not considered as well as the causal links between pairs of EEG-based parameters.

3.3.1 Steady state anesthesia

The main interactions between the physiological variables (HR, MAP, EEG based-parameters, CBF lin and CBF PP features) during steady state anesthesia are presented in

10.3389/fnetp.2022.912733

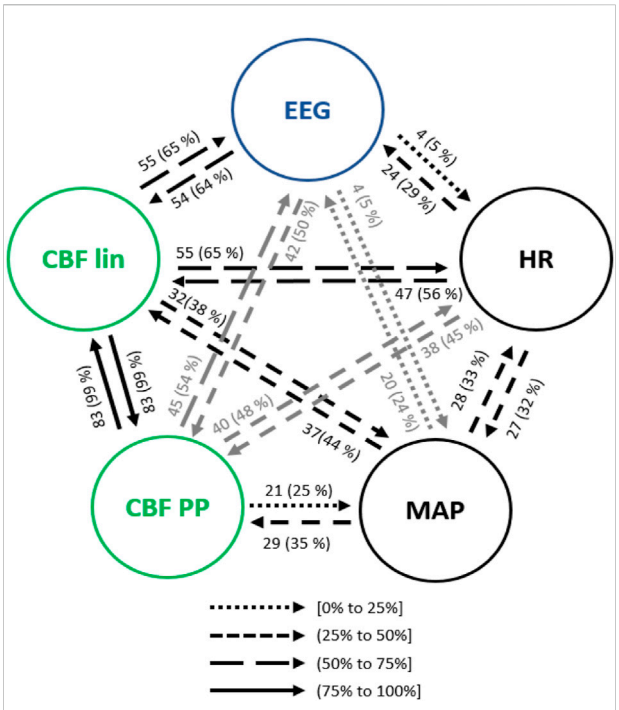


FIGURE 5
Main interactions between global hemodynamics (HR and MAP), EEG based-parameters, REG geometric features (CBF lin) and REG Poincaré plot (CBF PP) parameters during steady state anesthesia. The post hoc non-parametric U Mann-Whitney test and statistical significance level p-value < 0.005 were considered.

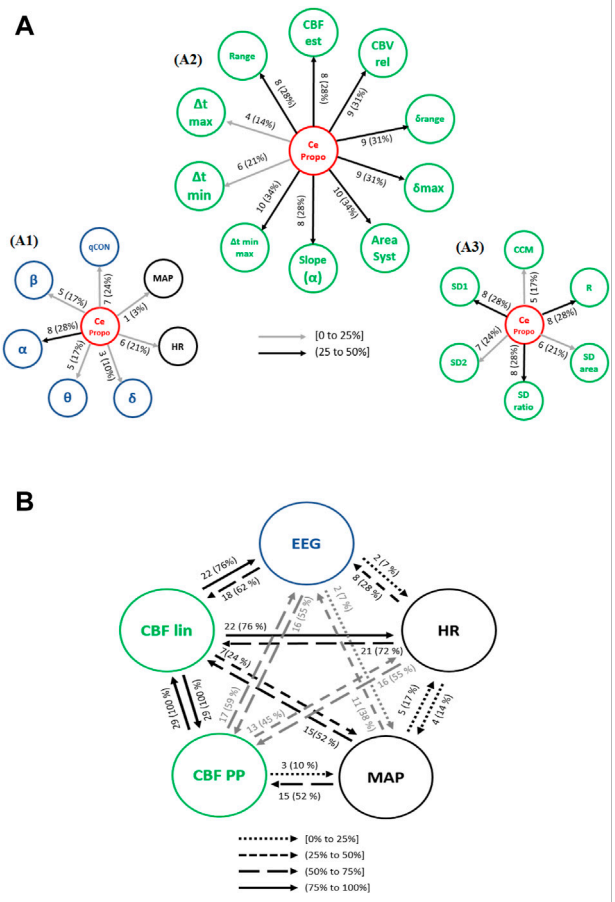


FIGURE 6
Causal interactions from **(A)** CePropo to: (A1) EEG based-parameters (δ, θ, α, ß) and global hemodynamics (HR, MAP), (A2) REG geometric features and (A3) REG Poincaré plot features. **(B)** Causal interactions between global hemodynamics, EEG based-parameters, REG geometric features (CBF lin) and REG Poincaré plot (CBF PP) parameters in propofol concentration. The post hoc non-parametric U Mann-Whitney test and statistical significance level p-value < 0.005 were considered.

Figure 5. Up to 99% of the analyzed patients showed a bilateral causal relationship between CBF lin and CBF PP parameters, since both sets of variables come from the same time series. Regarding the interactions between HR and CBF features, causalities from CBF to HR were more frequent than in the opposite direction, with the linear CBF (CBF lin) features showing a stronger role over the nonlinear ones (CBF PP). This relevance of the linear features is preserved in the causality study from and to MAP, even though in this case the causalities from MAP to the CBF features are more frequent than the opposite ones.

Both global (HR and MAP) and cerebral hemodynamics (CBF lin and CBF PP) presented causal relationships with EEG activity. EEG based-parameters had a similar occurrence of causality (5% of patients) towards HR and MAP, while HR presented a higher rate of causality towards EEG (29% of patients) than MAP (24% of patients). Regarding cerebral hemodynamics, the most relevant results rely on the 65% of causality from the CBF lin features to the EEG variables, which is one of the highest occurrences of interactions of the full system considered and therefore strongly suggests a modulation of EEG activity as a result of changes in the REG signals represented by their linear features. The Poincaré extracted features (CBF PP) showed a lower occurrence of causality (54% of patients) on EEG variables, but still higher than the ones provided by global hemodynamics HR and MAP of 29% and 24% of patients, respectively. Finally, the causality from EEG to CBF features was also higher for the linear features (CBF lin) when compared to the nonlinear parameters (CBF PP) extracted from REG signals, with 64% and 50% of patients, respectively.

No correlations were found between the causality indices and patient demographics and neither influences due to age, height, weight or BMI based on regression analysis.

3.3.2 Propofol infusion event

The propofol effect site concentration, CePropo, was added to the analysis since it is not constant in general anesthesia scenario. However,

it should only be considered as a causing variable, since it is collected from infusion pumps, resulting from the calculation of pharmacokinetic models and is not influenced by other physiological systems.

The interactions from the propofol effect site concentration (CePropo) towards all the collected physiological data are represented in Figure 6A. Causality in the opposite direction was not assessed since it does not have any clinical interpretation as previously stated. Among all the EEG bands (Figure 6A1), CePropo has the highest interaction with α (28% of patients), with similar results for its causality towards the qCON index (24% of patients). This indicates that the changes in propofol dosages are mainly affecting the α band and therefore projected in the overall depth of anesthesia assessment represented by the qCON index. The influence of CePropo in HR was detected in 21% of the patients, while causal relationships with MAP were limited to one patient. Regarding the effects of CePropo in the linear features extracted from REG signals (Figure 6A2), the causal relationships with higher occurrence were those towards Δtmin-max and AreaSyst, identified in 34% of the patients, followed by a 31% occurrence of causalities towards CBVrel, δmax and δrange. The less frequent interactions took place from CePropo to Δtmax and Δtmin. The Poincaré plot features (Figure 6A3) showed smaller occurrences, the higher ones associated to SD1, SDratio and R with 28% of patients, suggesting that CePropo is affecting the short-term variability of REG signals rather than the long-term one.

Besides the direct effects of propofol concentration changes in all the physiological variables under study, the causal relationships among hemodynamics and EEG might also be affected by the administration of the hypnotic drug. Figure 6B shows an overview of the existing causal interactions between global hemodynamics (HR, MAP), cerebral hemodynamics (CBF lin and CBF PP) and EEG related variables. Even though the detected interactions are similar to those during steady state anesthesia, several differences can be appreciated. For instance, the occurrence of causal interactions from HR and MAP towards CBF PP, CBF lin and EEG are higher, suggesting that changes in HR caused by propofol are projected in CBF and EEG. Additionally, causal effects from CBF lin to HR (76% of patients) and EEG (76% of patients) are also more frequent under propofol infusion, while the interactions between MAP and HR have a lower occurrence. Overall, changing the propofol effect site concentration elicits a higher number of interactions from both cerebral and global hemodynamics towards EEG.

The causality indices and patient demographics showed statistically significant correlations during changes in propofol effect site concentration (Table 3). Age proved to be correlated to the causality indices computed from REG features towards MAP, with correlations obtained for the REG slope (α), δrange, CBFest, SDratio and R. Among those, the linear parameters (REG slope (α), δrange and CBFest) presented increasing relations for increasing ages, while for the Poincaré based features (SDratio and R) the opposite behavior was detected. Moreover, qCON towards the REG slope (α) positively correlated with age. The influence of patient's height in the causality indices was only relevant for the causal links from SDratio and R to MAP, with taller patients related to higher values of the causality indices. In contrast, weight showed a more determinant role, patients with higher weight presented lower causality indices from CePropo to MAP, from Δtmin to MAP and from CBFest to the EEGα band. Nonetheless, the highest correlations were detected for the causality links from the EEGθ band to Δtmin and AreaSyst, with a positive correlation. Finally, BMI demonstrated to be relevant in the interactions between REG features and EEG. BMI was positively correlated with the causality relation form the EEGθ to Δtmin-max relation and AreaSyst, while it presented a negative correlation with the indices calculated from Δtmin-max

TABLE 3 Spearman correlation (ρ) between the causality indices and patient demographic.

From	To	Demographic	ρ
During changes of propofol effect site concentration			
CBFest	MAP	age↑	0.573
Slope (α)	MAP	age↑	0.523
δrange	MAP	age↑	0.548
qCON	Slope (α)	age↑	0.574
R	MAP	age↓	−0.535
SDratio	MAP	age↓	−0.553
R	MAP	height↑	0.543
SDratio	MAP	height↑	0.537
EEGθ	AreaSyst	weight↑	0.618
EEGθ	Δtmin	weight↑	0.621
CBFest	EEGα	weight↓	−0.537
Δtmin	MAP	weight↓	−0.577
EEGθ	AreaSyst	BMI↑	0.504
EEGθ	Δtmin-max	BMI↑	0.518
AreaSyst	EEGδ	BMI↓	−0.569
Δtmin-max	EEGδ	BMI↓	−0.524
During changes of remifentanil effect site concentration			
CCM	MAP	age↑	0.668
HR	EEGθ	weight↑	0.672
HR	SD1	weight↑	0.646
EEGß	δmax	weight↑	0.640
EEGα	Δtmax	weight↑	0.752
AreaSyst	HR	weight↓	−0.673
Δtmin-max	HR	weight↓	−0.707
CBVrel	qCON	weight↓	−0.698
EEGδ	R	weight↓	−0.639
EEGθ	Slope (α)	weight↓	−0.643
EEGß	δmax	BMI↑	0.637
EEGδ	R	BMI↓	−0.695
EEGδ	SDratio	BMI↓	−0.658

Increasing causality is denoted by ↑
Statistical significance p-value < 0.01.

10.3389/fnetp.2022.912733

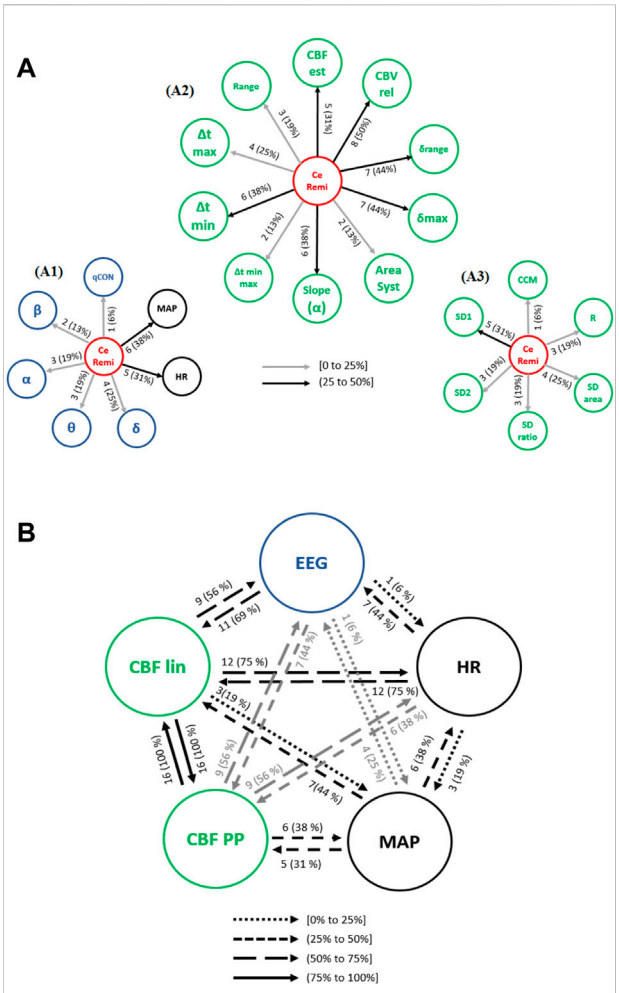


FIGURE 7
Causal interactions from **(A)** CeRemi to: (A1) EEG based-parameters (δ, θ, α, ß) and global hemodynamics (HR, MAP), (A2) REG geometric features and (A3) REG Poincaré plot features. **(B)** Causal interactions between global hemodynamics, EEG based-parameters, REG geometric features (CBF lin) and REG Poincaré plot (CBF PP) parameters in remifentanil concentration. The post hoc non-parametric U Mann-Whitney test and statistical significance level p-value < 0.005 were considered.

and AreaSyst to EEGδ. Influence of BMI was found in these causality indices on the level of adjusted R^2 = 0.7. However, no influences in causality indices due to age, height or weight based on regression analysis could be found.

3.3.3 Remifentanil infusion event

Causalities from CeRemi towards other variables should be taken into account since, as discussed for CePropo, CeRemi data are the result of the calculation of pharmacokinetic models and are not affected in any way by other physiological data, only depend on patient demographics.

Causal interactions from CeRemi towards EEG based-parameters, global hemodynamics (HR and MAP), linear features (CBF lin) and nonlinear features (CBF PP) are depicted in Figure 7A. The effects of CeRemi on EEG variables (Figure 7A1) have occurrences up to 25%, almost inexistent towards the qCON index, but slightly higher for α, θ and δ bands. However, causal relationships between CeRemi and global hemodynamics represented by HR and MAP were more frequent, reaching an incidence of 31% and 38%, respectively. Regarding the causal effects of CeRemi towards the linear features of CBF (Figure 7A2), the highest occurrences took place in the causality from CeRemi to CBVrel (up to 50%), followed by δrange and δmax (44% of patients). The weakest causality was detected towards Δtmin-max and AreaSyst, and this is one of the main differences when comparing causal effects elicited by CePropo and CeRemi. Finally, for the REG Poincaré plot features, the most frequent interaction was from CeRemi to SD1 (31% of patients), as detected as well in the CePropo analysis, suggesting that changes in remifentanil infusion did also affect short-term variability of REG signals.

The occurrence of causal interactions between HR, MAP, EEG and CBF linear and nonlinear parameters is presented in Figure 7B. When compared to steady state anesthesia, the causal effects of HR on EEG and CBF lin features are enhanced, as well as the effects of CBF PP on EEG. On the contrary, causal relationships of CBF lin features on EEG have lower occurrence. Moreover, when comparing CeRemi changes to CePropo changes, causality from HR to EEG is much more frequent under CeRemi analysis (44% of patients), while causality from MAP to EEG decreases, allowing to consider that CePropo modulates EEG changes through MAP while CeRemi influences EEG by means of HR. With respect to other significant differences, it should also be mentioned that CBF linear and nonlinear features have less frequent causal links with EEG variables, when compared to the analysis of CePropo changes. This finding is consistent with the fact that CePropo is acting at a cerebral level, reducing brain metabolism, while CeRemi has a less pronounced influence in EEG signals.

The causality indices obtained for several pairs of variables were highly correlated with patient demographics as summarized in Table 3, during changes in remifentanil effect site concentration. Age presented a positive correlation with the causality indices from CCM towards MAP, hence indicating that older patients presented higher causality indices between those two physiological parameters. Nonetheless, patient weight was the demographic variable showing more correlation value in the causal interactions detected under remifentanil dosage changes. For instance, the causality indices from Δtmin-max and AreaSyst towards HR showed a negative correlation with weight, suggesting that causality from REG to HR is enhanced in patients with lower weight (regressive analysis with adjusted R^2 = 0.6). A negative correlation was also obtained for the causal link from CBVrel to qCON, from EEGθ band to the slope of REG, and from the EEGδ band to the Poincaré descriptor R, while positive

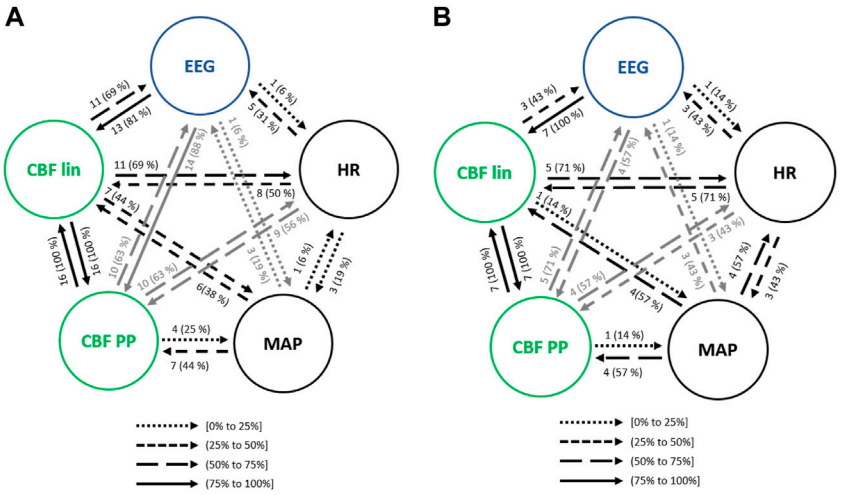


FIGURE 8
Causal interactions between global hemodynamics, EEG based-parameters, REG geometric features (CBF lin) and REG Poincaré plot (CBF PP) parameters **(A)** in atropine infusion and **(B)** in ephedrine infusion. The post hoc non-parametric U Mann-Whitney test and statistical significance level p-value < 0.005 were considered.

correlations were found for the interactions from EEGα to Δtmax, from EEGβ to δmax, from HR to SD1 and from HR to the EEGθ band. Some of those results were replicated for the BMI analysis, namely the causality from EEGβ to δmax and from EEGδ to R. Additionally, BMI presented a negative correlation from EEGδ to SDratio. Those correlations suggest that links between general hemodynamics, EEG activity and REG features under changes of remifentanil dosage are sensitive to the main characteristics of the patients being monitored, with weight being the key factor that influences causality from REG to EEG and HR to REG features, with a regression relation with adjusted $R^2 = 0.6$.

3.3.4 Atropine infusion event

Figure 8A presents the interactions between EEG parameters, HR, MAP and CBF extracted features. Causalities emerging from HR were lower towards MAP and CBF lin when compared to steady state anesthesia, but higher towards CBF PP and EEG. Regarding MAP, the causal link towards CBF PP (44% of patients) showed a higher occurrence for atropine infusion, while all other links were detected with a lower frequency. Finally, the analysis of the interactions between EEG and REG features was enhanced during the administration of atropine (CBF lin with 81% and CBF PP with 88% of patients), suggesting that this drug affects the electrical brain activity.

Several correlations between the causality indices and the demographic data of the patients were identified as significant (Table 4). Age presented a negative correlation with the causality indices from qCON to δrange and from MAP to SD1, and a

TABLE 4 Spearman correlation (ρ) between the causality indices and patient demographic.

From	To	Demographic	ρ
During atropine infusion			
SD2	EEGß	age↑	0.694
Δtmax	EEGß	age↑	0.780
Δtmin	EEGß	age↑	0.689
MAP	SD1	age↓	−0.736
qCON	δrange	age↓	−0.638
HR	EEGθ	height↑	0.664
EEGθ	SD1	height↑	0.693
EEGθ	SDarea	height↑	0.664
EEGß	δrange	height↑	0.715
EEGα	Δtmax	height↓	−0.748
CBVrel	EEGδ	weight↑	0.679
CBFest	EEGδ	BMI↑	0.723
CBVrel	EEGδ	BMI↑	0.749
SDarea	EEGδ	BMI↑	0.688
δmax	EEGδ	BMI↑	0.798
Δtmax	EEGß	BMI↑	0.692
During ephedrine infusion			
CCM	EEGα	height↓	−0.906
EEGδ	Range	weight↓	−0.955
EEGδ	Slope (α)	weight↓	−0.901
EEGδ	Range	BMI↓	−0.955
EEGδ	Slope (α)	BMI↓	−0.901

Increasing causality is denoted by ↑
Statistical significance p-value < 0.01.

positive one from the REG features Δtmax, Δtmin and SD2 towards the EEGβ band. Furthermore, regression analysis indicated that older patients present causal links from REG to EEG (R^2 = 0.6). Several correlations between height and the analyzed set of causal links were also found to be significant. For instance, the causality from the EEGα band to Δtmax had a negative correlation with height, while positive correlations were obtained from EEGβ to δrange, from EEGθ to SD1, from θ to SDarea and from HR to EEGθ. Therefore, the links between EEG and CBF features during atropine infusion seem to be dependent on patient height (also with regression analysis adjustment of R^2 = 0.6). Finally, weight was positively correlated with the causality index computed from CBVrel to the EEGδ band, as well as BMI. Additionally, BMI showed positive correlations from δmax, CBFest and SDarea towards the EEGδ band, and from Δtmax towards EEGβ. Increased BMI is hence related to enhanced causality from REG features towards electrical brain activity, with a regression analysis with adjusted R^2 = 0.6.

3.3.5 Ephedrine infusion event

As with the previous clinical scenarios analyzed, the interactions between the main sets of physiological variables were analyzed to assess the relationship between hemodynamics and brain activity (Figure 8B). One of the most relevant changes when compared to steady state anesthesia was the occurrence of the EEG causality towards CBF parameters, as well as the one from CBF lin to EEG (43% of patients) and all causal links emerging from HR and MAP, suggesting that the cardiovascular effects of ephedrine are also projected in brain activity. Some of those effects were also detected during the infusion of another vasoactive drug, atropine, even though in that case the causalities emerging from MAP and HR were in general lower, while those from EEG to CBF PP and from CBF lin to both MAP and EEG were enhanced.

Table 4 reflects the correlations between the causality indices and patient demographics. Age was not a relevant factor during ephedrine infusion. Decreasing height is correlated to increased causality between CCM and the EEGα band and also it is detected an influence to these causality indices from a regression with adjusted R^2 = 0.8. The highest correlation was detected between weight and the causality index from the EEGδ band towards the REG slope (ρ = −0.955, R^2 = 0.783), followed by the one from the EEGδ band towards the REG range (ρ = −0.901, R^2 = 0.804). Both correlations were also detected for BMI, suggesting that lower weight and BMI are associated to higher causality from EEG towards CBF features.

3.3.6 Trendelenburg positioning

The transition of anesthetized patients from a supine position to Trendelenburg was assessed for causality. Considering the interactions between hemodynamics and brain activity signals (Figure 9A), HR showed less influence in MAP when compared to steady state anesthesia, but higher causal effects on CBF features, up to 83% for the linear ones. On the contrary, MAP caused lower interactions than in steady state, except for CBF PP, which were significantly higher. Moreover, while causal links between EEG and CBF PP were enhanced during Trendelenburg positioning when compared to stable anesthesia, links between EEG and CBF lin features showed lower occurrence.

Regarding the influence of demographic characteristics of the patients in the causal relationships previously analyzed (Table 5), age showed a high negative correlation with the causality indices from HR to AreaSyst and Δtmin-max, indicating that the younger the patients the higher the causality from HR towards REG features (with regression adjusted R^2 = 0.8). The causality index from the depth of anesthesia index, qCON, towards Δtmin-max and AreaSyst was negatively correlated with height, as well as the causality index from CCM to the EEGα band and from HR to EEGβ (with a regression analysis with adjusted R^2 = 0.5), suggesting that taller patients presented weaker causal links among those pairs of variables. The role of weight was limited to two statistically significant correlations: one from δmax to qCON, presenting higher causality in patients with less weight, and a second one from MAP to SD2, in which taller patients had higher causality index associated (with a regression analysis with adjusted R^2 = 0.5). Finally, lower BMI was associated to an enhanced causality from several REG features (δmax, δrange and CBVrel) to the qCON index while higher BMI resulted in a stronger causality (with regression analysis with adjusted R^2 = 0.8) from EEGδ to Range, from MAP to SD2 and from SD1 to EEGθ.

3.3.7 Passive leg raise

Interactions among the physiological systems under study is presented in Figure 9B. Besides the bidirectional link between linear and nonlinear CBF features, the most frequent causality during passive leg raising takes place from CBF linear parameters towards EEG (77% of patients), suggesting that changes in cerebral hemodynamics are projected in brain activity. When compared to steady state anesthesia, higher causalities are detected, mainly from HR to CBF PP and EEG, from MAP to CBF features and, bilaterally, between CBF features and EEG. Additionally, causality from EEG to CBF PP is increased during patient positioning.

Since both Trendelenburg and passive leg raise provoke hemodynamic changes, it is worth comparing the causal interactions between both situations. Causalities emerging from MAP have higher occurrence under passive leg raising, as well as the interactions from CBF features to brain activity variables, and from EEG to CBF lin. However, causality from EEG to CBF PP is decreased, as well as from HR to CBF parameters. Furthermore, no statistically significant correlations were found between the causality indices and patient demographics, suggesting that the detected interactions were not dependent on patient characteristics.

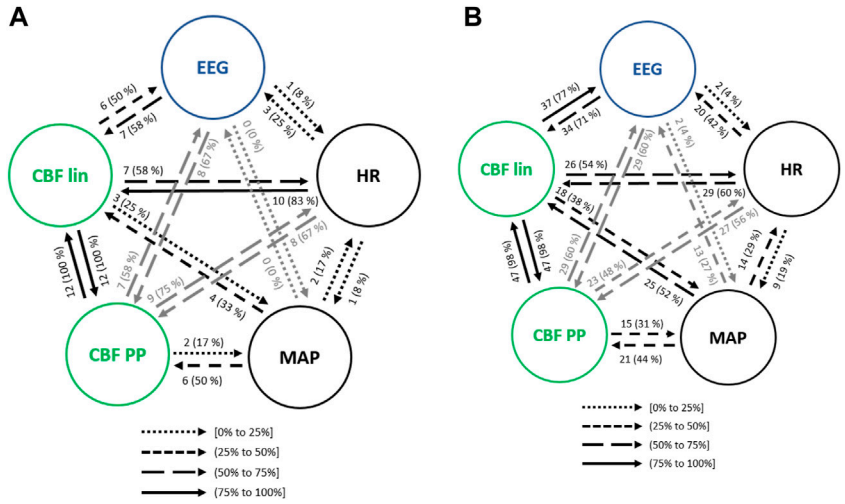


FIGURE 9
Causal interactions between global hemodynamics, EEG based-parameters, REG geometric features (CBF lin) and REG Poincaré plot (CBF PP) parameters **(A)** during Trendelenburg positioning and **(B)** during passive leg raising. The post hoc non-parametric U Mann-Whitney test and statistical significance level p-value < 0.005 were considered.

TABLE 5 Spearman correlation (ρ) between the causality indices and patient demographic during Trendelenburg positioning.

From	To	Demographic	ρ
HR	AreaSyst	age↓	−0.894
HR	Δtmin-max	age↓	−0.866
CCM	EEGα	height↓	−0.732
qCON	AreaSyst	height↓	−0.789
HR	EEGß	height↓	−0.732
qCON	Δtmin-max	height↓	−0.789
MAP	SD2	weight↑	0.872
δmax	qCON	weight↓	−0.746
SD1	EEGθ	BMI↑	0.769
EEGδ	Range	BMI↑	0.734
MAP	SD2	BMI↑	0.782
CBVrel	qCON	BMI↓	−0.769
δmax	qCON	BMI↓	−0.825
δrange	qCON	BMI↓	−0.769

Increasing causality is denoted by ↑
Statistical significance p-value < 0.01.

4 Discussion and conclusions

The geometric features extracted from REG waves collected during general anesthesia have provided statistically significant results. Several geometric features were able to detect differences between LOC and Anes states: Range, δmax, δrange, CBVrel and CBFest. The evolution of those values suggests a generally decreased CBF and instantaneous blood flow velocity during anesthesia, as previously reported by (Conti et al., 2006; Fodale et al., 2007). CBFest and CBVrel decreased during general anesthesia. This phenomenon has been related with the vasoconstriction associated to the propofol administration (Rasmussen et al., 2010). It should be noted that values for those two features are not recovered after extubation. This is probably caused by the effects of propofol in hemodynamics, since at the time of extubation it has not been eliminated from the body (Morgan et al., 1990). The reduction of CBF and related parameters in the anesthetic state might seem inconsistent with the slight increases detected during BSR. Intuitively, the lower the anesthetic depth, the lower the brain metabolism is, CBV and CBF. However, it has been proved in rats that hemodynamic fluctuations at the brain level occur during general anesthesia: cortical electrical activity is accompanied by oscillations in cerebral hemodynamics (Liu et al., 2010). This might explain the small and non-significant increase of the CBF related parameters during BSR.

The Poincaré plot features were computed for a range of τ values from 1 to 20 samples, after the preprocessing stage of the general anesthesia dataset. Statistical differences were found between Awake-LOC and LOC-Anes transitions, with a wide range of parameters showing statistically significant differences: SDratio, CCM and R in the Awake-LOC transition and all features but CCM in the LOC-Anes transition, however CCM presented statistical differences between Anes and BSR. Within the 1 to 20 samples interval of τ tested, the upper values concentrated the highest amount of statistically significant differences among anesthetic states. It is therefore stated that a value of τ = 20 samples (0.08 s) is the most appropriate one for

the analysis of REG signals during anesthesia. The performance of those parameters is dependent on the time lag τ used to reconstruct the signal attractor. Considering all results previously discussed from the information extracted of the REG signal features under anesthesia, REG analysis might be able to reflect CBF changes in REG waves. Concerning to global hemodynamics and EEG related parameters, qCON index differentiated across the different anesthesia states, but MAP just from LOC to Anes (p-value<0.01). Related to energies on the frontal EEG frequency bands, statistical differences were found between Awake vs LOC and BSR vs ROC. Similar results were found by (Sanjari et al., 2021) that applied transfer entropy on EEG signal for depth of anesthesia estimation, obtaining statistical differences between awake vs unconscious and unconscious vs. recovery EEG frequency bands.

Causal interaction analysis was applied to interpret how EEG, general hemodynamics and CBF evolve during general anesthesia under propofol and remifentanil. These interactions have been studied during steady state anesthesia, as well as during certain events occurring during surgery, such as anesthetic concentration changes, the administration of vasoactive drugs and patient positioning.

Even though literature on causal interactions involving REG signals is not available, several studies on brain activity and general hemodynamics describing heart-brain interactions have been published during both natural sleep and anesthesia. For instance (Faes et al., 2014), analyzed causal relationships among HRV and EEG during a full night sleep of subjects, observing a strong link between nonlinear beat-to-beat analysis and the power spectrum of the EEGδ band. Analogously, this EEGδ band was the one found in this study having a more frequent coupling with CBF measurements. Moreover, other studies do also support this link between hemodynamics and brain networks (Jurysta et al., 2003; Jurysta et al., 2006).

A brain-heart causality study during propofol anesthesia was published by (Won et al., 2019), concluding that causalities increased with depth of anesthesia and were stronger in the brain-heart direction than from the heart to the brain. Results obtained for the analysis herein presented suggest in fact that the most frequent interactions took place from cerebral hemodynamics to the EEG spectral densities (rather than in the opposite direction), that HR and MAP had closed loop relationships with cerebral hemodynamics and the depth of anesthesia index presented bilateral causal links with cerebral hemodynamics. Even though a larger physiological system was considered in this work, the obtained results are not consistent with those presented by (Won et al., 2019). Some differences exist in the study design, mainly based on the lowest age of the patients enrolled in Won's study, with a majority of males and receiving midazolam drug. Further data should be collected under the same circumstances to figure out the root cause of the differences between both studies, since patient demographics have shown to play an important role both in the occurrence of causality and its strength.

Overall, the analysis of causal interactions during steady state anesthesia showed that hemodynamics and EEG activity are closely linked, often under closed loop interactions, and even though there is no consensus on the direction and strength of those links, their existence has been published by several research groups and has turned neurocardiology into a relevant topic under analysis (Chen et al., 2017; Scherbakov and Doehner, 2018). Besides the study of causal interactions among heart and brain hemodynamics and EEG activity during stable anesthesia periods, changes in the concentration of propofol were also analyzed to assess its influence in these causal links. The effects of propofol on hemodynamics are well-known, characterized by a MAP and HR depression, cerebral vasoconstriction and reduced CBF while preserving autoregulation (Dagal and Lam. 2009). However, until now no information exists regarding the causality between hemodynamics and EEG during its infusion. The study performed in our work reveals that during a change in propofol dosage, the number of interactions between hemodynamics and brain activity increases. Changes in HR and MAP provoked changes in CBF and EEG, with CBF linear and nonlinear features causing EEG modulation. Moreover, one of the strongest links was found between the propofol effect site concentration and the EEGα band, which is consistent with the fact that propofol provokes a shift of the EEG energy towards this band (Schwilden et al., 1989). Additionally, a causal link between the propofol concentration and the qCON index was detected, as expected, since changes in hypnotic dosages should translate into changes in depth of anesthesia.

Several propofol pharmacokinetic-pharmacodynamic models are used in routine clinical practice for induction and maintenance of propofol anesthesia. The parameters used in those models are exclusively based on patient demographics but do not take into account hemodynamics. Given the causal relationships between brain activity and hemodynamics, the inclusion of HR, MAP or CBF data in the models would probably make them more patient-specific and improve their accuracies. Several studies have been published on this topic. Sahinovic et al. (Sahinovic et al., 2017) raised a concern on the use of propofol models in patients with brain tumours, since those might alter propofol kinetics and dynamics and loose accuracy. Furthermore, a new set of models called Physiologically-Based PK Models (PBPK) have been developed to account for the effects of hemodynamics in the currently used compartments models (Jones and Rowland, 2013), since hemodynamic variables such as cardiac output have shown to be determinant for predicting the effects of propofol infusion (Adachi et al., 2013). The results for propofol infusion in this work support the hypothesis that those new models should be key for anesthesia personalization, thus enhancing the accuracy of target-controlled infusion (TCI).

Similar conclusions can be drawn from the analysis of remifentanil concentration changes. The use of remifentanil is associated to depressed hemodynamics, preserving cerebral autoregulation but lowering CBF (Dagal and Lam. 2009), and its administration together with propofol is known to produce some

synergies in the modulation of the EEG waves and resulting depth of anesthesia index (Copot et al., 2015). The causal interactions between CeRemi and EEG related parameters revealed that the highest causality took place from CeRemi to EEGδ frequency band, followed by EEGα, EEGθ and EEGβ, but was almost inexistent with the qCON index. Those results are consistent with the EEG spectral analysis under remifentanil infusion published by Kortelainen et al. (Kortelainen et al., 2009), that highlighted the influence of remifentanil in the EEG spectrogram rather than to limit its effects to synergies with propofol. Hence, remifentanil modified the spectral content of the EEG of the patients under study while the qCON index remained unaffected. The causal relationships detected during CeRemi changes suggest that its causal effects in EEG, either directly or through hemodynamics, are less pronounced than those obtained for propofol, which is consistent with the fact that propofol is a hypnotic drug while remifentanil is an analgesic. Moreover, HR seems to be the link between CeRemi infusion and brain activity, while MAP played a more relevant role in propofol infusion. Together with the effects of propofol and remifentanil in the EEG activity, the causal relationships induced by vasoactive drugs such as atropine and ephedrine were also studied in order to find out to which extent those drugs could affect brain activity and depth of anesthesia. Both drugs are often administered during anesthesia to compensate bradycardia and/or hypotension provoked by hypnotics and analgesics, and are therefore producing HR and MAP increases to achieve hemodynamic stability.

Furthermore, the presented results provided information supporting the hypothesis that effects of atropine and ephedrine in EEG activity take place through the causal links between MAP, HR and CBF features towards EEG parameters, and vice versa. In a recently published case study (Jo et al., 2018), atropine was administered to a patient presenting very low depth of anesthesia values, including EEG suppression and a bradycardia episode. After the atropine infusion, hemodynamic stability was recovered together with recommended depth of anesthesia values. The authors related this episode to cerebral hypoperfusion, therefore suggesting that causal interactions exist between hemodynamics and brain activity, and that those are modulated through CBF.

Patient positioning was also considered as a potential factor influencing causal relationships between hemodynamics and EEG activity. Two different positions were assessed besides the standard supine position in steady state anesthesia: Trendelenburg and passive leg raising. Both positioning strategies are known to provoke changes in general hemodynamics, mainly in MAP (Fakhari et al., 2018), but information on their influence in EEG is scarce. Mallick et al. reported the dependence between the depth of anesthesia index and the steepness of the Trendelenburg position, establishing a relationship between both variables (Mallick et al., 2015). Considering the causal occurrences calculated in this work, HR and MAP do not seem to modulate directly EEG changes, but through alterations in CBF features that are further projected into EEG activity.

As part of the causality analysis, the role of patient demographics was assessed through correlation and hypothesis testing. Patient characteristics such as age, height, weight and BMI should be taken into account since those might enhance or prevent the existence of causal relationships and the intensity of the existing causal effects. For instance, during steady state anesthesia, lower ages were associated to a higher occurrence of causal links from CBF to EEG, as well as lower weight and BMI. However, the size of the database under study impaired a more detailed analysis of patient demographics in heart-brain links during anesthesia, being one of the limitations of this study. Other limitations that should be noted are the low number of recordings for some of the analyzed events, such as atropine or ephedrine infusions, and the concomitant effects of different factors, as for instance patient positioning taking place before or after a drug dosage change or the administration of a vasoactive drug.

During the causality study, linear and nonlinear CBF features have been independently considered in order to assess their individual performance. Even though they showed a 100% of causal effects among them in the majority of events under test, they revealed different occurrences of causal relationships with brain activity and global hemodynamics. During propofol infusion, bilateral causality between linear CBF features and HR was much more frequent than between Poincaré plot features of REG signals and HR, while those showed similar values during steady state anesthesia. In contrast, during atropine infusion, effects of MAP on CBF parameters were more frequent towards the Poincaré features. The use of a larger dataset would allow to further compare the performance of both algorithms, but results herein presented suggest that they are closely related to each other but the integration of the information contained in both sets of features improves the assessment of causality.

As a conclusion, results from this study confirm the hypothesis that during general anesthesia causal interactions among global hemodynamics, cerebral hemodynamics and EEG neural activity take place. And, as a consequence, clinical decisions made to achieve hemodynamical stability have effects at a neural level, as well as changes in anesthetic dosages would interfere both in global and brain hemodynamics. REG signals provided an assessment of brain hemodynamics, with both linear and nonlinear features contributing to the heart-brain interactions, revealing its potential as a monitoring tool for anesthesia management. Finally, CBF estimators demonstrated to contain information allowing to understand the coupling between hemodynamics and neural activity, and should therefore be integrated in routine clinical care, mainly in patients in which causal relationships might be impaired or altered due to pathological or intrinsic conditions.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The authors have not included the de-identified set of the data used in this work because of data restriction policies imposed by the Ethical Committee. Requests to access these datasets should be directed to cgz@quantiummedical.com.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of Hospital CLINIC de Barcelona. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Conceptualization: CG, EJ, MV-F. Data curation: CG. Formal analysis: CG, GG-H, MV-F. Investigation: CG, GG-H, MV-F. Methodology: CG, MV-F. Project administration: CG. Software: CG, GG-H. Supervision: EJ, MV-F. Writing—original draft: CG. Writing—review and edit: CG, GG-H, and MV-F.

Funding

This study was partly supported by the framework of the State Plan for Scientific and Technical Research and Innovation (DPI 2017-89827-R) and by the Biomedical Research Networking Center in Bioengineering, Biomaterials, and Nanomedicine (CIBER-BBN), an initiative of the Carlos III Health Institute (ISCIII). This work was also developed under the scope of the Industrial PhD program (2020 DI 102) by the Regional Catalan Government in collaboration with Quantium Medical S.L.U.

Conflict of interest

Authors CG, GG-H and EWJ were employed by Quantium Medical.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Adachi, Y. U., Satomoto, M., Higuchi, H., and Watanabe, K. (2013). The determinants of propofol induction time in anesthesia. *Korean J. Anesthesiol.* 65 (2), 121–126. doi:10.4097/kjae.2013.65.2.121

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 (1), e29072. doi:10.1371/journal.pone.0029072

Bodo, M., Pearce, F. J., and Armonda, R. A. (2004). Cerebrovascular reactivity: Rat studies in rheoencephalography. *Physiol. Meas.* 25 (6), 1371–1384. doi:10.1088/0967-3334/25/6/003

Bodo, M. (2010). Studies in rheoencephalography (REG). *J. Electr. Bioimpedance* 1 (1), 18–40. doi:10.5617/jeb.109

Chen, W., Wang, Z., Xie, H., and Yu, W. (2007). Characterization of surface EMG signal based on fuzzy entropy. *IEEE Trans. neural Syst. Rehabil. Eng.* 15 (2), 266–272. doi:10.1109/TNSRE.2007.897025

Chen, Z., Venkat, P., Seyfried, D., Chopp, M., Yan, T., Chen, J., et al. (2017). Brain--heart interaction: Cardiac complications after stroke. *Circ. Res.* 121 (4), 451–468. doi:10.1161/CIRCRESAHA.117.311170

Coben, R., and Mohammad-Rezazadeh, I. (2015). Neural connectivity in epilepsy as measured by Granger causality. *Front. Hum. Neurosci.* 9, 194. doi:10.3389/fnhum.2015.00194

Conti, A., Iacopino, D. G., Fodale, V., Micalizzi, S., Penna, O., and Santamaria, L. B. (2006). Cerebral haemodynamic changes during propofol-remifentanil or sevoflurane anaesthesia: Transcranial Doppler study under bispectral index monitoring. *Br. J. Anaesth.* 97 (3), 333–339. doi:10.1093/bja/ael169

Copot, D., De Keyser, R., and Ionescu, C. (2015). Drug interaction between propofol and remifentanil in individualised drug delivery systems. *IFAC-PapersOnLine* 48 (20), 64–69. doi:10.1016/j.ifacol.2015.10.116

Dagal, A., and Lam, A. M. (2009). Cerebral autoregulation and anesthesia. *Curr. Opin. Anaesthesiol.* 22, 547–552. doi:10.1097/ACO.0b013e32833020be

Dimitriev, D. A., Saperova, E. V., and Dimitriev, A. D. (2016). State anxiety and nonlinear dynamics of heart rate variability in students. *PLoS One* 11 (1), e0146131. doi:10.1371/journal.pone.0146131

Duggento, A., Bianciardi, M., Passamonti, L., Wald, L. L., Guerrisi, M., Barbieri, R., et al. (2016). Globally conditioned granger causality in brain--brain and brain--heart interactions: A combined heart rate variability/ultra-high-field (7 T) functional magnetic resonance imaging study. *Philos. Trans. A Math. Phys. Eng. Sci.* 374 (2067), 20150185. doi:10.1098/rsta.2015.0185

Fabozzi, F. J., Focardi, S. M., Rachev, S. T., and Arshanapalli, B. G. (2014). *The basics of financial econometrics: Tools, concepts, and asset management applications*. John Wiley and Sons.

Faes, L., Marinazzo, D., Jurysta, F., and Nollo, G. (2015). Linear and non-linear brain--heart and brain--brain interactions during sleep. *Physiol. Meas.* 36 (4), 683–698. doi:10.1088/0967-3334/36/4/683

Faes, L., Nollo, G., Jurysta, F., and Marinazzo, D. (2014). Information dynamics of brain–heart physiological networks during sleep. *New J. Phys.* 16 (10), 105005. doi:10.1088/1367-2630/16/10/105005

Faes, L., Nollo, G., and Porta, A. (2011). Information domain approach to the investigation of cardio-vascular, cardio-pulmonary, and vasculo-pulmonary causal couplings. *Front. Physiol.* 2, 80. doi:10.3389/fphys.2011.00080

Fakhari, S., Eissa Bilehjani, H. F., Pourfathi, H., and Chalabianlou, M. (2018). The effect of passive leg-raising maneuver on hemodynamic stability during anesthesia induction for adult cardiac surgery. *Integr. Blood Press. Control* 11, 57–63. doi:10.2147/IBPC.S126514

Fodale, V., Schifilliti, D., Conti, A., Lucanto, T., Pino, G., and Santamaria, L. B. (2007). Transcranial Doppler and anesthetics. *Acta Anaesthesiol. Scand.* 51 (7), 839–847. doi:10.1111/j.1399-6576.2007.01355.x

Fraser, A. M., and Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Phys. Rev. A Gen. Phys.* 33 (2), 1134–1140. doi:10.1103/physreva.33.1134

González, C., Jensen, E. W., Gambús, P. L., and Vallverdú, M. (2018). Poincaré plot analysis of cerebral blood flow signals: Feature extraction and classification methods for apnea detection. *PLoS ONE* 13 (12), e0208642. doi:10.1371/journal.pone.0208642

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424. doi:10.2307/1912791

Greco, A., Faes, L., Catrambone, V., Barbieri, R., Scilingo, E. P., Valenza, G., et al. (2019). Lateralization of directional brain-heart information transfer during visual

emotional elicitation. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 317, R25-R38. doi:10.1152/ajpregu.00151.2018

Hayashi, K., Mukai, N., and Sawa, T. (2015). Poincaré analysis of the electroencephalogram during sevoflurane anesthesia. *Clin. Neurophysiol.* 126 (2), 404–411. doi:10.1016/j.clinph.2014.04.019

Hayashi, K., Yamada, T., and Sawa, T. (2015). Comparative study of Poincaré plot analysis using short electroencephalogram signals during anaesthesia with spectral edge frequency 95 and bispectral index. *Anaesthesia* 70 (3), 310–317. doi:10.1111/anae.12885

Hornero, R., Aboy, M., Abasolo, D., McNames, J., Wakeland, W., Goldstein, B., et al. (2006). Complex analysis of intracranial hypertension using approximate entropy. *Crit. Care Med.* 34 (1), 87–95. doi:10.1097/01.ccm.0000190426.44782.f0

Hoshi, R. A., Pastre, C. M., Vanderlei, L. C. M., and Godoy, M. F. (2013). Poincaré plot indexes of heart rate variability: Relationships with other nonlinear variables. *Auton. Neurosci.* 177 (2), 271–274. doi:10.1016/j.autneu.2013.05.004

Jo, Y., Kim, J. M., Jeon, S. B., Park, S. U., Kam, H. J., Shim, W. H., et al. (2018). Sudden bispectral index reduction and suppression ratio increase associated with bradycardia in a patient undergoing breast conserving surgery. *J. Neurocrit. Care* 11 (1), 39–42. doi:10.18700/jnc.180041

Jones, H. M., and Rowland-Yeo, K. (2013). Basic concepts in physiologically based pharmacokinetic modeling in drug discovery and development. *CPT. Pharmacometrics Syst. Pharmacol.* 2 (8), e63. doi:10.1038/psp.2013.41

Juan-Cruz, C., Gómez, C., Poza, J., Fernández, A., and Hornero, R. (2017). "Assessment of effective connectivity in Alzheimer's disease using Granger causality," in *Converging clinical and engineering research on neurorehabilitation ii* (Springer), 763–767.

Jurysta, F., Lanquart, J. P., Van De Borne, P., Migeotte, P. F., Dumont, M., Degaute, J. P., et al. (2006). The link between cardiac autonomic activity and sleep delta power is altered in men with sleep apnea-hypopnea syndrome. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 291 (4), R1165–R1171. doi:10.1152/ajpregu.00787.2005

Jurysta, F., Van De Borne, P., Migeotte, P. F., Dumont, M., Lanquart, J. P., Degaute, J. P., et al. (2003). A study of the dynamic interactions between sleep EEG and heart rate variability in healthy young men. *Clin. Neurophysiol.* 114 (11), 2146–2155. doi:10.1016/s1388-2457(03)00215-3

Karmakar, C. K., Khandoker, A. H., Gubbi, J., and Palaniswami, M. (2009). Complex correlation measure: A novel descriptor for Poincaré plot. *Biomed. Eng. Online* 8, 17. doi:10.1186/1475-925X-8-17

Karmakar, C. K., Khandoker, A. H., Voss, A., and Palaniswami, M. (2011). Sensitivity of temporal heart rate variability in Poincaré plot to changes in parasympathetic nervous system activity. *Biomed. Eng. Online* 10 (1), 17. doi:10.1186/1475-925X-10-17

Khandoker, A. H., Karmakar, C., Brennan, M., Palaniswami, M., and Voss, A. (2013). *Poincaré plot methods for heart rate variability analysis*. Springer.

King, G. P., Jones, R., and Broomhead, D. S. (1987). Phase portraits from a time series: A singular system approach. *Nucl. Phys. B - Proc. Suppl.* (2), 379–390. doi:10.1016/0920-5632(87)90029-6

Kortelainen, J., Koskinen, M., Mustola, S., and Seppänen, T. (2009). Effects of remifentanil on the spectrum and quantitative parameters of electroencephalogram in propofol anesthesia. *Anesthesiology* 111 (3), 574–583. doi:10.1097/ALN.0b013e3181af633c

Liu, X., Zhu, X.-H., Zhang, Y., and Chen, W. (2010). Neural origin of spontaneous hemodynamic fluctuations in rats under burst--suppression anesthesia condition. *Cereb. cortex* 21 (2), 374–384. doi:10.1093/cercor/bhq105

Lu, C. W., Czosnyka, M., Shieh, J. S., Smielewska, A., Pickard, J. D., Smielewski, P., et al. (2012). Complexity of intracranial pressure correlates with outcome after traumatic brain injury. *Brain* 135 (8), 2399–2408. doi:10.1093/brain/aws155

Mallick, S., Das, A., Dutta, S., Chattopadhyay, S., Das, T., Banu, R., et al. (2015). A Prospective, double-blinded randomized controlled study comparing two different Trendelenburg tilts in laparoscopically assisted vaginal hysterectomy positioning. *J. Nat. Sci. Biol. Med.* 6 (1), 153–158. doi:10.4103/0976-9668.149115

Montgomery, L. D., Montgomery, R. W., and Guisado, R. (1995). Rheoencephalographic and electroencephalographic measures of cognitive workload: Analytical procedures. *Biol. Psychol.* 40 (1–2), 143–159. doi:10.1016/0301-0511(95)05117-1

Morgan, D. J., Campbell, G. A., and Crankshaw, D. P. (1990). Pharmacokinetics of propofol when given by intravenous infusion. *Br. J. Clin. Pharmacol.* 30 (1), 144–148. doi:10.1111/j.1365-2125.1990.tb03755.x

Moskalenko, Y. E., and Andreeva, J. (2015). Rheoencephalography: Past popularity, obvilion at present and optimistic furure. *Int. J. Adv. Life Sci. Technol.* 2 (1), 1–15. doi:10.18488/JOURNAL.72/2015.2.1/72.1.1.15

Nicolaou, N., and Georgiou, J. (2013). "Autoregressive model order estimation criteria for monitoring awareness during anaesthesia," in *IFIP international conference on artificial intelligence applications and innovations*, 71–80.

Nicolaou, N., and Georgiou, J. (2014). Neural network--based classification of anesthesia/awareness using granger causality features. *Clin. EEG Neurosci.* 45 (2), 77–88. doi:10.1177/1550059413486271

Nicolaou, N., Hourris, S., Alexandrou, P., and Georgiou, J. (2012). EEG-based automatic classification of 'awake'versus 'anesthetized'state in general anesthesia using Granger causality. *PLoS One* 7 (3), e33869. doi:10.1371/journal.pone.0033869

Porta, A., Castiglioni, P., Di Rienzo, M., Bassani, T., Bari, V., Faes, L., et al. (2013). Cardiovascular control and time domain granger causality: Insights from selective autonomic blockade. *Philos. Trans. A Math. Phys. Eng. Sci.* 371, 20120161. doi:10.1098/rsta.2012.0161

Porta, A., Faes, L., Bari, V., Marchi, A., Bassani, T., Nollo, G., et al. (2014). Effect of age on complexity and causality of the cardiovascular control: Comparison between model-based and model-free approaches. *PLoS One* 9 (2), e89463. doi:10.1371/journal.pone.0089463

Porta, A., Gelpi, F., Bari, V., Cairo, B., De Maria, B., Panzetti, C. M., et al. (2022). Monitoring the evolution of asynchrony between mean arterial pressure and mean cerebral blood flow via cross-entropy methods. *Entropy* 24 (1), 80. doi:10.3390/e24010080

Rasmussen, M., Juul, N., Christensen, S. M., Jónsdóttir, K. Y., Gyldensted, C., Vestergaard-Poulsen, P., et al. (2010). Cerebral blood flow, blood volume, and mean transit time responses to propofol and indomethacin in peritumor and contralateral brain RegionsPerioperative perfusion-weighted magnetic resonance imaging in patients with brain tumors. *Anesthesiology* 112 (1), 50–56. doi:10.1097/ALN.0b013e3181c38bd3

Riedl, M., Suhrbier, A., Stepan, H., Kurths, J., and Wessel, N. (2010). Short-term couplings of the cardiovascular system in pregnant women suffering from pre-eclampsia. *Philos. Trans. A Math. Phys. Eng. Sci.* 368, 2237–2250. doi:10.1098/rsta.2010.0029

Sahinovic, M. M., Eleveld, D. J., Miyabe-Nishiwaki, T., Struys, M. M. R. F., and Absalom, A. R. (2017). Pharmacokinetics and pharmacodynamics of propofol: Changes in patients with frontal brain tumours. *Br. J. Anaesth.* 118 (6), 901–909. doi:10.1093/bja/aex134

Sanjari, N., Shalbaf, A., Shalbaf, R., and Sleigh, J. (2021). Assessment of anesthesia depth using effective brain connectivity based on transfer entropy on EEG signal. *Basic Clin. Neurosci.* 12 (2), 269–280. doi:10.32598/bcn.12.2.2034.2

Scherbakov, N., and Doehner, W. (2018). Heart--brain interactions in heart failure. *Card. Fail. Rev.* 4 (2), 87–91. doi:10.15420/cfr.2018.14.2

Schulz, S., Haueisen, J., Bär, K. J., and Voss, A. (2020). The cardiorespiratory network in healthy first-degree relatives of schizophrenic patients. *Front. Neurosci.* 14, 617. doi:10.3389/fnins.2020.00617

Schwartz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 5, 461–464.

Schwilden, H., Stoeckel, H., and Schüttler, J. (1989). Closed-loop feedback control of propofol anaesthesia by quantitative EEG analysis in humans. *Br. J. Anaesth.* 62 (3), 290–296. doi:10.1093/bja/62.3.290

Siedlecka, J., Siedlecki, P., and Bortkiewicz, A. (2015). *Impedance cardiography–Old method, new opportunities. Part I*. doi:10.13075/ijomeh.1896.00451*Clin. Appl.*

Slupe, A. M., and Kirsch, J. R. (2018). Effects of anesthesia on cerebral blood flow, metabolism, and neuroprotection. *J. Cereb. Blood Flow. Metab.* 38 (12), 2192–2208. doi:10.1177/0271678X18789273

Takens, F. (1981)., 898. Springer, 366–381. doi:10.1007/BFb0091924Detecting strange attractors in turbulence. In: Dynamical systems and turbulence*Lect. Notes Math.*

Tauber, H., Streif, W., Gebetsberger, J., Gasteiger, L., Pierer, E., KnoflachFraedrich, G., et al. (2021). Cardiac output and cerebral blood flow during carotid surgery in regional versus general anesthesia: A prospective randomized controlled study. *J. Vasc. Surg.* 74 (3), 930–937.e2. doi:10.1016/j.jvs.2021.03.042

Voss, A., Fischer, C., Schroeder, R., Fingulla, H. R., and Goernig, M. (2010). Segmented Poincaré plot analysis for risk stratification in patients with dilated cardiomyopathy. *Methods Inf. Med.* 49 (5), 511–515. doi:10.3414/ME09-02-0050

Voss, A., Schroeder, R., Heitmann, A., Peters, A., and Perz, S. (2015). Short-term heart rate variability—Influence of gender and age in healthy subjects. *PLoS One* 10 (3), e0118308. doi:10.1371/journal.pone.0118308

Willems, J. L., Roelandt, J. O. S., De Geest, H., Kesteloot, H., and Joossens, J. V. (1970). The left ventricular ejection time in elderly subjects. *Circulation* 42 (1), 37–42. doi:10.1161/01.cir.42.1.37

Won, D. O., Lee, B. R., Seo, K. S., Kim, H. J., and Lee, S. W. (2019). Alteration of coupling between brain and heart induced by sedation with propofol and midazolam. *PLoS One* 14 (7), e0219238. doi:10.1371/journal.pone.0219238

Zanetti, M., Faes, L., Nollo, G., Zanetti, M., Faes, L., Nollo, G., et al. (2019). Information dynamics of the brain, cardiovascular and respiratory network during different levels of mental stress. *Entropy (Basel, Switz.* 21 (3), 275. doi:10.3390/e21030275

Zhang, P. (1993). On the convergence rate of model selection criteria. *Commun. Statistics - Theory Methods* 22 (10), 2765–2775. doi:10.1080/03610929308831184

frontiers | Frontiers in Network Physiology

TYPE Original Research
PUBLISHED 15 March 2023
DOI 10.3389/fnetp.2023.1106650

OPEN ACCESS

EDITED BY
Luca Faes,
University of Palermo, Italy

REVIEWED BY
Riccardo Pernice,
University of Palermo, Italy
Alireza Mani,
University College London,
United Kingdom

*CORRESPONDENCE
A. M. Brouwer,
anne-marie.brouwer@tno.nl

SPECIALTY SECTION
This article was submitted to Networks of Dynamical Systems, a section of the journal Frontiers in Network Physiology

RECEIVED 24 November 2022
ACCEPTED 22 February 2023
PUBLISHED 15 March 2023

CITATION
Difrancesco S, van Baardewijk JU, Cornelissen AS, Varon C, Hendriks RC and Brouwer AM (2023), Exploring the use of Granger causality for the identification of chemical exposure based on physiological data.
Front. Netw. Physiol. 3:1106650.
doi: 10.3389/fnetp.2023.1106650

Exploring the use of Granger causality for the identification of chemical exposure based on physiological data

S. Difrancesco[1], J. U. van Baardewijk[2], A. S. Cornelissen[3], C. Varon[4,5], R. C. Hendriks[4] and A. M. Brouwer[2]*

[1]Department Systems Biology, The Netherlands Organisation for Applied Scientific Research (TNO), Leiden, Netherlands, [2]Department Human Performance, The Netherlands Organisation for Applied Scientific Research (TNO), Soesterberg, Netherlands, [3]Department CBRN Protection, The Netherlands Organisation for Applied Scientific Research (TNO), Rijswijk, Netherlands, [4]Circuits and Systems (CAS) Group, Delft University of Technology, Delft, Netherlands, [5]Centre for Research and Engineering in Space Technologies—CREST, Université Libre de Bruxelles, Brussels, Belgium

Wearable sensors offer new opportunities for the early detection and identification of toxic chemicals in situations where medical evaluation is not immediately possible. We previously found that continuously recorded physiology in guinea pigs can be used for early detection of exposure to an opioid (fentanyl) or a nerve agent (VX), as well as for differentiating between the two. Here, we investigated how exposure to these different chemicals affects the interactions between ECG and respiration parameters as determined by Granger causality (GC). Features reflecting such interactions may provide additional information and improve models differentiating between chemical agents. Traditional respiration and ECG features, as well as GC features, were extracted from data of 120 guinea pigs exposed to VX (n = 61) or fentanyl (n = 59). Data were divided in a training set (n = 99) and a test set (n = 21). Minimum Redundancy Maximum Relevance (mRMR) and Support Vector Machine (SVM) algorithms were used to, respectively, perform feature selection and train a model to discriminate between the two chemicals. We found that ECG and respiration parameters are Granger-related under healthy conditions, and that exposure to fentanyl and VX affected these relationships in different ways. SVM models discriminated between chemicals with accuracy of 95% or higher on the test set. GC features did not improve the classification compared to traditional features. Respiration features (i.e., peak inspiratory and expiratory flow) were the most important to discriminate between different chemical's exposure. Our results indicate that it may be feasible to discriminate between chemical exposure when using traditional physiological respiration features from wearable sensors. Future research will examine whether GC features can contribute to robust detection and differentiation between chemicals when considering other factors, such as generalizing results across species.

KEYWORDS

Granger causality, chemical exposure, toxidrome detection, physiological data, support vector machine, machine learning

1 Introduction

Wearable sensor technology is rapidly evolving, leading to higher quality data and more physiological parameters that can be monitored simultaneously. This development offers new opportunities to quickly detect and identify toxic chemicals based on their effects on the human body, enabling timely (self-administered) treatment when medical evaluation is not immediately possible, such as in the military battle field. Various specialized devices are under investigation for alcohol or substance abuse or are already commercially available, such as the Secure Continuous Remote Alcohol Monitor (SCRAM®) ankle monitor, which electrochemically detects transdermal alcohol (Davis-Martin et al., 2021). Even though direct chemical detection is the gold standard, it only allows for the monitoring of a limited number of compounds, making it unsuitable for 'threat-agnostic' monitoring. Furthermore, chemical detection is difficult for compounds that are toxic at extremely low systemic levels, such as is the case for novel synthetic opioids as carfentanil (Uddayasankar et al., 2018). Also note that differential diagnosis in the clinic is not always straightforward, as exemplified in the 2018 Salisbury poisoning incident, in which a nerve agent poisoning was mistaken for an opioid overdose (Eddleston & Chowdhury, 2020; Haslam et al., 2022). Indirect detection by measuring the compound's (toxic) effects on the body (toxidrome) presents a promising approach for continuous, non-invasive monitoring of exposure to chemicals. Automatic algorithms may alert the possibly exposed individual or their colleague that quick countermeasures are required. In the battlefield such warnings could be especially helpful given that military personal likely ignore or suppress physical discomfort, and effects of chemicals are initially hidden for others by protective clothing and gas masks.

The effects of chemical intoxication on the body can be complex and multi-facetted. Machine learning models are suitable for complex pattern recognition analyses with relatively large numbers of parameters and previous studies showed that when applied to physiological data, they could detect various chemical intoxications. A study by Mahmud et al. (2018) employed various machine learning methods (decision tree, k-nearest neighbor, eXtreme Gradient Boosting) to detect opioid use based on data from a wrist-band with 99% accuracy. A study by Chang et al. (2021) showed that a neural network trained to recognize digoxin toxicity from electrocardiography (ECG) performed similarly to cardiologists and emergency room specialists, showing 84.6% sensitivity and 96.6% specificity. We previously showed that a machine learning model could accurately detect exposure to an opioid (fentanyl) or a nerve agent (VX) and differentiate between the two, based on continuously measured electroencephalography (EEG), ECG and respiration (whole-body plethysmography) data in guinea pigs (van Baardewijk et al., 2021).

While these studies successfully demonstrated the detection of chemical intoxication based on physiology, they all considered physiological parameters independently. However, under normal physiological conditions, the various biological systems of the body exhibit oscillatory patterns due to underlying feedback and feedforward mechanisms. For instance, heart rate is well-known to be regulated by many such mechanisms. In healthy people, successive beats do not occur at a constant rhythm, instead, (R-R) intervals show considerable variability. The largest contributor to heart rate variability (HRV) is respiratory sinus arrhythmia (RSA). The heart rate increases with inspiration and decreases with expiration, a mechanism by which the body optimizes pulmonary gas exchange (Hayano et al., 1996; Goldberger et al., 2013) and which is thought to be regulated mainly by central mechanisms (Gleb et al., 1936). These and other cardiorespiratory interactions vary under different circumstances, such as different breathing patterns (Stefanovska, 2002; Elstad et al., 2018; Lukarski et al., 2022). Various pathological conditions have been linked to changes in HRV, such as congestive heart failure, diabetes, and depression (Musialik-Łydka et al., 2003; Wang & Wang, 2011; Young & Benton, 2018; Hartmann et al., 2019). HRV has also been implicated as a useful marker for substance abuse (Koenig et al., 2015), withdrawal symptoms (Levin et al., 2019; Garland & Howard, 2021), and exposure to fine particulate matter (Riediker et al., 2018).

Even though the precise mechanisms of HRV remain poorly understood, these studies highlight the fact that the various physiological systems of our body do not function in an isolated manner. Instead, they coordinate and synchronize their functions to maintain a given physiological state. This holistic view of physiology is investigated in the field of Network Physiology (Bashan et al., 2012; Bartsch et al., 2015). Quantifying the interactions of physiological features under different (healthy and intoxicated) circumstances may improve the detection of toxic chemicals as well as increase our understanding of physiological mechanisms.

One method to quantify the (causal, i.e., time ordered) relationships between physiological features is Granger causality (GC), named after the econometrician who first described it in 1969 (Granger, 1969). This technique has been frequently applied in the financial sector, among others for investigating causal relationships between market factors, economic changes, stock prices, and stock price predictions (Výrost et al., 2015; Gao et al., 2018; Gherghina et al., 2020; Thakkar & Chaudhari, 2021). A variety of studies that used methods based on GC to quantify the interactions between physiological signals, mainly focused on heart rate, respiration, and arterial blood pressure. GC in these studies show how well the future of a physiological signal (e.g., heart rate variability) can be predicted from the present and the past of another signal (e.g., respiration) by means of linear vector autoregressive (VAR) models, and result in directionality and strength of interaction (Seth et al., 2015). A study by Faes et al. (2015) applied GC to map directional interactions in brain-brain and brain-heart networks in different sleep states, exemplifying the added value of GC in neuroscience (Porta & Faes, 2015; Seth et al., 2015). A study by Rozo et al. (2021) demonstrated that different methods to quantify RSA, based on GC principles, captured the cardiorespiratory changes expected during different non-REM sleep stages. Interactions between respiration, blood pressure and heart rate have been found to be influenced by factors such as body position (Mary et al., 2019) and deep versus normal breathing (Mary et al., 2018). These interactions have been used to make distinctions between healthy and diseased conditions under conditions such as congestive heart failure (Radovanović et al., 2018), and pre-eclampsia (Riedl et al., 2010). Recent studies successfully used GC between brain and heart signals to characterize epileptic seizures (Pernice et al., 2022) and GC between activity in different brain areas to distinguish between patients with cognitive impairment associated with epilepsy and healthy controls (Jiang et al., 2021).

Acknowledging the potential diagnostic value of physiological interactions in the context of exposure to toxics, and acknowledging GC as a way to quantify such interactions, we here determine the GC interactions within and between both respiration and ECG parameters under healthy and intoxicated

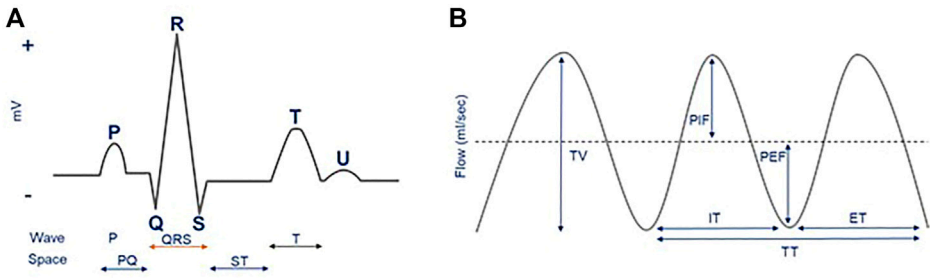


FIGURE 1
ECG **(A)** and respiration **(B)** features illustrated in schematic raw ECG and respiratory signals. In the current study, we used R-R interval (RR-I), ST elevation (ST-E), R height (R-H) and QRS duration (QRS) as traditional ECG features. For respiration, the traditional features that were used were tidal volume (TV), peak inspiratory flow (PIF), peak expiratory flow (PEF), inspiratory time (IT), expiratory time (ET), and total time (TT).

(fentanyl or VX) conditions. RSA, that we already discussed above, is the most studied form of cardiorespiratory GC interactions, despite the identification of other forms (cardiorespiratory phase synchronization: Bartsch et al. (2007), and time delay stability; Bartsch et al., 2014). One of the reasons for this is that RSA can be directly estimated using predictability and casual measures based on GC, applied to the raw respiratory signal and the tachogram, derived from the ECG. Such measures are often used to estimate the information transferred from a driver, often the respiration, to a target signal (e.g., the tachogram). Here, however, we study the cardiorespiratory interactions not by predicting one (close to) raw signal from the other, but by looking at the effect that specific respiratory higher order features have on the morphological and rhythm features of the ECG, such as the effect of peak inspiratory flow on the interval between successive heart beats, and *vice versa*. Interactions between these features are also examined within modality. Relying on higher order features is important from the perspective of our envisioned ultimate application of using wearables for diagnostics in the field, where the quality of the raw physiological signals and their synchronization is likely compromised. Since it is currently unknown how exposure to fentanyl and VX affect GC interactions, we first provide an overview of the interactions for each condition. Next, we evaluate the contribution of traditional ECG and respiration features as well as GC features in machine learning models that aim to differentiate between exposure to fentanyl and VX over 45 min following exposure as well as over the first 15 min of exposure (which we considered as a cut-off for a timely treatment in an exposure scenario). The study here is an updated and extended version of a previous proceedings paper [van Baardewijk et al. (2022)].

2 Materials and methods

2.1 Sample

Data comprised four existing physiological datasets of freely moving guinea pigs, exposed to VX ($n = 62$) or fentanyl ($n = 71$). The animal procedures were as described previously (Joosen et al., 2017). Briefly, VX was obtained from the in-house synthesized TNO stocks. Purity upon issue was >98%. Fentanyl citrate (European Pharmacopoea grade) was purchased from Spruyt-Hillen (IJsselstein, Netherlands). Purity was >99%. VX was either dissolved in 2-propanol (IPA) to the required concentration or applied as neat agent. The VX doses applied were 1–2 mg/kg dermally, corresponding to approximately 1.5–3 times the 24 h LD50 values in guinea pigs (Rice et al., 2015). The fentanyl doses ranged from 0.05 to 8 mg/kg (intravenous bolus) and 0.4–32 mg/kg (subcutaneous), selected to elicit varying degrees of respiratory depression. Fentanyl was dissolved in phosphate-buffered saline (PBS) to the required concentration before administration. For continuous measurements, animals were surgically equipped with ECG leads. Two leads were sutured in the superficial muscles under the skin right below the right collar bone and between the second and third rib (configuration II). ECG data were transmitted wirelessly to a hardware system (Data Sciences International (DSI), St. Paul, MN, United States) using F40-EET (nominal sampling rate 240 Hz) or HD-S02 (nominal sampling rate 375 Hz) telemetry devices. Unrestrained respiratory plethysmography (URP) data were obtained using whole-body plethysmography cages, connected to a Universal XE signal conditioner (DSI). Telemetry and plethysmography data were upsampled simultaneously at 1,000 Hz using the Ponemah Physiology Platform (v5.41) software, in order to combine the modalities into a single dataset. Under typical conditions, the synchronization error of the two modalities was within 150 ms. For each animal, at least 30 min of data were acquired before exposure. The final sample included 120 animals; nine animals were excluded because they belonged to a placebo group, four animals were excluded because they died during the experiment.

2.2 Preprocessing: From raw data to the extraction of ECG and respiratory features

Physiological data were preprocessed using Ponemah® Software. The signals were inspected visually to identify and exclude artifacts related to

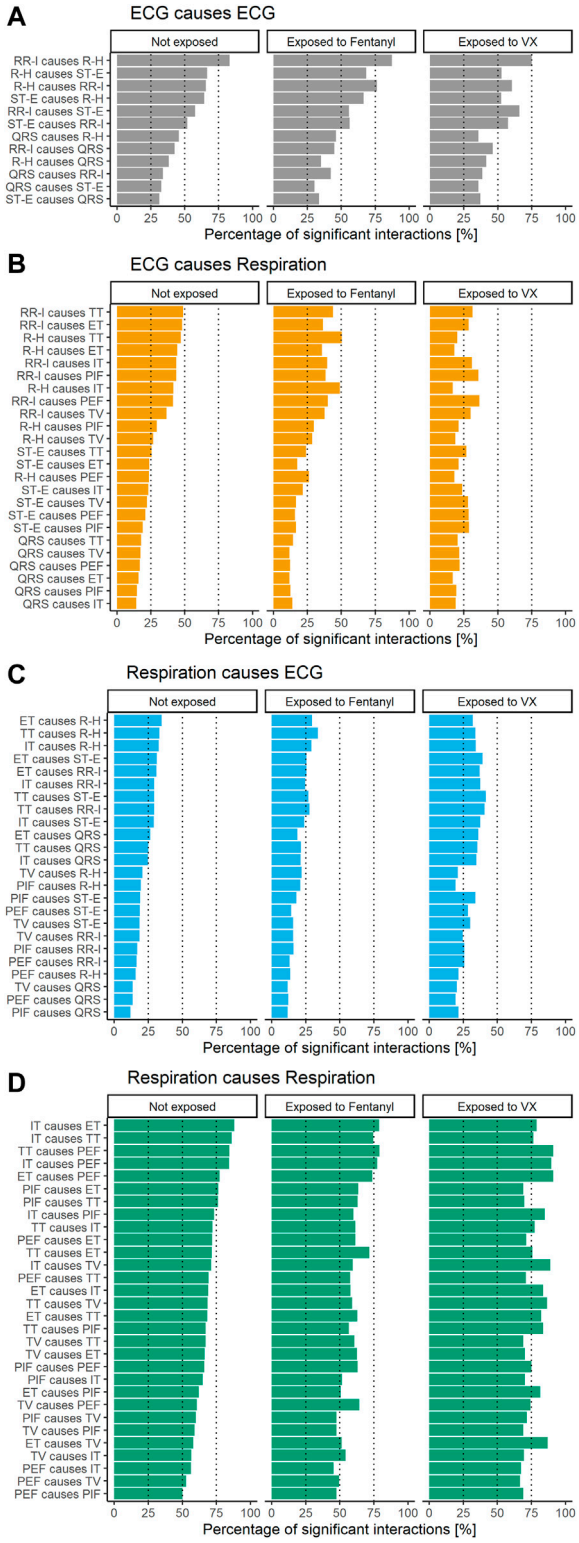


FIGURE 2
Percentage of 5-minute windows with significant GC per condition and feature combination for ECG causing ECG **(A)**, ECG causing respiration **(B)**, respiration causing ECG **(C)**, and respiration causing respiration **(D)**.

movements and sudden ambient pressure changes. All derived features were subsequently exported in a beat-to-beat format for further processing. The following four traditional ECG features were extracted from ECG data: R-R interval (RR-I), ST elevation (ST-E), R height (R-H) and QRS duration (QRS). For respiration, the six traditional features that were extracted from URP data were: tidal volume (TV), peak inspiratory

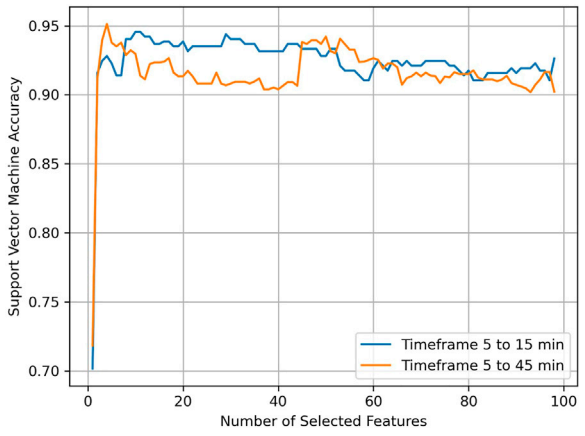


FIGURE 3
SVM accuracy as a function of number of selected features.

flow (PIF), peak expiratory flow (PEF), inspiratory time (IT), expiratory time (ET), and total time (TT; TT = IT + ET). These ECG and respiratory features are illustrated in Figures 1A, B, respectively.

To identify and remove signal artifacts, z-scores were determined for 20 s moving windows (shifted in steps of 1 s). Datapoints with a z-score higher than 3 or lower than −3 were removed.

Data around the moment of exposure (from 5 min before to 5 min after) were excluded to prevent any handling effects related to administering the chemical potentially influencing the data.

2.3 Traditional ECG and respiratory features

Extracted features were aggregated over successive 5-min windows for data from 30 min before exposure to 45 min after exposure. Within-animal centering of features was done by subtracting the baseline, which was defined as the average feature value as recorded during the first 15 min (i.e., from 45 min until 30 min before exposure). Missing data were linearly interpolated. Such features were used as input for Support Vector Machine (SVM) classification analysis (Cortes & Vapnik, 1995) as described later.

2.4 Granger causality (GC) features

In this study, a bivariate formulation of GC was used. In such a formulation a system consists of two variables X and Y. Y causes X, in the Granger sense, if the past of Y (Y_n^l) provides information about the future of variable X, given the past of X (X_n^k), where Y_n and X_n denote the present value of Y and X respectively and:

$$X_n^k = [X_{n-1}, X, \ldots, X_{n-k+1}], \quad (1)$$

$$Y_n^l = [Y_{n-1}, Y_{n-2}, \ldots, Y_{n-l+1}], \quad (2)$$

where k and l represent the time lags.

For extracting GC features between each pair of features, the different time series (i.e., extracted features) were resampled at aligned points in time. For subsequent windows of 100 ms, data points in a specific window were averaged. Average data points were linearly interpolated and data were resampled at 10 Hz. The time series were then detrended by means of differencing.

GC features were determined for each animal, each 5-min window, and each of the 90 combinations of traditional ECG and respiratory features (four ECG and six respiratory features as described in 2.2 gives a total of 10 features; leading to 10*10–10 = 90 unique combinations). To determine the optimal GC lag, the Vector Autoregression (VAR) on the healthy data of all animals was calculated by varying the lag from 1 to 50 100 ms-windows. The optimal lag value was the one with the lowest average Akaike information criterion (AIC) (Akaike, 1974). In this case, the optimal lag was found to be 35 100 ms-windows, i.e. 3.5 s.

Statistical significance of GC was determined with an SSR F-test, resulting in a p-value ($\alpha = 0.05$) for each combination in each 5-min window. F-statistic values were used as GC features in the prediction models. Percentages of statistically significant GC features were plotted to give an impression of which interactions between ECG and respiratory features were more present than others in the current sample.

To extract Granger causality features and to identify the optimal lag as just described, we used Python and the statsmodels module version 0.12.2. Specifically, the "VAR" class from the statsmodels module was used to calculate the VAR. For each lag a fit was done and from the result the AIC value was used. The function "grangercausalitytests" from the statsmodels module was used for calculation of the Granger Causality. GC for each pair of features was calculated by putting the values of the two features in a two-column dataframe and using that as the first input for this function. A list of one element was used for the "maxlag" parameter, with the optimal lag as the only element, so only this lag was used as parameter for the test. From the output of this function, the p-value from "ssr_ftest" was used for statistical significance as described above.

2.5 Feature selection, classification of exposure and classification evaluation

Using classification analysis we explored whether GC could support classification of respiration and ECG data into either exposure to VX or fentanyl. Twenty percent of the data (21 of the animals) was set aside as a test set to evaluate the final model after the training phase (using data of 99 animals). The proportion of animals exposed to either VX and fentanyl was held constant between the training and the test set.

A standard scaler was used to standardize the data before using the SVM. After that, feature selection, classification and evaluation of the classification was performed in the dataset stratified by time from the exposure (i.e., 5–45 min and 5–15 min from exposure).

Because the final training set was composed of 100 features (4 traditional ECG features, 6 traditional respiratory features, and 90 GC features), feature selection was performed. The minimum Redundancy Maximum Relevance (mRMR) algorithm was used to rank features by their importance. mRMR ranks features high if they

10.3389/fnetp.2023.1106650

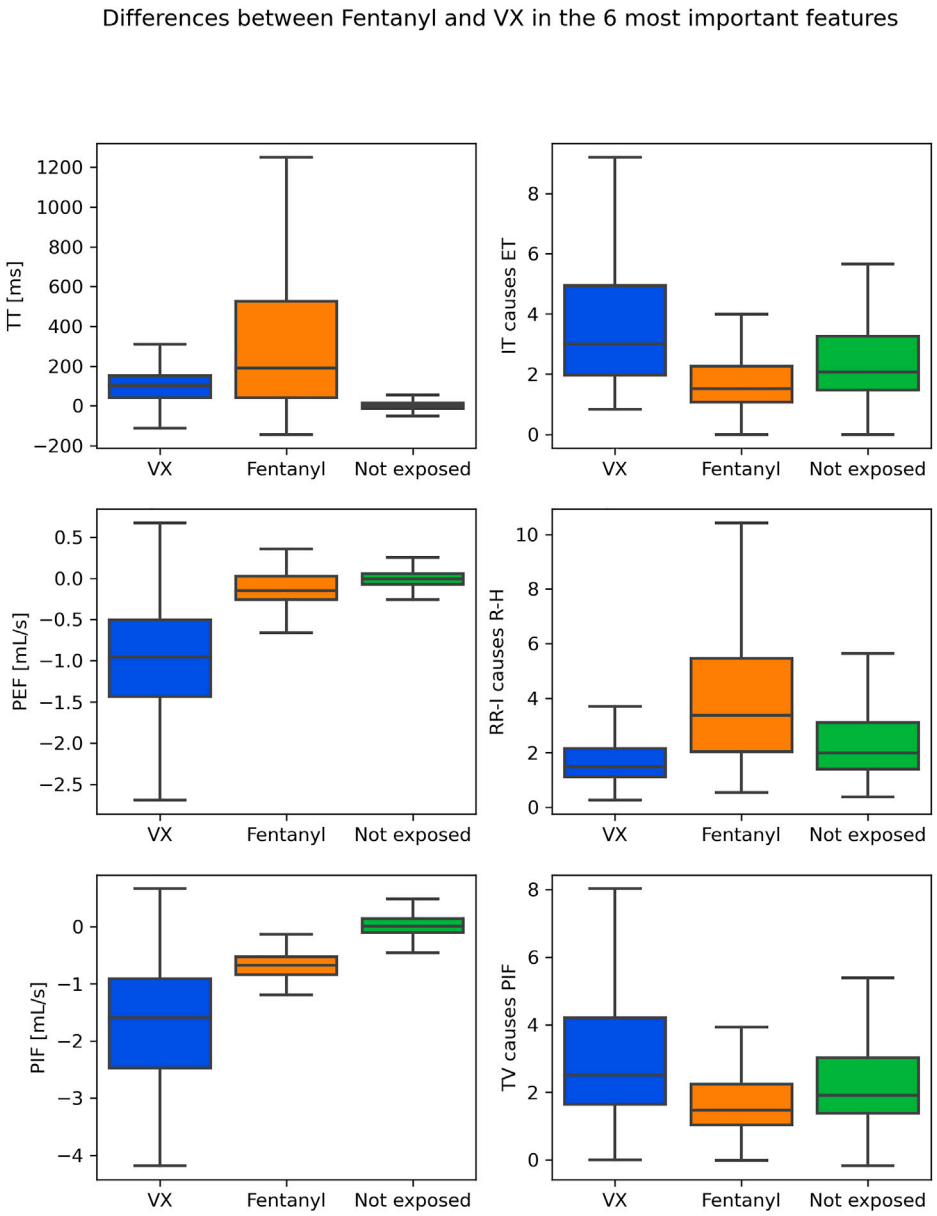


FIGURE 4
Boxplots for VX (blue) and fentanyl (orange) exposure in the most important features in the first 45 min from exposure: top 3 traditional features on the left, top 3 GC features on the right. For reference, values for the not exposed condition are shown as well (green).

are mutually far away from each other (i.e., minimum redundancy) while still correlating strongly with the to be classified variable (i.e., maximum relevance) (Peng et al., 2005). After that, an SVM was used to classify the exposure to the type of chemical based on varying numbers of features. Specifically, an SVM (with standard hyperparameters) was trained by adding one of the ranked features at the time, following their rank and starting with the first. Leave-one-group-out (LOGO) cross validation was applied to evaluate the accuracy of the classification model by leaving one animal out at each iteration on the training set. The average accuracy varying by the number of ranked features was plotted to qualitatively select the optimal number of features. A similar approach has been used before (Peng et al., 2005). The top traditional as well as GC features were selected.

A grid-search was then used for tuning the SVM and optimize the following hyperparameters: kernel, C or "regularization parameter", and gamma. The kernel defines whether the decision boundary is linear or not. Here, a linear and a radial basis function were used as candidate kernels. The constant C represents the tradeoff between minimizing the training set error and maximizing the margin. Gamma is a parameter for nonlinear kernels; gamma controls the influence of each feature on the decision boundary. C and gamma were initialized on different scales ranging from 0.0001 to 100. LOGO was used in this step to optimize the SVM as well as performing internal validation for the best set of parameters. Finally, the tuned SVM was evaluated on the test set (i.e., external validation).

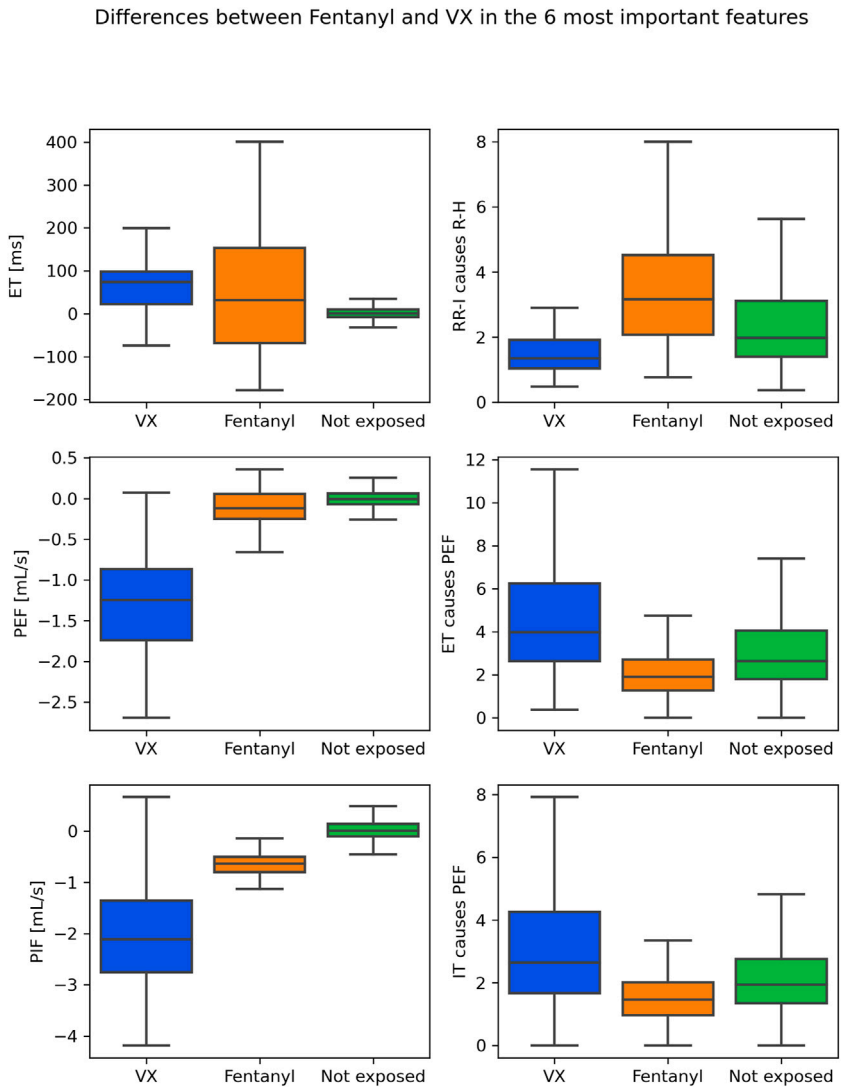


FIGURE 5
Boxplots for VX (blue) and fentanyl (orange) exposure in the most important features in the first 15 min from exposure: top 3 traditional features on the left, top 3 GC features on the right. For reference, values for the not exposed condition are shown as well (green).

Python 3.9 was used to perform the analyses. The pymrmr library was used to rank features and the sklearn library was used to build the classification model.

3 Results

Figure 2 shows the percentage of 5-min windows with significant GC. Not exposed conditions are shown on the left, exposure to fentanyl in the middle and exposure to VX on the right. Results are shown separately for ECG causing ECG features a), ECG causing respiration features b), respiration causing ECG features c) and respiration causing respiration features d). All feature combinations were GC related for more than 25% of the 5-min windows within modality (ECG causing ECG and respiration causing respiration features). Most feature combinations were GC related for more than 20% of the cases between modalities (ECG causing respiration and respiration causing ECG features). Patterns appear to differ between the three exposure conditions.

The feature selection step showed that with only the top three features (ranked PEF, TT, PIF for the first 45 min; ranked PEF, PIF, ET for the first 15 min) the accuracy of the SVM on the training set already reached a high accuracy for both the 45 and the 15 min case (Figure 3). The following top three features in the feature selection step were all GC features. For the first 45 min from exposure, these were (ranked) IT causes ET, RR-I causes R-H, PIF causes TV. For the first 15 min from exposure, these were (ranked) RR-I causes R-H, TT causes PEF, PIF causes TV.

Figures 4, 5 provide insight into how the features that came out as the most relevant features for differentiating between fentanyl and VX differed between these chemical exposure conditions. The left panels in Figure 4 show how the top three (traditional) variables differed for the first 45 min after exposure. TT, PEF and PIF were all lower for fentanyl than for VX. The left

TABLE 1 SVM hyperparameters tuning.

	Kernel	C	ɣ
5 to 45 min from exposure			
Traditional features[a]	rbf	90	0.02
GC features[b]	linear	0.09	5
Traditional + GC features[a+b]	rbf	90	0.01
5 to 15 min from exposure			
Traditional features[c]	rbf	100	0.10
GC features[d]	rbf	1.50	0.50
Traditional + GC features[c+d]	rbf	100	0.10

[a]Features = TT, PEF, PIF.
[b]Features = IT causes ET, RR-I causes RH, TV causes PIF.
[c]Features = ET, PEF, PIF.
[d]Features = RR-I causes R-H, ET causes PEF, IT causes PEF.

panels in Figure 5 show how the three highest ranked features (PEF, PIF and ET) differed between the chemical exposure conditions when only the first 15 min from exposure were considered. As for the 45 min, PEF and PIF were lower for fentanyl than for VX. ET was higher for fentanyl than for VX. The right panels in Figures 4, 5 indicate how the top GC features differed between conditions. The GC for ECG features (RR-I causes R-H in both the 45- and 15 min case) was low in VX compared to fentanyl. In contrast, the GC for respiration features (with slightly different GC combination popping up for the 45- and 15-min case) was low in fentanyl compared to VX. To provide insight as to how the exposure conditions compare to the not exposed condition, Figures 4, 5 include these data as well.

Table 1 shows optimal hyper parameter settings used in the SVM. Table 2 shows internal validation and external validation results. Accuracy of the SVM using traditional features alone was high for both the first 45 min and the first 15 min (respectively 95% and 97% for the test sets). Results were comparable when using traditional features alone versus traditional and GC features in both the 45 and 15 min time frames, indicating that GC features did not add to traditional features. Still, models using only GC features had an accuracy of 79%, confirming the impression from Figure 2 that relations between physiological features differed between exposure to VX and fentanyl.

4 Discussion

We explored the relation between different ECG and respiration parameters under healthy conditions, after exposure to fentanyl, and after exposure to VX using Granger causality. We were especially interested in whether the interactions between physiological signals, quantified using Granger's method, could be used to improve discrimination by machine learning models between exposure to VX and fentanyl relative to using traditional features alone. Quantification of (cardiorespiratory) interactions may be useful in improving machine learning models designed to detect acute chemical intoxication and discriminate between chemicals based on non-invasive physiological data, as well as improve our understanding of chemicals' toxic effects.

SVM classification showed that it was already possible to discriminate between VX and fentanyl with high accuracy (95%) after 15 min, using traditional features. While models using GC features alone showed that these features contained information, adding them did not result in improved classification of the already high accuracy reached by using only traditional features. Respiration features were the most important to discriminate between the two types of exposure. This is consistent with the different pharmacological mechanisms by which both compound classes exert their toxic effects and the used administration routes. Opioids, such as fentanyl, directly bind to the mu opioid receptor (MOR), disrupting the central respiratory drive, controlled by various respiratory centers in the brainstem (Pattinson, 2008; van der Schier et al., 2014). In the current data sets, fentanyl poisoning occurred *via* intravenous and subcutaneous exposure, leading to rapid intoxication. Nerve agents, such as VX, cause a cholinergic crisis, leading to a wide palette of signs and symptoms typical for nicotinic and muscarinic overstimulation. In the current data sets, VX poisoning occurred *via* dermal exposure, resulting in a steady progression of toxicity, where bradycardia and hypothermia are typically observed first and respiratory distress at a more severe state (Hamilton et al., 2004; Mumford et al., 2008). Our finding that it was already possible to discriminate between the chemicals in the first 15 min from the exposure with good accuracy demonstrate the feasibility to discriminate between chemical exposure when using respiration data that may be measured using wearable sensors. This

TABLE 2 Internal and external validation of SVM prediction of chemical exposure.

	Internal validation	External validation
	Accuracy, mean (SD)	Accuracy
5 to 45 min from exposure		
Traditional features[a]	0.954 (0.215)	0.950
GC features[b]	0.754 (0.425)	0.790
Traditional + GC features[a+b]	0.953 (0.203)	0.940
5 to 15 min from exposure		
Traditional features[c]	0.970 (0.260)	0.970
GC features[d]	0.767 (0.614)	0.790
Traditional + GC features[c+d]	0.958 (0.256)	0.970

[a]Features = TT, PEF, PIF.
[b]Features = IT causes ET, RR-I causes RH, TV causes PIF.
[c]Features = ET, PEF, PIF.
[d]Features = RR-I causes R-H, ET causes PEF, IT causes PEF.

is important to provide timely interventions to reverse the effects of chemicals' exposure. It remains to be seen how the accuracy of models such as these varies with other compounds and other dosages.

The fact that classification accuracy was already high when using traditional features alone made it difficult for GC features to further improve that. Models based on GC features alone showed that these features contained information, but they performed poorly compared to models based on traditional features alone. This may be explained by two main limitations of this work. Firstly, the GC features were computed under the assumptions of stationarity and joint Gaussian distribution. As a result, only the linear interactions could be captured, thereby ignoring possible nonlinearities that could be strongly affected by the exposure. Therefore, future work should focus on the quantification of these possibly nonlinear interactions (Rozo et al., 2021). Secondly, the interactions between the features were assumed to be constant throughout the 15 or 45 min after the exposure. It is, however, still unknown whether such interactions change towards the general health deterioration caused by the exposure. Future studies will investigate if such dynamic changes are stronger and occur faster in the cross-modality features when compared to traditional ECG and respiratory features.

While the current work did not demonstrate a large contribution of Granger causality features for the purpose of distinguishing between exposure to different toxic chemicals, these features may add value for the purpose of generalizing results across species and across movement conditions. Automatic and early detection of exposure to toxic chemicals can save human lives, but studying the physiological effects of these chemicals can only be done in animals where it is questionable how well these models generalize to humans. Also, large variations in body movement and posture may make it hard to automatically detect and exposure to chemicals. In future work we hope to examine how traditional as well as interaction features vary across species, movement and exposure conditions in order to select the features that are insensitive to variations in species and movement. As mentioned before, the cardiorespiratory interactions can be analyzed from the GC perspective using the raw respiratory signal and the tachogram. This study, instead, highlights that the GC relations between respiration and ECG are also prominent when higher order variables are used, which are derived from the raw signal (e.g., RR-I, TT), suggesting that exact synchronization and high-quality raw signals may not be essential. This is an advantage since such high-quality signals are known to be difficult to record using wearables and under ambulatory conditions. Future work will examine how traditional GC relations between respiration and ECG compare for these different approaches, and for data from wearables compared to high-end equipment.

Data availability statement

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author.

Ethics statement

The animal study was reviewed and approved by TNO ethics committee—All experiments were carried out according to the EU Legislation for testing on experimental animals (EU Directive 2010/63/EU) at the TNO CBRN Protection Department, Rijswijk, The Netherlands.

Author contributions

SD and JB developed the methodology, performed analyses and contributed to the results and interpretation. AC contributed to data curation and pre-processing. SD, JB, AC, CV, and AB contributed to the original draft—writing. RH contributed to the original draft—review and editing. AB provided supervision.

Funding

This study was funded by the Dutch Ministry of Defence.

Acknowledgments

We thank Marloes Joosen for starting up the collaboration on detection of exposure using physiological data and for providing support. We also thank Sarthak Agarwal who started the Granger causality and classification analyses during his internship at TNO.

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.

Publisher's note

References

Akaike, H (1974). A new look at the statistical model identification. *IEEE Trans. Automatic Control* 19 (6), 716–723. doi:10.1109/tac.1974.1100705

Bartsch, L, Ma, I, Ma, Q. D., and Ivanov, P. C. (2014). Three independent forms of cardio-respiratory coupling: Transitions across sleep stages. *Comput. Cardiol.* 41, 781–784.

Bartsch, K, Penzel, H, and Havlin, S. (2007). Experimental evidence for phase synchronization transitions in the human cardiorespiratory system. *Phys. Rev. Lett.* 98 (5), 054102. doi:10.1103/PhysRevLett.98.054102

Bartsch, R. P., Liu, K. K. L., Bashan, I., Bashan, A., and Ivanov, P. C. (2015). Network physiology: How organ systems dynamically interact. *PLoS ONE* 10 (11), 0142143. doi:10.1371/journal.pone.0142143

Bashan, B, Kantelhardt, H, Kantelhardt, J. W., Havlin, S., and Ivanov, P. C. (2012). Network physiology reveals relations between network topology and physiological function. *Nat. Commun.* 3, 702. doi:10.1038/ncomms1705

Chang, L, Tsao, L, Chen, T, Lin, L, Chen, J. T., Tsai, C. S., et al. (2021). Detecting digoxin toxicity by artificial intelligence-assisted electrocardiography. *Int. J. Environ. Res. Public Health* 18 (7), 3839. doi:10.3390/ijerph18073839

Cortes, C, and Vapnik, V (1995). Support-vector networks. *Mach. Learn.* 20 (3), 273–297. doi:10.1007/bf00994018CiteSeerX 10.1.1.15.9362

Davis-Martin, R. E, and Boudreaux, E. D (2021). Alcohol use disorder in the age of technology: A review of wearable biosensors in alcohol use disorder treatment. *Front. Psychiatry* 12. doi:10.3389/fpsyt.2021.642813

Eddleston, C (2020). Organophosphorus poisoning: The wet opioid toxidrome. *Lancet* 6736, 2020–2022.

Elstad, O. C, Smith, B, Ben-Tal, A., and Ramchandra, R. (2018). Cardiorespiratory interactions in humans and animals: Rhythms for life. *Am. J. Physiology. Heart Circulatory Physiology* 315 (1), H6–H17. doi:10.1152/ajpheart.00701.2017

Faes, M, Jurysta, N, Jurysta, F., and Nollo, G. (2015). Linear and non-linear brain-heart and brain-brain interactions during sleep. *Physiol. Meas.* 36 (4), 683–698. doi:10.1088/0967-3334/36/4/683

Gao, H, Sun, H, Sun, X., Hao, X., and An, F (2018). Modelling cointegration and Granger causality network to detect long-term equilibrium and diffusion paths in the financial system. *R. Soc. Open Sci.* 5 (3), 172092. doi:10.1098/rsos.172092

Garland, H, and Howard, M. O. (2021). Prescription opioid misusers exhibit blunted parasympathetic regulation during inhibitory control challenge. *Psychopharmacology* 238 (3), 765–774. doi:10.1007/s00213-020-05729-z

Gherghina, A, Armeanu, D. Ş., and Joldeş, C. C. (2020). Stock market reactions to COVID-19 pandemic outbreak: Quantitative evidence from ARDL bounds tests and granger causality analysis. *Int. J. Environ. Res. Public Health* 17 (18), 6729. doi:10.3390/ijerph17186729

Gleb, P, and Roessler, R (1936). Respiratory variations of the heart rate - II—the central mechanism of the respiratory arrhythmia and the inter-relations between the central and the reflex mechanisms. *Proc. R. Soc. Lond. Ser. B - Biol. Sci.* 119 (813), 218–230. doi:10.1098/rspb.1936.0006

Goldberger, G, and Shvilkin, A (2013). Chapter 13-sinus and escape rhythms. *Goldb. Clin. Electrocardiogr. (Ninth Ed. Goldb. AL, Goldb. ZD, Shvilkin, A., Eds*, 114–120. doi:10.1016/B978-0-323-08786-5.00013-0W.B. Saunders

Granger, C. W. J (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37 (3), 424–438. doi:10.2307/1912791https://econpapers.repec.org/RePEc:ecm:emetrp:v:37:y:1969:i:3:p:424-38

Hamilton, H, Conley, S, Caneva, L, Sawyer, T. W., Caneva, D. C., and Lundy, P. M. (2004). Clinical aspects of percutaneous poisoning by the chemical warfare agent VX: Effects of application site and decontamination. *Mil. Med.* 169 (11), 856–862. doi:10.7205/milmed.169.11.856

Hartmann, S, Sander, H, Sander, C., and Hegerl, U. (2019). Heart rate variability as indicator of clinical state in depression. *Front. Psychiatry* 9, 735. doi:10.3389/fpsyt.2018.00735

Haslam, J. D., Russell, P., Hill, S., Emmett, S. R, and Blain, P. G (2022). Chemical, biological, radiological, and nuclear mass casualty medicine: A review of lessons from the Salisbury and amesbury novichok nerve agent incidents. *Br. J. Anaesth.* 128 (2), e200–e205. doi:10.1016/j.bja.2021.10.008

Hayano, Y, Okada, M, Mukai, S., and Fujinami, T. (1996). Respiratory sinus arrhythmia. A phenomenon improving pulmonary gas exchange and circulatory efficiency. *Circulation* 94 (4), 842–847. doi:10.1161/01.cir.94.4.842

Jiang, Z, Nie, L, Liu, H., and Zheng, J. (2021). Identification and effective connections of core networks in patients with temporal lobe epilepsy and cognitive impairment: Granger causality analysis and multivariate pattern analysis. *International Journal of Neuroscience*, 1–12. doi:10.1080/00207454.2021.2017926

Joosen, M. J, Van Den Berg, R. M, De Jong, A. L, Van Der Schans, M. J, Noort, D, and Langenberg, J. P (2017). The impact of skin decontamination on the time window for effective treatment of percutaneous VX exposure. *Chemico-Biological Interact.* 267, 48–56. doi:10.1016/j.cbi.2016.02.001

Koenig, M, Hillecke, T, and Jarczok, M, N (2015). Heart rate variability and cocaine: A systematic review of human studies. *Archives Neurosci.* 2 (1), e60035. doi:10.5812/archneurosci.18798

Levin, W, Jones, C, and Jones, J. D. (2019). Changes in cardiac vagal tone as measured by heart rate variability during naloxone-induced opioid withdrawal. *Drug Alcohol Dependence* 204, 107538. doi:10.1016/j.drugalcdep.2019.06.040

Lukarski, S, Stavrov, D., and Stankovski, T. (2022). Variability of cardiorespiratory interactions under different breathing patterns. *Biomed. Signal Process. Control* 71, 103152. doi:10.1016/j.bspc.2021.10315210.1016/j.bspc.2021.103152

Mahmud, F, Wang, C, Wang, H., Carreiro, S., and Boyer, E. (2018). Automatic detection of opioid intake using wearable biosensor. Poceedings of the 2018 Int. Conf. Comput. Netw. Commun. ICNC March 2018, 784–788.Maui, HI, USA doi:10.1109/ICCNC.2018.83903348390334

Mary, S, Singh, D., and Deepak, K. (2019). Assessment of interaction between cardio-respiratory signals using directed coherence on healthy subjects during postural change. *IRBM* 40 (3), 167–173. doi:10.1016/j.irbm.2019.04.002

Mary, S, Singh, D., and Deepak, K. (2018). Impact of respiration on cardiovascular coupling using Granger causality analysis in healthy subjects. *Biomed. Signal Process. Control* 43, 196–203. doi:10.1016/j.bspc.2018.03.008

Mumford, P, Price, M. E., and Wetherell, J. R. (2008). A novel approach to assessing percutaneous VX poisoning in the conscious Guinea-pig. *J. Appl. Toxicol. JAT* 28 (5), 694–702. doi:10.1002/jat.1324

Musialik-Łydka, S, Sredniawa, B., and Pasyk, S. (2003). Heart rate variability in heart failure. *Kardiologia Pol.* 58 (1), 10–16.

Pattinson, K. T (2008). Opioids and the control of respiration. *Br. J. Anaesth.* 100 (6), 747–758. doi:10.1093/bja/aen094

Peng, L, Long, F., and Ding, C. (2005). Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Analysis Mach. Intell.* 27 (8), 1226–1238. doi:10.1109/TPAMI.2005.1592005.159

Pernice, F, Feucht, B, Mangione, S, Benninger, F., and Schiecke, K. (2022). Pairwise and higher-order measures of brain-heart interactions in children with temporal lobe epilepsy. *J. Neural Eng.* 19 (4), 045002. doi:10.1088/1741-2552/ac7fba

Porta, A., and Faes, L. (2015). Wiener–Granger causality in network physiology with applications to cardiovascular control and neuroscience. *Proc. IEEE* 104 (2), 282–309. doi:10.1109/jproc.2015.2476824

Radovanović, P, Milašinović, K, Milasinovic, G., Kircanski, B., and Platisa, M. M. (2018). Bidirectional cardio-respiratory interactions in heart failure. *Front. Physiology* 9, 165. doi:10.3389/fphys.2018.00165

Rice, H., Dalton, C. H., Price, M. E., Graham, S. J., Green, A. C., Jenner, J., et al. (2015). Toxicity and medical countermeasure studies on the organophosphorus nerve agents VM and VX. *Proc. R. Soc. A* 471, 20140891. doi:10.1098/rspa.2014.0891

Riediker, F, Bochud, M, and Rousson, V (2018). Exposure to fine particulate matter leads to rapid heart rate variability changes. *Front. Environ. Sci.* 6. https://www.frontiersin.org/articles/10.3389/fenvs.2018.00002.

Riedl, S, Stepan, K, Stepan, H., Kurths, J., and Wessel, N. (2010). Short-term couplings of the cardiovascular system in pregnant women suffering from pre-eclampsia. *Philosophical Trans. R. Soc. A Math. Phys. Eng. Sci.* 368 (1918), 2237–2250. doi:10.1098/rsta.2010.0029

Rozo, A, Morales, J, Moeyersons, J, Joshi, R, Caiani, E. G, Borzée, P, Buyse, B, Testelmans, D, Van Huffel, S, and Varon, C (2021). Benchmarking transfer entropy methods for the study of linear and nonlinear cardio-respiratory interactions. *Entropy* 23 (8). doi:10.3390/e23080939

Seth, B, Barrett, A. B., and Barnett, L. (2015). Granger causality analysis in neuroscience and neuroimaging. *J. Neurosci.* 35 (8), 3293–3297. doi:10.1523/JNEUROSCI.4399-14.2015

Stefanovska, A (2002). Cardiorespiratory interactions. *Nonlinear Phenom. Complex Syst.* 5 (4), 462–469.

Thakkar, C (2021). Fusion in stock market prediction: A decade survey on the necessity, recent developments, and potential future directions. *Int. J. Inf. Fusion* 65, 95–107. doi:10.1016/j.inffus.2020.08.019

Uddayasankar, L, Oleschuk, E, Oleschuk, C., Eschun, G., and Ariano, R. E. (2018). The pharmacokinetics and pharmacodynamics of carfentanil after recreational exposure: A case report. *Pharmacother. J. Hum. Pharmacol. Drug Ther.* 38 (6), e41–e45. doi:10.1002/phar.2117

van Baardewijk, J. U., Agarwal, S., Cornelissen, A. S., Joosen, M. J., Kentrop, J., Varon, C., Brouwer, A. M, et al. (2021). Early detection of exposure to toxic chemicals using continuously recorded multi-sensor physiology. *Sensors (Basel, Switz.* 21 (11), 3616. doi:10.3390/s21113616

van Baardewijk, J. U, Agarwal, S, Cornelissen, A. S, Varon, C, Hendriks, R. C, Kentrop, J, Joosen, M. J. A, and Brouwer, A. M (2022). Quantifying interactions between physiological signals to identify exposure to different Chemicals. *Proc. Jt. 12th Int. Conf. Methods Tech. Behav. Res. 6th Seminar Behav. Methods*, Volume 2 72–82.

Van Der Schier, R., Roozekrans, M., Van Velzen, M., Dahan, A, and Niesters, M (2014). Opioid-induced respiratory depression: Reversal by non-opioid drugs. *F1000Prime Rep.* 6, 79. doi:10.12703/P6-79

Výrost, L, Lyocsa, S., and Baumohl, E. (2015). Granger causality stock market networks: Temporal proximity and preferential attachment. *Phys. A Stat. Mech. Its Appl.* 427, 262–276. doi:10.1016/j.physa.2015.02.017

Wang, W, and Wang, C. (2011). Detrended fluctuation analysis of pathological cardiac signals. *Sheng wu yi xue gong cheng xue za zhi* 28 (3), 484–486.

Young, B, and Benton, D. (2018). Heart-rate variability: A biomarker to study the influence of nutrition on physiological and psychological health? *Behav. Pharmacol.* 29 (2), 140–151. doi:10.1097/FBP.0000000000000383and 3-Spec Issue

frontiers | Frontiers in Network Physiology

TYPE Original Research
PUBLISHED 31 May 2023
DOI 10.3389/fnetp.2023.1085347

OPEN ACCESS

EDITED BY
Michal Javorka,
Comenius University, Slovakia

REVIEWED BY
Michele Allegra,
Dipartimento di Fisica e Astronomia "G. Galilei", Italy
Teodor Buchner,
Warsaw University of Technology, Poland
Sven Braeutigam,
University of Oxford, United Kingdom

*CORRESPONDENCE
Michel Besserve,
michel.besserve@tuebingen.mpg.de

RECEIVED 31 October 2022
ACCEPTED 11 May 2023
PUBLISHED 31 May 2023

CITATION
Shao K, Logothetis NK and Besserve M (2023), Information theoretic measures of causal influences during transient neural events.
Front. Netw. Physiol. 3:1085347.
doi: 10.3389/fnetp.2023.1085347

Information theoretic measures of causal influences during transient neural events

Kaidi Shao[1,2,3], Nikos K. Logothetis[1,2,4] and Michel Besserve[2,5]*

[1]International Center for Primate Brain Research (ICPBR), Center for Excellence in Brain Science and Intelligence Technology (CEBSIT), Chinese Academy of Sciences (CAS), Shanghai, China, [2]Department of Cognitive Neurophysiology, Max Planck Institute for Biological Cybernetics, Tübingen, Germany, [3]Graduate School of Neural and Behavioral Sciences, International Max Planck Research School, Eberhard-Karls University of Tübingen, Tübingen, Germany, [4]Centre for Imaging Sciences, Biomedical Imaging Institute, The University of Manchester, Manchester, United Kingdom, [5]Department of Empirical Inference, Max Planck Institute for Intelligent Systems, Tübingen, Germany

Introduction: Transient phenomena play a key role in coordinating brain activity at multiple scales, however their underlying mechanisms remain largely unknown. A key challenge for neural data science is thus to characterize the network interactions at play during these events.

Methods: Using the formalism of Structural Causal Models and their graphical representation, we investigate the theoretical and empirical properties of Information Theory based causal strength measures in the context of recurring spontaneous transient events.

Results: After showing the limitations of Transfer Entropy and Dynamic Causal Strength in this setting, we introduce a novel measure, relative Dynamic Causal Strength, and provide theoretical and empirical support for its benefits.

Discussion: These methods are applied to simulated and experimentally recorded neural time series and provide results in agreement with our current understanding of the underlying brain circuits.

KEYWORDS

information theory, causal strength, graphical models, transfer entropy, structural equations, neural oscillations

1 Introduction

During both wakefulness and sleep, the mammalian brain is able to implement numerous functions key to our survival with extraordinary reliability. This implies precise coordination of transient *mechanisms* at multiple spatiotemporal scales ensuring both the synergy between brain regions contributing to the same task, and the non-interference between network activities in charge of different functions. Evidence for such transient mechanisms is provided by the variety of neural events that can be observed in brain activity across multiple structures. Such phenomena may occur in response to stimuli, as has been observed for gamma oscillations (Tallon-Baudry and Bertrand, 1999; Fries, 2015), and may play a role in the dynamic encoding of information. However, key phenomena can also occur spontaneously, as exemplified by the variety of events occurring during sleep. These include Sharp Wave-Ripples (SWR) complexes that occur in the hippocampus during the same sleep stages, and take the form of a slow deflection (the sharp wave, SW) superimposed with a fast short-lived oscillation (the ripple). SWRs have been extensively studied and a large set of evidence supports their key

role in episodic memory consolidation and the recall of previous experiences (Lee and Wilson, 2002; Diba and Buzsaki, 2007; Ego-Stengel and Wilson, 2010).

In order to understand how these transient phenomena operate mechanistically, causality measures based on observed neural time series can be very helpful to quantify the underlying transient influences between brain structures. Several measures of causality have been proposed, starting in the econometrics literature with Granger causality (GC) (Granger, 1969), relying on vector auto-regressive models. This measure can be generalized to an information-theoretic quantity: Transfer Entropy (TE) (Schreiber, 2000). In the present work, we focus on "model-free" quantities such as TE that are defined independently of the specific functional relationships entailed by a particular model of the dynamics. TE and GC have been used to assess the significance of causal links, but also the "strength" of these links. However, whether they are appropriate quantities to measure such strength is debated (Janzing et al., 2013; Stokes and Purdon, 2017).

Structural Causal Models (SCM, see Supplementary Section SA for background) also allow causal strength measures to be evaluated by their ability to reflect the magnitude of the changes resulting from removing the causal links. In this context, the relevance of causality measures has been investigated by Ay and Polani (2008), who discuss how to account for the effect of knockout experiments, and introduce a measure of *information flow*, emphasizing its desirable properties; Janzing et al. (2013) provide interesting theoretical justifications for this kind of measure and extend it to define *causal strength* (CS) of an arbitrary set of arrows in a graphical model. With respect to TE, information flow and CS have the benefit to be local, in the sense that they depend only on the direct causes of the observed effects and their associated mechanisms. This makes CS a good candidate to measure transient connectivity changes during non-stationary neural events, as they would be able to restrict themselves to causal influences that take place at a specific time, associated to specific arrows in the "unrolled" causal graph describing time-varying interactions.

However, we will argue that, even in simple unidirectional settings where a "source" region is driving events in a target region, such TE and CS may not reflect well a key element for neuroscientists: the role played by transient dynamics. Based on the potential outcome framework (Rubin, 1974), causal reasoning has also been used to provide intuitive measures of the *causal impact* of a specific phenomenon happening at a given time point (Brodersen et al., 2015), by comparing it to a putative scenario where this phenomenon does not happen and called *synthetic control* in economics (Abadie, 2021). This inspired us to take into account the peri-event change of signals compared to a pre-event stage as another component of causal influence.

Therefore, we look at causal influences through the lens of interventions in SCMs to propose a principled quantification of the strength of causal interactions in peri-event time series, i.e., datasets collected specifically around the times of occurrence of an identified phenomenon in neural signals. Based on information theoretic analyses, we assess the relevance and issues raised by a time-varying implementation of GC, TE and causal strength (DCS, where D stands for *Dynamic*), and extend DCS to a novel measure, the relative DCS (rDCS), to quantify causal influences reflected by both the connectivity and the event-related change at the cause. We show theoretically that rDCS is effective in uncovering dynamic causal influences for task-dependent events that are often accompanied with a deterministic component, as well as for spontaneous events. We also demonstrate how choices made for aligning peri-event time series collected across multiple occurrences of these events may bias causality measures, and we propose a way to align the detected events favoring the recovery of the ground truth causal direction for a uni-directionally coupled system. The benefits of rDCS over TE and DCS is demonstrated by both simulated toy models and neurophysiological recordings of SWRs. Overall, our results suggest that rDCS helps better quantify the causal interactions between transient dynamical events, and thus uncover elementary mechanisms that shape brain activities.

2 Methods

2.1 General principles for the analysis of event-related causal interactions

2.1.1 Structural Causal Models (SCM)

Mathematically, an SCM for variables $\{V_1, \ldots, V_d\}$ is a collection of so-called "structural assignments" of the form

$$V_j := f_j\left(\mathbf{PA}_j, N_j\right), \quad j = 1, \ldots, d. \tag{1}$$

where the right hand side function f_j determines the assignment of the value of V_j on the left-hand side based on the values of parents variables $\mathbf{PA}_j \subset \{V_1, \ldots, V_d\}$ and of the so-called exogenous random variable N_j. The SCM is associated to a directed graph, the *causal graph*, where each variable V_j is represented by a node, and the parent-child relations between them is indicated by a "parent → child" arrow. While SCMs do not necessary include time information, we can exploit them to study dynamic interactions between two subsystems by considering causal relations linking variables representing one subsystem at past time points, to variables representing the other subsystem at a future time point. As an example, Figure 1A shows two uni-directionally coupled brain regions whose activities are represented by time series $\{\ldots, X_t^1, X_{t+1}^1, \ldots\}$ and $\{\ldots, X_t^2, X_{t+1}^2, \ldots\}$ and the corresponding SCM links the past of X^2 to the future of X^1. Typically, such a model also includes dependencies of each regional activity on its own past, and those dependencies can involve multiple time steps, leading to causal graphs of the form exemplified in Figure 2A. We focus on information theoretic causality measures, which are typically "model free", in the sense that they can be expressed independently from the choice of functions f_j in the assignments of Eq. 1. However, model-free estimation of information theoretic quantities is a challenging problem that we will not address in this paper. Instead, estimation of the relevant quantities will rely on the following linear time-inhomogeneous Structural Vector Autoregressive (SVAR) model.

$$X_t^1 := \mathbf{a}_t^\top \boldsymbol{X}_{p,t}^1 + \mathbf{b}_t^\top \boldsymbol{X}_{p,t}^2 + \eta_t^1, \quad \eta_t^1 \sim \mathcal{N}\left(k_t^1, \sigma_{1,t}^2\right), \tag{2a}$$

$$X_t^2 := \mathbf{c}_t^\top \boldsymbol{X}_{p,t}^1 + \mathbf{d}_t^\top \boldsymbol{X}_{p,t}^2 + \eta_t^2, \quad \eta_t^2 \sim \mathcal{N}\left(k_t^2, \sigma_{2,t}^2\right). \tag{2b}$$

where:

- η_t^k is the innovation for channel k at time t, sampled independently from the past values of X and from innovations at other time points and/or channels,
- $\boldsymbol{X}_{p,t}^k = [X_{t-1}^k, \ldots, X_{t-p}^k]^\top$ is the vector collecting past p samples of the time series,

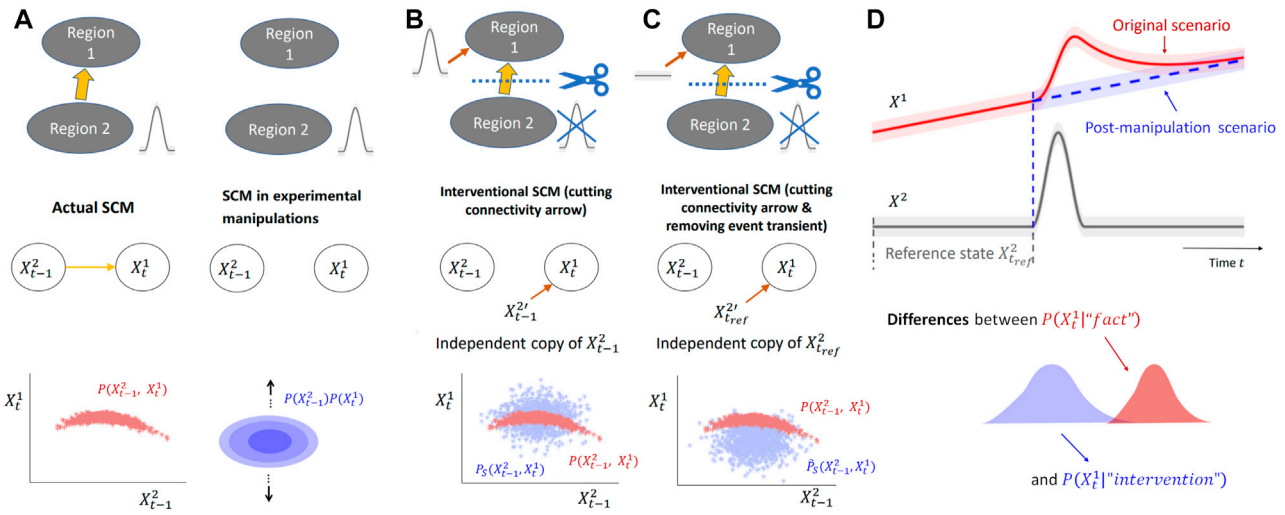


FIGURE 1
Analysis of event-related causality via interventions in SCMs. **(A)** (Top) Diagrams representing two brain regions with uni-directional connectivity from Region 2 to Region 1 and the post-manipulation scenario where the connectivity is removed. Region 2, as the "cause region", exhibits transient events (grey) that influence Region 1 by propagating along the anatomical connection. (middle) SCMs underlying the diagrams, where X^1_t and X^2_t denotes states of Region 1 and Region 2. (bottom left) joint distribution of the two nodes in the SCM above reflecting dependencies between them. (bottom right) Hypothetical Gaussian joint distribution of two nodes in the SCM above with black arrows indicating varied covariance. **(B)** (Top) An experimental manipulation of the two-region diagram in **(A)** related to the intervention in measuring causal strength: cutting the anatomical connectivity. (middle) A corresponding intervention of the SCM in **(A)** represents cutting the causal arrow and feeding the effect node X^1_t with an independent copy of the cause node X^2_{t-1}. (bottom) joint distribution of the corresponding intervention distribution in contrast to the joint distribution in (A, bottom left). **(C)** (Top) Another experimental manipulation of the two-region diagram in **(A)**: cutting the anatomical connectivity and removing the event-based signal changes at Region 2. (Bottom) The corresponding intervention of the SCM in **(A)** represents cutting the causal arrow and feeding the effect node X^1_t with an independent copy of a reference state of the cause node $X^2_{t_{ref}}$. (bottom) joint distribution of the corresponding intervention distribution in contrast to the joint distribution in (A, bottom left). **(D)** (Top) A time course of observed peri-event signals of Region 1 (X^1_t, red) and Region 2 (X^2_t, grey) reflecting the original scenario. The blue dashed time course represents the post-intervention scenario where X^1_t evolves without the influence from X^2_t. The interval marked by grey dashed lines refers to the reference state before the occurrence of events in X^2_t. (Bottom) Visualization of the difference between the original and post-intervention scenarios at each time point.

- the t subscript in all parameters ($\mathbf{a}_t$, $\mathbf{b}_t$, $\mathbf{c}_t$, $\mathbf{d}_t$, k^1_t, k^2_t, $\sigma^2_{1,t}$, $\sigma^2_{2,t}$) comes from our time-inhomogeneity assumptions and is not standard in the GC literature.

Note that the constraint of independence between exogenous variables η^k_t associated to different channels entails the assumption of no contemporaneous effects, in contrast to, e.g., Moneta et al. (2011). Figure 2A exemplifies the associated causal graph for $p = 2$ and $c_t = 0$, such that X^2 is dependent only on X_2 itself, encodes unidirectional causation from X^2 to X^1. The methods presented in this paper can be applied to time series generated by any other Markovian time series model (e.g., non-linear models (Marinazzo et al., 2011a)). However, the choice of SVAR 1) allows expressing the link between Granger causality and Transfer Entropy, 2) facilitates the estimation of all information theoretic quantities (which is inherently hard in a model-free setting, see, e.g., McAllester and Stratos (2020), as they get an analytic expression based on second order statistics), 3) avoids issues related to the non-parametric estimation of information theoretic quantities (e.g., finite sample bias), 4) allows to build easily interpretable models of transient oscillations.

2.1.2 Interventions in SCMs

One key question in causality is estimating the effect of (possibly imaginary) *manipulations* (see, e.g., Woodward, 2001) of the system of interest from data, which boils down to comparing two "worlds" or scenarios (Shpitser and Pearl, 2008): the original world where no manipulation is performed, and the "post-manipulation" world.

Both original and post-manipulation worlds typically cannot be measured simultaneously (e.g., "treatment" and "no treatment" in the same patient). However, estimating their differences arguably forms the basis of causal investigations in empirical sciences, for example, by performing randomized experiments on multiple instances of a system designed with mutually exclusive treatments to infer the outcome of manipulations of this system. However, even performing carefully controlled experiments on close to identical instances of a system is often challenging in reality, as many physical and physiological phenomena cannot be easily reproduced or manipulated. This is typically the case for *spontaneous* transient neural events investigated in this paper, where neurophysiological experimental techniques limit the understanding and control of their conditions of occurrence, as well as the ability to precisely modify some aspects of network activity to test assumptions on the underlying mechanisms.

Under additional model assumptions, such as the absence of unobserved confounders, the framework of SCMs (as briefly introduced in Section 2.1.1 and Supplementary Section SA), can be leveraged to infer the outcome of manipulations based on observational data only. Assuming those assumptions are met (see also Discussion for examples), the SCM inferred from data

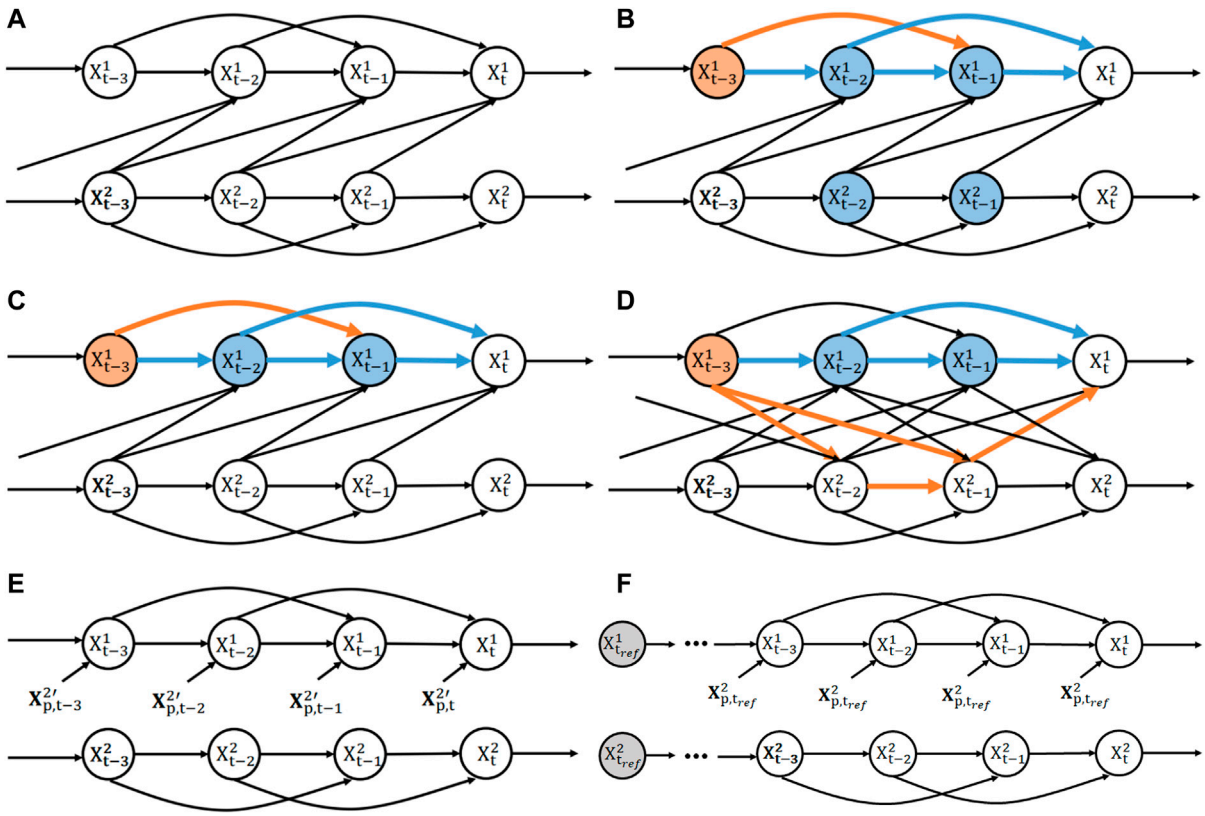


FIGURE 2
D-separation of bi-variate SVAR(2) model. **(A)** Structural causal model of a bi-variate SVAR(2) model defined in Eq. 2a with uni-directional coupling from X^2 to X^1. **(B)** Conditioning on both past states of X^1 and X^2 blocks all paths from X^1_{t-3} to X^1_t. Blue nodes represents conditioned nodes while blue arrows marks blocked paths. Orange arrows marks the unblocked paths. **(C)** Conditioning on past states of X^1 alone blocks all paths from X^1_{t-3} to X^1_t in the uni-directional case. Color codes are the same as **(B)**. **(D)** Conditioning on past states of X^1 alone does not block all paths from X^1_{t-3} to X^1_t in the bi-directional case. Color codes are the same as **(B)**. **(E)** The intervention implemented in devising CS is to break the causal arrows and send an independent copy $\boldsymbol{X}^{2}_{p,t}$ to X^1_t at each time point. This diagram applies to both CS and DCS (Section 2.2.3). **(F)** The intervention implemented in devising rDCS is to break the causal arrows and send an independent copy of the stationary state $\boldsymbol{X}^{2}_{p,t_{ref}}$ (marked by grey) to X^1_t at each time point.

can be modified using a family of operations named *interventions* to model the manipulations of the system described by the SCM (Pearl, 2000; Peters et al., 2017). Intervening typically refers to modifying the structural equation associated to one node in the SCM, to study the modifications it brings about in the system. When interventions are performed, the only affected mechanistic relations (represented by arrows in an SCM) are the ones between the intervened nodes and their parent nodes. For instance, one can impose a fixed deterministic value to a node, or that this node's variable is drawn from a given distribution, independently from other variables in the SCM (Janzing et al., 2013; Correa and Bareinboim, 2020; Peters et al., 2017, Chapter 3). Both such interventions lead to an intervened causal graph where the arrows between the node intervened upon and its parents are removed. For a better understanding of the manipuation modelled in these two cases, consider examples from experimental manipulations in electrophysiology. The first case involves fixing a value, as in voltage clamp techniques used to study channel conductance. By fixing the membrane potential of a single neuron, the causal arrow between the membrane potential and extracellular ion concentration is broken. An example of the second case is injecting current in plasticity studies to maintain certain firing rates in a patch-clamped neuron, ensuring that the pharmacological shutdown of certain ion channels does not cause a change in neuronal firing patterns.

Importantly, while an intervention modifies the graph associated to an SCM, the variables' joint distribution can still be obtained by exploiting interventional knowledge, (unintervened) observational data and prior assumptions related to the unaffected conditionals.

For the SCM introduced in Eq. 1, intervening on V_k consists in replacing its structural assignment by a new one:

$$V_k := \tilde{f}_k\left(\widetilde{\mathbf{PA}_k}, \widetilde{N_k}\right). \tag{3}$$

where the new function $\widetilde{f_k}$, set of parents $\widetilde{\mathbf{PA}_k}$, and/or the distribution of the exogenous variable $\widetilde{N_k}$ may be differ from the original ones. The resulting distribution $\tilde{P}_V$ is called *intervention distribution* and denoted $P_V^{\mathrm{do}(V_k := \tilde{f}_k(\widetilde{\mathbf{PA}_k}, \widetilde{N_k}))}$ (see e.g., Peters et al. (2017, chapter 6)), where the superscript indicates that we refer to the distribution resulting from the modification of the SCM's kth equation by Eq. 3, which is called a "do operation". Meanwhile, the other structural equations and the distribution of their associated exogenous variables are kept unchanged.

2.1.3 From causal strength to a measure of event-based causal influence

Typically, one brain region influences another through axonal propagation of spiking activity in afferent neurons, and synaptic transmission to dendrites of the target region. In the simplest bivariate case (i.e., if we focus on two brain regions with direct synaptic projections), this relationship can be represented in an SCM by Figure 1A(left) where two nodes representing the neural activities of the two regions are linked by a uni-directional arrow. For the sake of simplicity, we ignore for now the influence of the past of X^2 on itself, such that the only causal link is from the past (at $t-1$) of region 2 to the present (at t) of region 1. The joint probability can be causally factorized as $P(X^2_{t-1}, X^1_t) = P(X^1_t|X^2_{t-1})P(X^2_{t-1})$, where $P(X^1_t|X^2_{t-1})$ reflects the stochastic map or causal mechanism between the child and parent of the arrow: for example, Figure 1A(bottom left) shows a typical example how a nonlinear structural equation (i.e., the arrow) induces dependencies between a normally-distributed X^2_{t-1} and X^1_t.

A natural manipulation to study the causal mechanism is to shut down the synaptic transmission during event occurrence and compare the outcomes, e.g., via physically cutting the pre-synaptic dendrite attached to the synapse, or pharmacological blockade of the relevant ion channel. Suppose that the experiment could be done, the data obtained in this hypothetical manipulated scenario can be modeled by another SCM without the arrow between X^2_{t-1} and X^1_t, as seen in Figure 1A(right middle), such that they are independent from each other with $\tilde{P}(X^2_{t-1}, X^1_t) = \tilde{P}(X^1_t)\tilde{P}(X^2_{t-1})$ due to the Markov properties (Supplementary Section SA). In a model-free setting and in absence of data where this (technically challenging) manipulation is actually done, the choice of factorized $\tilde{P}$, such as the marginal mean and covariances, is non-trivial (see illustration in Figure 1A(bottom right)). Janzing et al. (2013) introduce a well-grounded way to make this choice to emulate this experimental ablation of connectivity, that generalize to arbitrary causal graphs.

This approach will be thoroughly discussed in Section 2.2.3 but here we explain it briefly in this simplified example to provide the readers with an intuitive understanding of the principles. Figure 1B illustrates the intervention performed in the SCM: cut the causal arrow from X^2_{t-1} to X^1_t and feed X^1_t with an independent copy of X^2_{t-1} (denoted $X^{2'}_{t-1}$), where *independent copy* means that $X^{2'}_{t-1}$ is a random variable statistically independent from any exogenous variables within the graph, and with the same marginal distribution as X^2_{t-1}. The idea of this intervention is to achieve the independence of variables in the post-manipulation world by exploiting the observational conditional and marginal probabilities available to us. That is:

- choose $\tilde{P}(X^2_{t-1})$ to be the observed marginal distribution of the cause, i.e., $\tilde{P}(X^2_{t-1}) = P(X^2_{t-1})$, because it is the only one available to us in a model-free setting. Consider alternative choices: we could set the cause to a constant, but which constant to choose is unclear without additional information on the system. For example, even taking the observational average would not be realistic if X^2_{t-1} is binary,
- replace the cause → effect mechanism $P(X^1_t|X^2_{t-1})$ by the operation of feeding the effect node X^1_t with an independent copy of the cause node X^2_{t-1} at the same time $t-1$, such that the mechanism becomes $\tilde{P}(X^1_t) = \int P(X^1_t|X^{2'}_{t-1})P(X^{2'}_{t-1})dX^{2'}_{t-1} = P(X^1_t)$. Here again, the choice of observational density guarantees that the resulting mechanism is well defined for arbitrary SCMs.

The strength of the causal arrow is then quantified by the Kullback-Leibler (KL) divergence D_{KL} between original (unintervened) and intervened joint densities $D_{KL}(P\|\tilde{P}) = D_{KL}(P(X^2_{t-1}, X^1_t)\|P(X^2_{t-1})P(X^1_t))$, which in this simple case thus boils down to the mutual information between the two nodes. Figure 1B(bottom) illustrates the contrast between the actual joint distribution and the intervention distribution for such intervention: by comparing these two distributions one could quantify how much the causal mechanism tilts the Gaussian shape, while they still largely overlap. In a context where nodes correspond to single neurons, this can be thought of as a proxy for the experiment of cutting the axon of afferent neurons, while injecting a current to maintain the baseline level of excitation in the target neuron, such that it is kept in naturalistic conditions.

However, we will argue that this choice of intervention is not an ideal way to measure the event-based causal influences between two brain regions. Going back to the manipulation experiment, despite the transient activities occurring at the pre-synaptic neurons as an input to the synapse, due to the cutoff of afferents or blockage of ion channels, the activity of the post-synaptic neuron (i.e., the effect variable) is expected to remain at baseline level without being influenced by the event occurring in the cause. In this context, the operation of feeding the effect node X^1_t with an independent copy of the cause node X^2_{t-1} at the same time $t-1$ still implicitly incorporates the influence of the event-related transient changes undergone by X^2 at the time $t-1$ on X^1_t, as the distribution of X^2_{t-1} may strongly differ from what it is during baseline activity (without the occurrence of events). Therefore we propose instead to reconstruct the baseline state by replacing the independent copy of the marginal distribution of the event-related activity X^2_{t-1} by the marginal distribution of a baseline state in $X^2_{t_{ref}}$, at a reference time point t_{ref} where the event of interest has not yet occurred (Figure 1D(Top)), such that the new intervention distribution becomes $\tilde{P}(X^2_{t-1}, X^1_t) = P(X^2_{t-1}) \int P(X^1_t|X^{2'}_{t_{ref}}) P(X^{2'}_{t_{ref}})dX^{2'}_{t_{ref}} = P(X^2_{t-1}) \tilde{P}(X^1_t)$. The difference between feeding independent copies of different marginal distributions and the resulting baseline joint probability are illustrated in Figures 1B, C. We thus argue that this is a better reference scenario for testing the influence of an event between two brain regions because it accounts for the event-related variations of the input distribution relative to baseline activity.

2.2 Candidate time-varying causality measures

We now present the time-varying versions of commonly-adopted causal strength measures of a given direction of causation $X^2 \to X^1$ and discuss their properties in the context of transient event-based causality analysis, in light of the above principles. The candidate measures include time-varying extensions of Granger causality (GC), Transfer Entropy (TE) and Causal Strength (CS) (Janzing et al., 2013). To make the comparison

quantitative, time series are modeled using the bivariate linear SVAR model of Eq. 2a and Eq. 2b. As we will see, all measures boil down to comparing the "full" bivariate model to a model where the contribution of the cause time series to the effect is removed in some way. Generalization to more than two observed time series is possible in all cases and briefly mentioned for each approach.

2.2.1 Granger causality

Granger causality (GC), as well as its information-theoretic extension, Transfer Entropy (TE) is rooted in Wiener's principle of causality. For the bivariate case, Granger (1969) defines the statement of (Granger-)causality from X^2 to X^1 whenever knowledge of $\boldsymbol{X}^2_{p,t}$, in addition to $\boldsymbol{X}^1_{p,t}$, yields a strictly better prediction of X^1_t. This can be interpreted as a comparison between two prediction scenarios:

- Scenario 1: predict X^1_t using both $\boldsymbol{X}^1_{p,t}$ and $\boldsymbol{X}^2_{p,t}$,
- Scenario 2: predict X^1_t using only $\boldsymbol{X}^1_{p,t}$,

where $\boldsymbol{X}^1_{p,t}$ and $\boldsymbol{X}^2_{p,t}$ refer to the respective p previous past points of each time series, without further specification, such that in our notation p can be potentially infinite.

The predictive model describing the first scenario is referred to as the *full model* (Geweke, 1984) and corresponds to Eq. 2a of the SVAR model, where the first variable X^1 is dependent on both variables X^1 and X^2. An estimate of the innovation variance of X^1_t ($\sigma^2_{1,t}$ in Eq. 2a) is the mean squared residual error ($\hat{\sigma}^2_{1,t}$) of the forecast of X^2_t under the assumption that both $\boldsymbol{X}^1_{p,t}$ and $\boldsymbol{X}^2_{p,t}$ contribute to X^1_t. Under Scenario 2 where X^1_t is predicted only by $\boldsymbol{X}^1_{p,t}$, we have a *reduced model*

$$X^1_t = \mathbf{a}'^{\top}_t \boldsymbol{X}^1_{p',t} + \eta'^1_t, \quad \eta'^1_t \sim \mathcal{N}\left(k'^1, \sigma'^2_{1,t}\right). \tag{4}$$

where the model order p', the coefficient $\mathbf{a}'$, the innovations mean k'^1 and innovations variance σ'^2_1 are different from the corresponding terms in Eq. 2a and are classically re-estimated.

If X^2 Granger-causes X^1, then the full model should fit the data more accurately compared to the reduced model as measured by the estimated variance $\widehat{\sigma'}^2_{1,t}$, which should be larger than $\hat{\sigma}^2_{1,t}$. Then the amount of Granger causality can be defined as the log ratio of the residual variance between the reduced model and the full model, which leads to estimating the magnitude of Granger causality as

$$\mathrm{GC}\left(X^2_t \to X^1_t\right) = \frac{1}{2}\log\left(\frac{\widehat{\sigma'}^2_{1,t}}{\hat{\sigma}^2_{1,t}}\right), \tag{5}$$

where the factor 1/2 is chosen for consistency with TE (see Section 2.2.2). While the above linear SVAR model is the most widely used, Granger causality has been extended to non-linear models following the same predictive approach (e.g., Marinazzo et al., 2008; Marinazzo et al., 2011b; Diks and Wolski, 2016; Wismüller et al., 2021, for a recent review see Shojaie and Fox, 2022).

2.2.2 Transfer Entropy

TE is an information-theoretic implementation of Wiener's principle, where a comparison between the prediction performance of the above two scenarios is quantified with conditional entropy. TE quantifies to which amount X^2 causes X^1 in the Granger sense and is defined by the entropy difference

$$\mathrm{TE}\left(X^2_t \to X^1_t\right) = H\left(X^1_t|\boldsymbol{X}^1_{p,t}\right) - H\left(X^1_t|\boldsymbol{X}^1_{p,t}, \boldsymbol{X}^2_{p,t}\right). \tag{6}$$

Interestingly, using the Kullback-Leibler (KL) divergence D_{KL} between two probability densities $D_{KL}(p\|q) = \int p(x)\log\frac{p(x)}{q(x)}dx$, TE can be rewritten as an expected KL-divergence between the corresponding conditional probabilities, thereby contrasting the two above mentioned scenarios:

$$\mathrm{TE}\left(X^2_t \to X^1_t\right) = \mathbb{E}_{\boldsymbol{X}^1_{p,t}, \boldsymbol{X}^2_{p,t}}\left[D_{KL}\left(p\left(X^1_t|\boldsymbol{X}^1_{p,t}, \boldsymbol{X}^2_{p,t}\right)\|p\left(X^1_t|\boldsymbol{X}^1_{p,t}\right)\right)\right]. \tag{7}$$

As noticed by Barnett et al. (2009), under stationary Gaussian SVAR assumptions the analytic expression of Gaussian entropy applied to Eq. 6 leads to $\mathrm{GC}(X^2_t \to X^1_t) = \mathrm{TE}(X^2_t \to X^1_t)$ in the limit of unbiased variance estimation, such that TE appears as a strict generalization of GC, and can be estimated by GC in the context of Gaussian SVAR models. TE and GC statistics are two commonly used measures of causal strength for investigating interactions between brain regions (e.g., Vicente et al., 2011; Besserve et al., 2010, 2015; Ding et al., 2006; Wibral et al., 2013, 2014; Barrett et al., 2010; Wen et al., 2013; Shorten et al., 2021; Cekic et al., 2018, with several widely applied toolboxes such as Barnett and Seth, 2014; Montalto et al., 2014; Lizier, 2014; Wollstadt et al., 2019). They generalize easily to more than two signals by including also the past of additional signals in the prediction equations when assessing causality for a specific pair, as is done in conditional pairwise Granger causality (Barrett et al., 2010; Faes et al., 2011; Runge et al., 2012; Barnett and Seth, 2014). Based on the observational conditional distribution of the neural signals being analyzed, these two measures estimate a quantity that is easily interpretable from a forecasting perspective. However, they have some limitations with regard to their interpretability as interventions in the SCM framework and in the time varying setting that interests us in the present paper.

A key issue is that the reduced model ignores but *does not remove* the influence of past values of $\boldsymbol{X}^2$ ($\boldsymbol{X}^2_{p,t}$) on $\boldsymbol{X}^1_t$ by marginalizing with respect to them. It can be shown that such change does not preserve the SCM structure, and leads to violations of the Markov properties due to the implicit dependency on the mechanisms relating $\boldsymbol{X}^2_{p,t}$ and $\boldsymbol{X}^1_{p,t}$, which manifest themselves through the $p(\boldsymbol{X}^2_{p,t}|\boldsymbol{X}^1_{p,t})$ term of the marginalization equation (Ay and Polani, 2008; Janzing et al., 2013):

$$p\left(X^1_t|\boldsymbol{X}^1_{p,t}\right) = \int p\left(X^1_t|\boldsymbol{X}^1_{p,t}, \boldsymbol{X}^2_{p,t}\right)p\left(\boldsymbol{X}^2_{p,t}|\boldsymbol{X}^1_{p,t}\right)\mathrm{d}\boldsymbol{X}^2_{p,t}. \tag{8}$$

As a consequence, the reduced model cannot be generally interpreted as an intervention on the original SCM that would result in a model where arrows associated to the causal influence of interest would be removed. In addition, in case of bi-directional coupling, the reduced model of Eq. 4 is misspecified (in a generic case) for any finite order. This can be seen easily by exploiting the d-separation criterion (Supplementary Section SA), as illustrated in Figure 2. Figure 2B shows the estimation in the full model, where conditioning on both $\boldsymbol{X}^1_{p,t}$ and $\boldsymbol{X}^2_{p,t}$ blocks all the paths from X^1_{t-3} to X^1_t such that X^1_{t-3} and X^1_t are conditionally independent. For such a uni-directionally-

coupled system, a finite order for the reduced model also guarantees such conditional independence, as seen in Figure 2C where all paths are blocked by conditioning. However, in the same system with bi-directional coupling, for any $k > p$ (i.e., $k > 2$), there is always a path from X^1_{t-k} to X^1_t going through nodes of X^2 that is unblocked by $(X^1_{t-p}, \ldots, X^1_{t-1})$. As Figure 2D shows, 2 paths from X^1_{t-3} to X^1_t are not blocked by conditioning on $\boldsymbol{X}^1_{p,t}$. Under faithfulness assumptions, this implies that there is conditional dependence between X^1_t and its remote past samples, no matter how many finite past states we are conditioning on. This further implies that to minimize the forecast error of X^1_t in the reduced model one should ideally exploit the past information of this time series up to $p = +\infty$.

This issue has been both raised and addressed in the literature, in particular by resorting to Autoregressive Moving Average models and state space models for defining an appropriate reduced model (e.g., (Barnett and Seth, 2015; Solo, 2016)). However, this remains an important limitation when extending TE to time-varying versions, where the model is assumed to be stationary at best locally in time. For example, when defining a non-stationary SVAR model as Eq. 2a, we assume a different linear model in each 1-point time window. The non-locality of TE is particularly problematic for such a time-varying model assumption because of the implicit influence of past activities on this quantity.

2.2.3 Dynamic causal strength

To overcome the limitations of TE and GC, Ay and Polani (2008) have proposed a measure of *information flow* to quantify the influence of some variables on others in a system, which has been further studied and generalized in Janzing et al. (2013) as a measure of the *Causal Strength* (CS) of an arbitrary set of arrows in a graphical model. In the present paper, we define CS in the context of time-inhomogeneous vector autoregressive processes and their associated unrolled causal graph, and thus call it *Dynamic Causal Strength* (DCS).

DCS can be naturally defined using the SCM interventional formalism (Pearl (2000); Peters et al. (2017), see Section 2.1.2). Briefly, interventions are performed on nodes in order to remove the specific arrows from the causal graph whose influence we wish to quantify. In agreement with Ay and Polani (2008) and Janzing et al. (2013), in the context of inhomogeneous SVAR models (as illustrated in Figure 2A), an appropriate intervention to remove the causal influence from $\boldsymbol{X}^2_{p,t}$ to X^1_t can be designed as the following intervention (shown in Figure 2E): *remove the arrow* $\boldsymbol{X}^2_{p,t} \to X^1_t$ *by injecting instead* $\boldsymbol{X}^2_{p,t}{}'$, *an independent copy of* $\boldsymbol{X}^2_{p,t}$ *with the same joint distribution, in the original mechanism* $P(X^1_t|\boldsymbol{X}^1_{p,t}, \boldsymbol{X}^2_{p,t})$. The intervention distribution p^{DCS} models the post-interventional world after removing the causal arrow from $\boldsymbol{X}^2_{p,t}$ to X^1_t and results in the entailed conditional probability

$$p^{DCS}\left(X^1_t|\boldsymbol{X}^1_{p,t}\right) = p^{do\left(X^1_t := f\left(\boldsymbol{X}^1_{p,t}, \boldsymbol{X}^2_{p,t}{}', \eta^1_t\right)\right)}\left(X^1_t|\boldsymbol{X}^1_{p,t}, \boldsymbol{X}^2_{p,t}\right) = \int p\left(X^1_t|\boldsymbol{X}^1_{p,t}, \boldsymbol{X}^2_{p,t}\right) p\left(\boldsymbol{X}^2_{p,t}\right) \mathrm{d}\boldsymbol{X}^2_{p,t},$$

which does not depend on $p(\boldsymbol{X}^2_{p,t}|\boldsymbol{X}^1_{p,t})$ anymore, in comparison to Eq. 8. DCS then quantifies the KL divergence between the distributions of $X^1_t|(\boldsymbol{X}^1_{p,t}, \boldsymbol{X}^2_{p,t})$ obtained in both worlds, such that

$$\mathrm{DCS}(X^2_t \to X^1_t) = \mathbb{E}_{\boldsymbol{X}^1_{p,t}, \boldsymbol{X}^2_{p,t}}\left[D_{KL}\left(p\left(X^1_t|\boldsymbol{X}^1_{p,t}, \boldsymbol{X}^2_{p,t}\right) \mid p^{DCS}\left(X^1_t|\boldsymbol{X}^1_{p,t}\right)\right)\right]. \tag{9}$$

A parametric formulation under linear Gaussian model assumptions is given in Supplementary Section SD.5. Generalization to more than two time series is also straightforward following (Janzing et al., 2013): p^{DCS} and DCS are simply computed by also including conditioning on the past of all other time series, in addition to $\boldsymbol{X}^1_{p,t}$.

Remark: In contrast with Janzing et al. (2013), but in line with Ay and Polani (2008), we do not use jointly independent copies of each component of $\boldsymbol{X}^2_{p,t}$, that is, the copy preserves the dependency between the successive past time points of X^2. Indeed, Janzing et al. (2013) require having copies with jointly independent components in order to assess the individual strength of each arrow in the causal graph, which would correspond to the influence of each time lag in our setting. In contrast, this is not a requirement for us as we are only interested in assessing the overall effect of the whole past of a given time series on the another. One benefit of our choice is that it is consistent with the definition of TE: one can easily check that in absence of dependency of X^1_t on its own past, both TE (based on Eq. 6) and DCS reach the same value: the mutual information of $\boldsymbol{X}^2_{p,t}$ and X^1_t. Given that successive samples may be strongly correlated in practice, our choice avoids unnecessary discrepancies between these two measures to focus on their key difference. In additional, our choice can be seen as in line with Janzing et al. (2013) when considering a state representation of the time series' causal graph, where the node of variable k at time t would be the vector $[X^k_t, X^k_{t-1}, \ldots, X^k_{t-p+1}]^\top$.

2.3 Near deterministic behavior of TE and DCS

The analysis of transient neural events leads us to analyze signals that have limited stochasticity in several respects: on the one hand, strongly synchronized oscillatory signals can be represented by SVAR models with low innovation variance, relative to the variance of the measured signal. Moreover, when a study focuses on a reproducible type of transient pattern, it is often characterized by a waveform that has little variability across collected trials. Such a situation can be modeled with a time-varying deterministic innovation, exhibiting strong variation of its mean across time, but no or little variance. We investigate the theoretical properties of TE and DCS in this regime, showing a benefit of DCS with respect to TE, but also remaining limitations.

2.3.1 TE behavior for strongly synchronized signals

Besides, it has also been pointed out that the definition of TE in Eq. 7 has some other non-intuitive implications (Ay and Polani, 2008; Janzing et al., 2013). In particular, there are situations in which $\mathrm{TE}(X^2 \to X^1)$ almost vanishes, although the influence is intuitively clear. How frequent are the practical situations in which we have these detrimental effects is unclear; however, theoretical analysis

suggests that this can happen when the time series are strongly correlated.

To see this, we can derive from Eq. 7 the case where X^2 is a deterministic function of X^1 such that TE vanishes. Take the special case where X_t^2 is proportional to X_t^1 such that $X_t^2 = kX_t^1$, representing a time-wise synchronization of the two signals, the conditional variance will be

$$\begin{aligned}\Sigma_{\boldsymbol{X}_p^2|\boldsymbol{X}_p^1} &= \Sigma_{\boldsymbol{X}_p^2} - \Sigma_{\boldsymbol{X}_p^2\boldsymbol{X}_p^1}\Sigma_{\boldsymbol{X}_p^1}^{-1}\Sigma_{\boldsymbol{X}_p^1\boldsymbol{X}_p^2} \\ &= \Sigma_{\boldsymbol{X}_p^2} - k\Sigma_{\boldsymbol{X}_p^2}\cdot\left(\frac{1}{k^2}\Sigma_{\boldsymbol{X}_p^2}^{-1}\right)\cdot k\Sigma_{\boldsymbol{X}_p^2} = 0\end{aligned}$$

Plugging into Supplementary Equation S8 in Supplementary Section SD.4 yields,

$$\begin{aligned}\mathrm{TE}\left(X_t^2 \to X_t^1\right) &= \log\frac{\mathbf{b}_t^\top \mathrm{Cov}\left[\boldsymbol{X}_{p,t}^2|\boldsymbol{X}_{p,t}^1\right]\mathbf{b}_t + {\sigma_{1,t}}^2}{{\sigma_{1,t}}^2} \\ &= \log\frac{{\sigma_{1,t}}^2}{{\sigma_{1,t}}^2} = \log 1 = 0\end{aligned}$$

However, a strong correlation between two observed time series does not necessarily imply that causal interactions between them are weak, from an SCM perspective. We will investigate this case in Section 3.1 and compare with the results of DCS to show that DCS does not suffer from this non-intuitive vanishing problem.

2.3.2 Insensitivity of TE and DCS to deterministic perturbations

While several intuitive properties make DCS a good candidate to quantify causal influences, we exhibit a counterintuitive property common to TE and DCS in the context of peri-event time series. Transient neural events are mainly investigated in two types of analyses: 1) stimulus-triggered (or response-triggered) data that are temporally aligned by task (or response) onset and 2) event-triggered data where occurrences of a type of brain-activity pattern are detected along the time course of the recordings (manually or algorithmically) and used to create peri-event trials.

In both cases, neural activities are likely to have a deterministic component appearing in the peri-event ensembles, due the similarity of the response to successive stimuli in case 1), or due to the similarity of the neural patterns detected in the recordings in case 2). Here we will show that, in a linear setting, TE and DCS are insensitive to such a deterministic component. Specifically, TE and DCS values are unaffected by interventions on the innovations' mean at any time point.

First, we exhibit the role played by a deterministic perturbation in an example.

Example 1. *Consider the bi-variate SVAR(1) model in the following form*

$$X_t^1 := aX_{t-1}^1 + bX_{t-1}^2 + \eta_t^1, \tag{10a}$$

$$X_t^2 := \eta_t^2, \tag{10b}$$

with $a, b \neq 0$ and a stationary innovation for X^1, $\eta_t^1 \sim \mathcal{N}(0,1)$, but a non-stationary innovation for X^2, $\eta_t^2 \sim \mathcal{N}(\alpha\delta_{t,t_0}, \sigma_{2,t}^2)$, with

$$\delta_{t,t_0} = \begin{cases} 1, & \text{for } t = t_0, \\ 0, & \text{otherwise.}\end{cases}$$

When varying α, this models a intervention on the second time series. Then it can be easily shown that the expected time course of X^1 is

$$\mathbb{E}\left[X_t^1\right] = \begin{cases} \alpha b a^{t-t_0+1}, & t \geq t_0 + 1 \\ 0, & \text{otherwise.}\end{cases}$$

This witnesses the causal influence of $X_{t_0}^2$ on values of X_t^1 at subsequent times, which for large α results in large deviations from the baseline expectation of X_t^1 for t prior to t_0. Intuitively, one may expect that a quantification of the magnitude (strength) of the causal influence of X^2 on X^1 should be larger for larger α, as a transient of larger magnitude propagates from X^2 to X^1. From a neuroscientific perspective, this could model an experimental setting where one brain region is electrically stimulated with increasing strength to detect whether it is anatomically connected to another. Obviously, the magnitude of the stimulation is expected to be critical to elicit a response in the target region. However, TE and DCS actually turn out to be insensitive to such stimulation.

We will show this in the more general setting of the SVAR(p) model of Eq. 2a and Eq. 2b.

Proposition 1. *For linear SVAR models defined by Eq. 2a and Eq. 2b, TE and DCS measures are invariant to deterministic perturbations, i.e., to changes in the mean of the innovation's distributions* (k_t^1, k_t^2).

Proof. Without loss of generality, we will show invariance to an elementary intervention at a single time t_0 that transforms $\eta_{t_0}^2$ to $\eta_{t_0}^2 + \alpha$, which boils down to changing the mean parameters of the innovation $k_{t_0}^2$ in Eq. 2a and Eq. 2b. By linearity and symmetry of the problem for channel 1 and 2, invariance to deterministic perturbations results from combining several elementary interventions.

To compute how the intervention distribution of the new variables denoted $(\tilde{X}^1, \tilde{X}^2)$ changes with respect to the distribution of the original variables, we can examine the difference with respect to (X^1, X^2) that has the same innovations, except for $\eta_{t_0}^2$ for which we remove a constant α. (X^1, X^2) is then distributed according to the original distribution (before intervention), and the difference $(U, V) = (\tilde{X}^1 - X^1, \tilde{X}^2 - X^2)$ follows the equations

$$\begin{aligned}U_t &= \mathbf{a}^\top \mathbf{U}_{p,t} + \mathbf{b}^\top \mathbf{V}_{p,t}, \\ V_t &= \mathbf{c}^\top \mathbf{U}_{p,t} + \mathbf{d}^\top \mathbf{V}_{p,t} + \delta_{t,t_0},\end{aligned}$$

which is a set of linear deterministic difference equations with a unique solution making $\boldsymbol{X}$ and $\tilde{\boldsymbol{X}}$ coincide before the intervention[1] (U_t, V_t). As a consequence, by linearity, the interventional density $\tilde{p}$ is a shifted version of the original:

$$\tilde{p}\left(X_t^1, \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) = p\left(X_t^1 - U_t, \boldsymbol{X}_{p,t}^1 - \mathbf{U}_{p,t}, \boldsymbol{X}_{p,t}^2 - \mathbf{V}_{p,t}\right),$$

which implies the same for conditional marginal densities, e.g.,

$$\tilde{p}\left(X_t^1|\boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) = p\left(X_t^1 - U_t|\boldsymbol{X}_{p,t}^1 - \mathbf{U}_{p,t}, \boldsymbol{X}_{p,t}^2 - \mathbf{V}_{p,t}\right)$$

and

$$\tilde{p}\left(X_t^1|\boldsymbol{X}_{p,t}^1\right) = p\left(X_t^1 - U_t|\boldsymbol{X}_{p,t}^1 - \mathbf{U}_{p,t}\right).$$

As a consequence TE on the intervention distribution writes

1 Because initial conditions of this deterministic linear system are set to zero before the intervention at t_0

$$\mathrm{TE}\left(\tilde{X}_t^2 \to \tilde{X}_t^1\right) = \int \tilde{p}\left(X_t^1, \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) \log \frac{\tilde{p}\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right)}{\tilde{p}\left(X_t^1 | \boldsymbol{X}_{p,t}^1\right)} \, \mathrm{d}X_t^1 \mathrm{d}\boldsymbol{X}_{p,t}^1 \mathrm{d}\boldsymbol{X}_{p,t}^2$$
$$= \int p\left(X_t^1 - U_t, \boldsymbol{X}_{p,t}^1 - \mathbf{U}_{p,t}, \boldsymbol{X}_{p,t}^2 - \mathbf{V}_{p,t}\right)$$
$$\times \log \frac{p\left(X_t^1 - U_t | \boldsymbol{X}_{p,t}^1 - \mathbf{U}_{p,t}, \boldsymbol{X}_{p,t}^2 - \mathbf{V}_{p,t}\right)}{p\left(X_t^1 - U_t | \boldsymbol{X}_{p,t}^1 - \mathbf{U}_{p,t}\right)} \mathrm{d}X_t^1 \mathrm{d}\boldsymbol{X}_{p,t}^1 \mathrm{d}\boldsymbol{X}_{p,t}^2.$$

And by change of variable we get the invariance property:

$$\mathrm{TE}\left(\tilde{X}_t^2 \to \tilde{X}_t^1\right) f = \int p\left(X_t^1, \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) \log \frac{p\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right)}{p\left(X_t^1 | \boldsymbol{X}_{p,t}^1\right)} \, \mathrm{d}X_t^1 \mathrm{d}\boldsymbol{X}_{p,t}^1 \mathrm{d}\boldsymbol{X}_{p,t}^2$$
$$= \mathrm{TE}\left(X_t^2 \to X_t^1\right),$$

which can be generalized to arbitrary deterministic perturbations. The same reasoning can be applied to DCS leading to invariance as well (Supplementary Section SB) and this concludes the proof.

Arguably, this result is not what we would expect from an event-related measure of influence, because in the above example of Eq. 10a and Eq. 10b, setting a large α intuitively leads to a large influence of X^2 on X^1 provided $b \neq 0$. Provided that TE and DCS can be made arbitrarily small by reducing the innovation's variance $\sigma_{2,t}^2$ (according to their analytical expression in Supplementary Section SD), TE and DCS may detect no influence despite this strong effect on the mean of X_t^2. Although this invariance result is rigorously derived for linear SVAR models, it uncovers an issue for non-linear models as well, the magnitude of the causal influence associated to deterministic perturbation then depending chiefly on the non-linear properties of the system under study, and not on the magnitude of the changes triggered by the perturbation. Moreover, linear SVAR(p) models being able to approximate nonlinear dynamics, this suggests that deterministic causal influences cannot be detected by TE or DCS for a broad class of models in practice.

As elaborated above, this is in contrast to what would be expected in the neuroscientific context, and directly relates to the observational, event-related setting that we investigate: the deterministic component is due to the alignment of the data with respect to an event of interest, and we do not have a different condition to contrast the occurrence of this event with what would have happened in its absence. This analysis calls for building a synthetic baseline condition that would allow deterministic changes to be detected.

2.4 A novel measure: relative Dynamic Causal Strength

2.4.1 Motivation

Following the guidelines for event-based causality (presented in Section 2.1), we propose a novel measure, the relative Dynamic Causal Strength (rDCS), as a modification of DCS. This measure aims at taking into account the influence of event-based changes in the cause signals independent from the connectivity (the mechanism), and notably those driven by deterministic exogenous inputs. In the specific problem we are investigating, the cause is the past states of X^2, denoted $\boldsymbol{X}_{p,t}^2$, while the mechanism can be represented by the model in Eq. 2a and symbolized by the corresponding causal arrow in the causal graph. DCS only deletes the causal arrow in the post-intervention scenario but preserves the event-related change in the cause itself.

In the case where X^2 is driven by a deterministic exogenous input in a transient window, the cause exhibits significant changes relative to baseline; thus, intuitively, the causal effect should also be enhanced even if the causal arrow remains the same (i.e., the coefficient **b** stays unchanged). Apart from intervening on the causal arrow, further intervention can be implemented on the cause node to construct a post-intervention scenario where the cause receives no time-varying innovations. Therefore, inspired by causal impact (Section 2.1.3) which characterizes the difference between the current state and a baseline state, we propose (additionally to DCS) to replace the marginal of $\boldsymbol{X}_{p,t}^2$ by the marginal of $\boldsymbol{X}_{p,t_{ref}}^2$, that we denote $p_{ref}\left(\boldsymbol{X}_{p,t_{ref}}^2\right)$, for a reference time t_{ref}. The reference time t_{ref} is typically chosen to be a stationary period before the occurrence of the transient deterministic perturbations and statistics of $\boldsymbol{X}_{p,t_{ref}}^2$ can be averaged by statistics of $\boldsymbol{X}_{p,t}^2$ within this period. This leads to the *relative Dynamic Causal Strength* (rDCS)

$$\mathrm{rDCS}\left(X_t^2 \to X_t^1\right) = \mathbb{E}_{\boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2}\left[D_{KL}\left(p\left(X_t^1 | \boldsymbol{X}_{p,t}^2, \boldsymbol{X}_{p,t}^2\right) \middle\| p^{do\left(X_t^1 := f\left(\boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t_{ref}}^2, \eta_t^1\right)\right)}\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right)\right)\right] \tag{11}$$

with

$$p^{do\left(X_t^1 := f\left(\boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t_{ref}}^2, \eta_t^1\right)\right)}\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) = \int p\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) p_{ref}\left(\boldsymbol{X}_{p,t}^2\right) \mathrm{d}\boldsymbol{X}_{p,t}^2 \tag{12}$$

The implementation of rDCS given a SVAR model is derived in Supplementary Section SD.6. Generalization to more than two time series can be done in the same way as for DCS, by including extra conditioning on the past of other time series for all quantities.

Intuitively, the term *relative* originates from the comparison between the current past state $\boldsymbol{X}_{p,t}^2$ and the reference past state $\boldsymbol{X}_{p,t_{ref}}^2$. It is then natural to predict that in the uni-directional case, $\mathrm{rDCS}(X^2 \to X^1) = \mathrm{DCS}(X^2 \to X^1)$ for any reference time t_{ref} if X^2 is stationary because stationarity implies that the marginal distributions of $\boldsymbol{X}_{p,t_{ref}}^2$ and $\boldsymbol{X}_{p,t}^2$ are identical. As a particular case, this result implies that a transient loss of causal link from X^2 to X^1 will lead to rDCS = 0, while for a stationary bivariate system, DCS = rDCS is constant.

2.4.2 Sensitivity of rDCS to deterministic perturbations

The definition of rDCS implies sensitivity to deterministic perturbations. Indeed, taking the example in Section 2.3.2, the reference state $\boldsymbol{X}_{p,t_{ref}}^2$ is unaffected by the deterministic perturbation. Consequently, the translational invariance does not hold for the intervention distribution because

$$\mathrm{rDCS}\left(\tilde{X}_t^2 \to \tilde{X}_t^1\right) = \int \tilde{p}\left(X_t^1, \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) \log \frac{\tilde{p}\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right)}{\int \tilde{p}\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) \tilde{p}_{ref}\left(\boldsymbol{X}_{p,t}^2\right) \mathrm{d}\boldsymbol{X}_{p,t}^2}$$
$$\mathrm{d}X_t^1 \mathrm{d}\boldsymbol{X}_{p,t}^1 \mathrm{d}\boldsymbol{X}_{p,t}^2 = \int p\left(X_t^1 - U_t, \boldsymbol{X}_{p,t}^1 - \mathbf{U}_{p,t}, \boldsymbol{X}_{p,t}^2 - \mathbf{V}_{p,t}\right) \log \frac{p\left(X_t^1 - U_t | \boldsymbol{X}_{p,t}^1 - \mathbf{U}_{p,t}, \boldsymbol{X}_{p,t}^2 - \mathbf{V}_{p,t}\right)}{\int p\left(X_t^1 - U_t | \boldsymbol{X}_{p,t}^1 - \mathbf{U}_{p,t}, \boldsymbol{X}_{p,t}^2 - \mathbf{V}_{p,t}\right) p_{ref}\left(\boldsymbol{X}_{p,t}^2\right) \mathrm{d}\boldsymbol{X}_{p,t}^2}$$
$$\mathrm{d}X_t^1 \mathrm{d}\boldsymbol{X}_{p,t}^1 \mathrm{d}\boldsymbol{X}_{p,t}^2 = \int p\left(X_t^1, \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) \log \frac{p\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right)}{\int p\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) p_{ref}\left(\boldsymbol{X}_{p,t}^2 + \mathbf{V}_{p,t}\right) \mathrm{d}\boldsymbol{X}_{p,t}^2}$$
$$\mathrm{d}X_t^1 \mathrm{d}\boldsymbol{X}_{p,t}^1 \mathrm{d}\boldsymbol{X}_{p,t}^2 \neq \int p\left(X_t^1, \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) \log \frac{p\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right)}{\int p\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) p_{ref}\left(\boldsymbol{X}_{p,t}^2\right) \mathrm{d}\boldsymbol{X}_{p,t}^2}$$
$$\mathrm{d}X_t^1 \mathrm{d}\boldsymbol{X}_{p,t}^1 \mathrm{d}\boldsymbol{X}_{p,t}^2 = \mathrm{rDCS}\left(X_t^2 \to X_t^1\right),$$

because $\tilde{p}_{ref}$ is not translated by the deterministic perturbation in the way $\tilde{p}\left(\boldsymbol{X}_{p,t}^2\right)$ is (as the perturbation is happening after the

reference time), such that the denominators do not allow equating the integrated terms by change of variables in the generic case. Therefore rDCS is capable of uncovering transient causal influences between stimulus-triggered events exhibiting a deterministic waveform.

2.5 Alignment for spontaneous events

The relevance of peri-event time-varying causal analysis using the proposed rDCS, as well as TE and DCS, depends on the modeling assumptions of peri-event data. In particular, we assume that the neural events we want to study reflect a sequence of continuously changing hidden states and that values at each peri-event time point t' are sampled *i.i.d.* across trials from the same ground truth distribution (Shao et al., 2022) at t'. This is easily justified for stimulus-evoked events, as addressed in Section 2.3.2 and Section 2.4.2, with an intrinsic reference time for occurrence (i.e., the triggering time). However, analyzing spontaneous events whose occurrence times are not known *a priori*, such as transient events observed during sleep, requires 1) a *selection* procedure to identify them and 2) a procedure to choose a reference time point for each detected event, which is used to *align* all of them on a common *peri-event* time grid. The idea of reference points for alignment is similar to the anchor points in *Phase Rectified Signal Averaging* (Bauer et al., 2006). In contrast with such work, we focus on a transient phenomenon at the time scale of a peri-event time window instead of a very fast increase in the signal amplitude. Given a signal exhibiting spontaneous events, common procedures involve 1) *selecting* the events by thresholding a filtered version of this signal (that amplifies the events' features of interest); 2) aligning events according to the local peak of this same filtered signal to best reflect the evolution of the underlying state. The result may only approximately recover the ground truth distribution of the events, as it is influenced by the choice of filtered signal and putative signal perturbations.

Importantly, selection may lead to a biased estimation of event statistics and peri-event dynamics, due to *selecting* data based on a specific detection signal, resulting in a misleading characterization of causal interactions (e.g., wrong causal directions as seen in Supplementary Figure S6B). We will thus study how event *selection* affects the estimation of causal influence and propose an appropriate procedure on this basis. To model the effect of selection, we use an SCM-based perspective on selection bias (Bareinboim and Pearl, 2012; Bareinboim et al., 2014). We can modify the SCM in Figure 2A to incorporate an additional node S representing the selection variable, which is a binary variable indicating whether the time window, with specified reference time point, is selected (Supplementary Section SC for background). Typically S is defined by testing whether a continuous random variable D goes over a predefined threshold. D is itself a function of the time series nodes within the peri-event time window, corresponding, for example, to the aforementioned filtering operation. A practical example is the detection of oscillatory events using a band-pass filter, where the dependency of D (and thus S) on other nodes reflects the dependency of the filtered signals on past samples of $\boldsymbol{X}$ through the coefficients of a causal Finite Impulse Response (FIR) filter.

In practice, we can *a priori* choose S to depend either on the cause variables $\boldsymbol{X}^2$ (Figure 3A) or on the effect variables $\boldsymbol{X}^1$ (Figure 3B). Assuming that the filter (i.e., for constructing the continuous RV) is well chosen, and the selection threshold is high enough, choosing windows satisfying $S = 1$ will typically "over-select", i.e., exclude some peri-event time series that would actually be relevant for our analysis. Figure 3C(left, top right) illustrates how thresholding selects only a subset of peri-event trajectory samples at $t' = 0$ in a simulated scenario. This over-selection can then be modeled as sampling peri-event data from a conditional peri-event distribution $p(X|S)$, while we are interested in analyzing a ground truth distribution $p(X)$. This conditioning may induce a so-called *selection bias* in the estimation of quantities we are interested in, notably the conditional distributions that enter the calculations of TE, DCS and rDCS. The impact of such bias on those quantities as been investigated in Bareinboim and Pearl (2012); Bareinboim et al. (2014) within the SCM framework, as we describe in the following.

For simplicity and consistency with the Results section, we will restrict ourselves to models with a unidirectional causal effect (either $X^1 \to X^2$ or $X^2 \to X^1$) and assume that S is only dependent on a finite number of past peri-event times ($t' \leq 0$) as in the case of a causal FIR filter (for other cases, refer to Supplementary Section SC.2). Figures 3A, B illustrate in this setting that the conditional associated to causal arrow ($X^2 \to X^1$) can be recovered at any peri-event time only when the selection node depends on the cause variable (Supplementary Section SC.2 for justification). Specifically, this means that $P(X_t^1 \mid \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2, S) = P(X_t^1 \mid \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2)$ for the SCM in Figure 3A. For the opposite direction, $P(X_t^2 \mid \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2, S) \neq P(X_t^2 \mid \boldsymbol{X}_{p,t}^2)$ for negative peri-event time $t' \leq 0$. For the case where S depends on the effect variable, $P(X_t^1 \mid \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2, S) \neq P(X_t^1 \mid \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2)$ for negative peri-event time $t' \leq 0$ and $P(X_t^2 \mid \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2, S) \neq P(X_t^2 \mid \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2)$ for $t' < 0$ (see also Supplementary Section SC.2 and Supplementary Figure S4). The S-dependent and S-independent conditionals are visualized in Figure 3E for an example SVAR(1) model, as described in Section 2.3.2, where the innovations η_t^1 and η_t^2 are drawn from a uniform-distribution. Similarly, the conditional model of the post-intervention scenario for rDCS with selection node depending on the cause satisfies

$$
\begin{aligned}
&p^{do\left(X_t^1 := f\left(\boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t_{ref}}^2, \eta_t^1\right)\right)}\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2, S\right) \\
&= \int p_{ref}\left(\boldsymbol{X}_{p,t}^2\right) p\left(X_t^1 \mid \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2, S\right) \mathrm{d}\boldsymbol{X}_{p,t}^2 \\
&= \int p_{ref}\left(\boldsymbol{X}_{p,t}^2\right) p\left(X_t^1 \mid \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right) \mathrm{d}\boldsymbol{X}_{p,t}^2 \\
&= p^{do\left(X_t^1 := f\left(\boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t_{ref}}^2, \eta_t^1\right)\right)}\left(X_t^1 | \boldsymbol{X}_{p,t}^1, \boldsymbol{X}_{p,t}^2\right)
\end{aligned}
$$

Therefore, the KL divergence for the ground truth direction $X^2 \to X^1$ can be estimated correctly when selecting the event based on the ground-truth cause variable "$S(X^2)$", while this does not hold for the opposite direction $X^2 \to X^1$ nor when selecting based on the ground-truth effect "$S(X^1)$". As the true causal direction is unknown, we thus propose that, to investigate the dominant causal direction

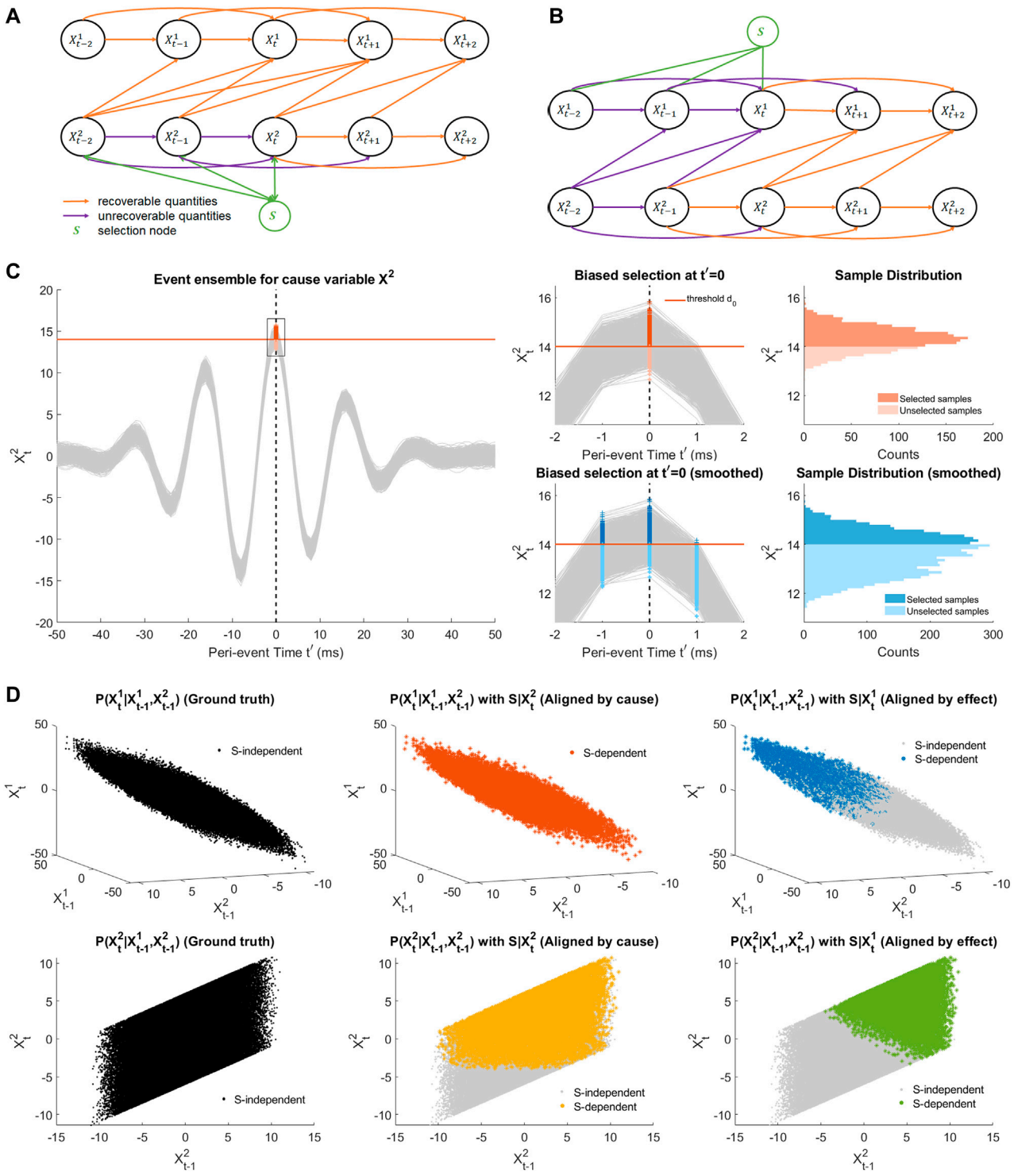


FIGURE 3
Illustration of selection bias due to thresholding and alignment. **(A)** SCM of a bi-variate SVAR(2) model with uni-directional coupling from X^2 to X^1 and a selection node S depending on states of the cause variable before peri-event time ($t' <= 0$). The selection node S represents partial selection of samples due to thresholding of the filtered cause signal (as the detection signal). Orange arrows makes the recoverable arrows with the current selection node, while purple arrows indicates the unrecoverable ones. **(B)** The same SCM as in **(A)** with the selection node depending in a similar way on the effect signal. **(C)** An example event ensemble for the cause variable X_t^2 in **(A,B)** and the detection threshold. **(D)** Zoomed event ensenbles for **(C)** (left) and histograms for selected samples compared to the full sample (right). Top panel illustrates selection bias at ground truth peri-event time $t' = 0$. The orange and shaded distributions represents histograms at a single time $t' = 0$. Bottom panel shows selection bias at the peri-event time $t' = 0$ for detected events aligned by the peak. The dataset aligned in this way reflects at $t' = 0$ is a local average of the state trajectories in a neighborhood of the target ground truth state. **(E)** Illustration of recoverability when aligning by the cause. Subplots show joint distributions of the lagged variables and the putative effect variable of a SVAR(1) model with uniformly distributed innovations, with left column for the ground-truth alignment, middle column for aligning by the cause and right column for aligning by the effect. The conditional is only recoverable for the top middle panel. See Supplementary Figure S4 for the cases of peri-event time $t'=0$ and $t'>0$.

between two event ensembles, we should focus on comparing the causality measures (TE, DCS and rDCS) for each direction when the events are aligned on the putative cause, i.e., $X^2 \rightarrow X^1|S(X^2)$ compared to $X^1 \rightarrow X^2|S(X^1)$. Although rDCS is expected to be biased for the second case when aligned by the effect variable, for uni-directionally coupled systems, as the ground truth rDCS is zero, we expect the bias to still lead to a comparatively small rDCS value relative to ground truth, such that the contrast between two directions is preserved.

Other factors may affect the estimation of causal strength. Since rDCS is defined as the expectation over the KL divergence over the past states $\boldsymbol{X}^1_{p,t}$ and $\boldsymbol{X}^2_{p,t}$ (Eq. 11), reliable estimation of $\mathrm{rDCS}(X^2_t \rightarrow X^1_t)$ also depends on the unbiased sampling of the joint probability of $\boldsymbol{X}^1_{p,t}$ and $\boldsymbol{X}^2_{p,t}$. We argue here that this is approximately satisfied as the conditioning is made on a specific detection signal rather than on these variables, such that they are mildly affected by it.

Next, the above mentioned alignment procedure may affect causal strength estimation. *Perfect alignment* (considered *ground truth*) refers to the condition where the ground-truth hidden states are identical for all trials at each peri-event time t' in an extracted event ensemble, as shown in Figure 3D for $t' = 0$. In this scenario, no further alignment is needed as all trials are intrinstically aligned. In order to study the influence of the aforementioned selection bias specifically due to thresholding, we may still apply selection to the perfectly aligned dataset, resulting in excluding below threshold samples from the estimation procedure. We refer to this situation as *single-time selection* of events where trials are aligned based on known ground-truth reference time. This setting assumes that one knows the hidden state, which is possible only for stimulus triggered or simulated events, but impossible for experimentally observed spontaneous events. In the latter case, by thresholding over the whole observed signal one typically end up selecting successive sliding time windows that all have a detection signal exceeding the threshold (e.g., Figure 3D(Bottom right)). Selecting all these points can be interpreted as smoothing the ground truth state over all these neighboring state space points, an alignment scenario which we name as *smoothed alignment* of events. In practice, a common alternative is to further select among above-threshold overlapping peri-event time-windows the local peaks as the reference points, which can be understood as a non-uniform subsampling of the smoothed alignment and can be unified into the same scenario category.

2.6 Data processing pipeline

The whole analysis procedure can be conducted in two phases: event selection and causal analysis. We will elaborate on the detailed steps in each phase in the following.

- Phase 1: Event Selection
  1. *Filtering*: given a bi-variate signal (as a simple case), for different purposes of study, one would need to find an appropriate filter to apply to the original signals such that certain features of the underlying system can be amplified. For example, to locate the Sharp Wave-Ripples (introduced in the Introduction and analyzed in Section 3.3) that are prominent in the ripple band [80–250]Hz, one would use a bandpass filter such that the irrelevant components are attenuated. Events can be also detected with a template matching procedure, which is another type of filtering (Supplementary Figure S2).
  2. *Thresholding*: a certain threshold is determined beforehand (up to the specific feature of the question) and applied to the filtered signal. As the filtered signal is designed to amplify the feature, time points where the filtered signals are over the threshold are candidate reference points. Reference points define the peri-event time t' and are used to extract peri-event data as multiple trials.
  3. *Alignment*: the thresholding procedure can be applied to either the cause or effect signals. One can select all candidate reference points obtained by filtering either signal (for the smoothed alignment case) or the time points of local peaks (of the filtered signal) as reference points. Then the bi-variate peri-event trials are extracted in a fixed-length window surrounding the reference points, thus forming the peri-event ensemble for further analysis.
- Phase 2: causal analysis
  1. *Model order selection*: as mentioned in Section 2.2.1, our estimation of information theoretic quantities is based on time-inhomogeneous SVAR models. One thus needs to determine the optimal SVAR model order that best reflects the underlying dynamics. A common approach for model order selection is the Bayesian Information Criterion (BIC), which we have extended to the time-varying case in Shao et al. (2022) using the extracted event ensembles obtained in the first phase.
  2. *SVAR model estimation*: Shao et al. (2022) also provide a way to estimate the SVAR model with the extracted event ensemble and the optimized model order. Thus we will obtain an estimate of the autoregressive parameters, i.e., the autoregressive coefficients and innovation mean and variance.
  3. *Computation of causality measures*: with the estimated autoregressive parameters and the signals second order statistics, we can estimate the time-varying causality measures as detailed in Supplementary Section SD: TE based on Supplementary Equation S8, DCS on Supplementary Equation S9 and rDCS on Supplementary Equation S10.

Notably, the causal analysis procedure can be applied to event ensembles obtained with any type of alignment. However, as elaborated in Section 2.5, we propose to compare the causality measures in two different directions from the event ensembles where trials are aligned by the putative causes. To facilitate the application of this analysis framework, we have made available the code that performs the aforementioned experimental procedure (see https://github.com/KaidiShao/event_causality_frontiers).

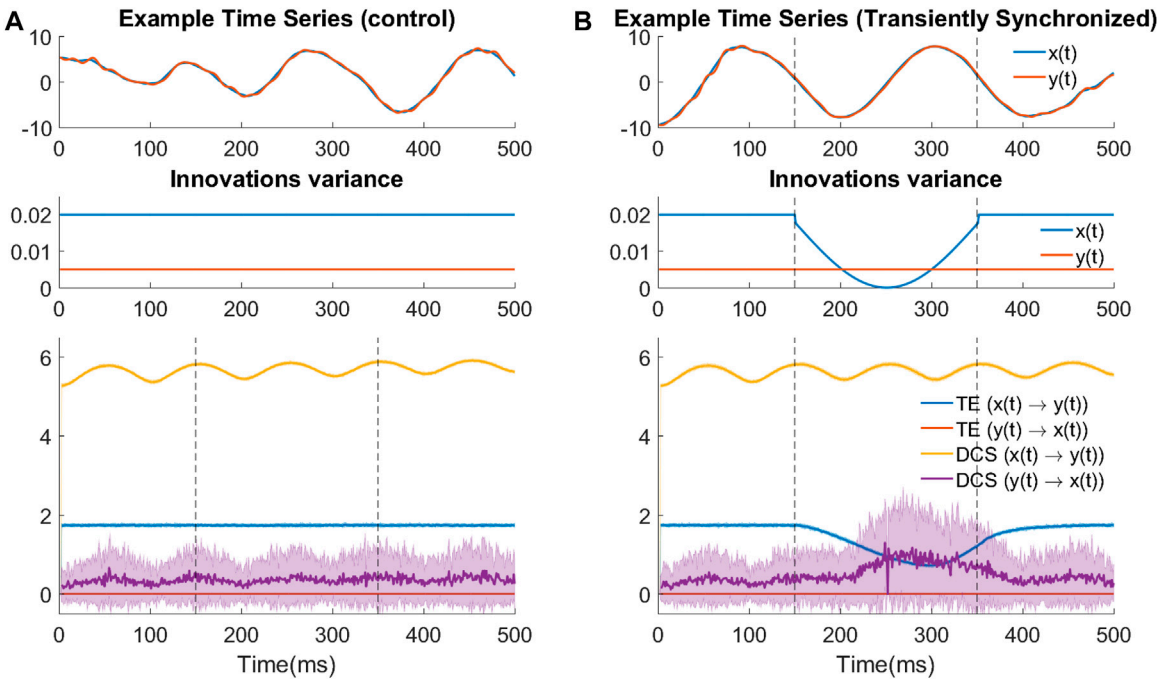


FIGURE 4
TE fails when the signals are strongly synchronized. **(A)** Control experiments where synchrony is not changed. (top) Example trace of the bivariate signal in the control experiment. (middle) Time-varying design of innovation's variance for both variables in the control experiment. (bottom) Time-varying TE and DCS results in the control experiment. **(B)** TE underperforms during transient increased synchrony induced by a tiny change in noise variance. The transient change can be seen as an event. Subfigure designs are the same as **(A)**.

3 Results

In this section, we first focus on illustrating the properties of TE, DCS and rDCS with simulated toy models. The problem of vanishing TE occurring with synchronized signals and the benefits of DCS in the same situation will be investigated in Section 3.1. Next, we simulate a simple uni-directionally coupled SVAR(4) system with rhythmic perturbations of the cause variable to generate transient events, where we will show that rDCS is able to reflect the change of causal effects due to this perturbation while TE and DCS fail. We also study the influence of the alignment method in the same example, as well as in experimental *in vivo* recordings from uni-directionally coupled hippocampal regions during SWRs.

3.1 The case of strongly-correlated signals

As mentioned in Section 2.3.1, TE does not capture well causal influences when the cause and effect signals are strongly correlated with each other, contray to DCS. Here, to illustrate such contrast, we simulate a bivariate dynamical system in the form of two synchronized continuous harmonic oscillators $x(t)$ and $y(t)$, with uni-directional coupling (i.e., $x(t)$ driving $y(t)$):

$$\begin{cases} \dfrac{d^2x}{dt^2} = -2\zeta_x\omega_x\dfrac{dx}{dt} - \omega_x^2 x + n_x, \\ \dfrac{d^2y}{dt^2} = -2\zeta_y\omega_y\dfrac{dy}{dt} - \omega_y^2 y + cx + n_y. \end{cases} \tag{13}$$

In this system, $x(t)$ is designed as an under-damped oscillator ($\zeta_x = 0.015722$), which approximately oscillates at a period $T_x = 200$ samples corresponding to natural (angular) frequency $\omega_x = 2\pi/T_x = 0.0314$ rad/sample. To achieve synchrony, $y(t)$ is also designed as an under-damped oscillator ($\zeta_y = 0.2$) whose intrinsic oscillation gradually vanishes and finally follows the oscillation of $x(t)$ with a coupling strength of $c = 0.098$. For $y(t)$, $T_y = 20$, $\omega_y = 2\pi/T_y = 0.314$. We also add small Gaussian innovations to both oscillators: $n_x \sim \mathcal{N}(0, 0.02)$, $n_y \sim \mathcal{N}(0, 0.005)$. Adding this noise allows fitting a SVAR model to the signals to assess the causal interactions with TE and DCS. SVAR parameter estimation would fail with deterministic signals by causing the covariance matrix estimates to be singular.

Using the Euler method with a time step of 1 and random initial points ($\mathcal{N}(0, 1)$), we simulated 2000 trials of this uni-directionally coupled system with 1000-point length. We discarded the first 500 points to ensure that the time series reach a sufficient level of synchronization. We can see this system as a stationary SVAR(2) process because numerical simulation with the Euler method generates data as a function of the last two past states. The idea of using a SVAR(2) model is elaborated on in the Supplementary Section SE. Notably, modeling simulated data with a SVAR(2) model is also possible if the numerical integration method is switched to Runge-Kutta, despite the SVAR(2) parameters having a more complex form than the continuous formulation of the system.

Figure 4A shows the results of time-varying TE and DCS for assessing the causal effects between $x(t)$ and $y(t)$. Calculation is performed in both the ground truth direction ($x(t) \rightarrow y(t)$) and the

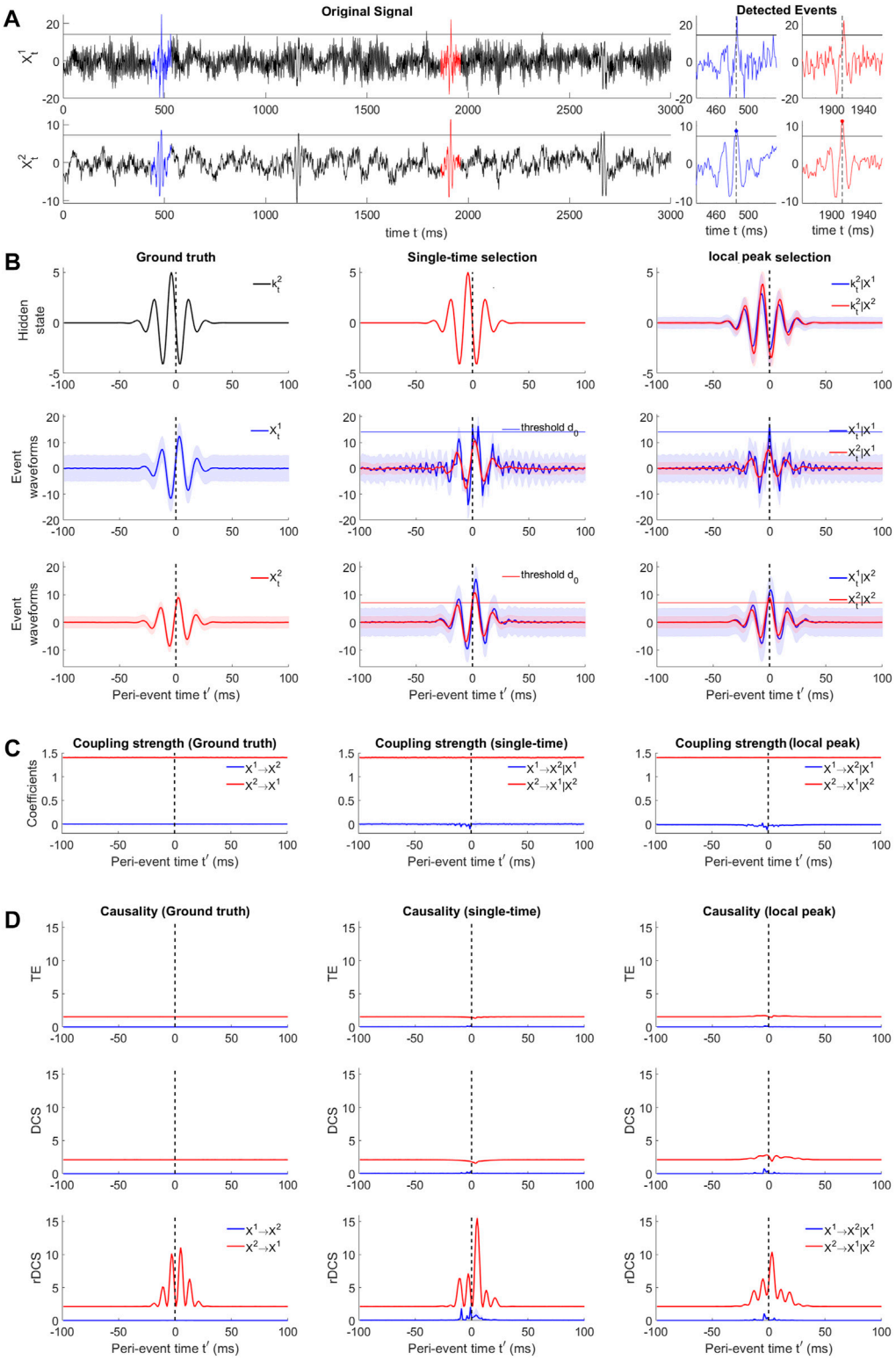


FIGURE 5
Causal analysis for simulated perturbation events with non-zero innovations. **(A)** Example signal traces of the bi-variate SVAR(4) system (black). Blue and red traces mark two example events detected by thresholding over the cause X_t^2. Blue and red dots show other reference points. **(B)** (Top) Hidden states for *ground-truth alignment* (left), *single time selection* of the ground truth event ensembles due to thresholding (middle) and events aligned by local peaks over threshold (right). (Middle) ground truth event ensemble for X_t^1 (left) and bi-variate ensembles of the other two selections aligned by X_t^1 (middle, right). Thin blue line represents the threshold in X_t^1. (Bottom) Same settings as in (middle) but aligned by X_t^2. **(C)** Example elements of coupling strength in the ground truth directions $X_t^2 \to X_t^1$ (red) and the opposite direction $X_t^1 \to X_t^2$ (blue) for 3 types of event ensembles aligned by putative cause. **(D)** TE (left), DCS (middle) and rDCS (right) for all 3 types of event ensembles aligned by putative cause.

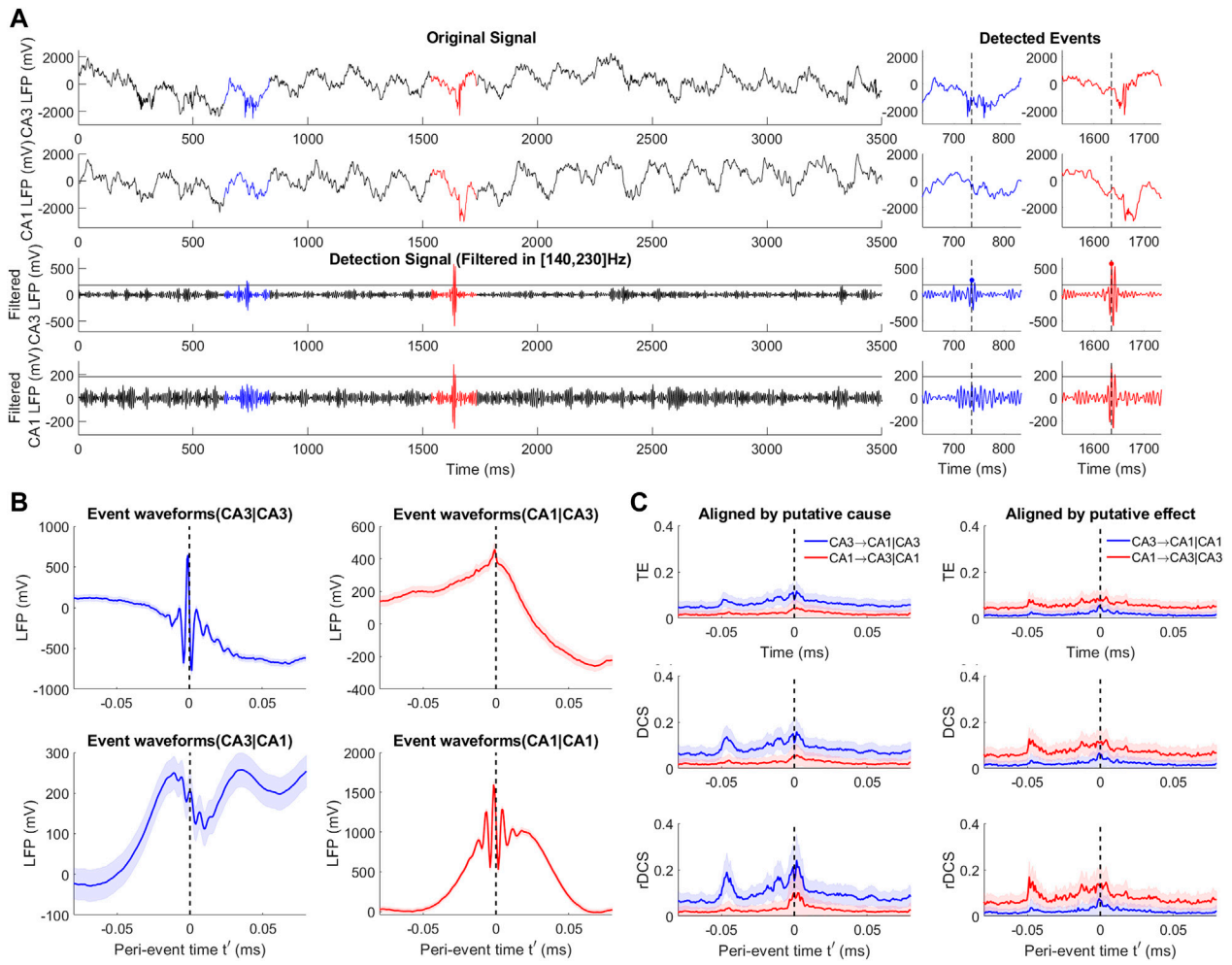


FIGURE 6
Event-based causal analysis for SWRs in rodent hippocampal CA3 and CA1 regions. **(A)** Examples signal traces of the original signals and bandpass filtered signals of CA3 and CA1 regions (black). Blue and red traces mark two example events detected by thresholding over the cause CA3 and aligned by the local peak. Blue and red dots show other reference points. **(B)** Event waveforms of SWR event ensembles at CA3 (left) and CA1 (right) regions aligned by CA3 (Top) and CA1 (Bottom) signals. Shades repensent the ensemble standard error averaged over 1024 channel pairs. **(C)** Peri-event causality measured by TE (Top), DCS (middle) and rDCS (Bottom) for event ensembles aligned by the putative cause (left) and the putative effect (right). Shades reflect standard deviation of 100 repeated bootstrapped ensembles.

opposite direction. We first look at the control experiment. Consistent with the system's stationarity, TE is constant in both directions while being higher in the ground-truth direction. DCS in the ground-truth direction stays at a relatively high level, despite some small oscillation under a frequency similar to the intrinsic oscillation frequency of $x(t)$.

With respect to the detection of causal direction, both measures are able to detect the correct direction (i.e., causation for $x(t) \to y(t)$ is much larger than in the opposite direction). It is also reasonable that DCS in both directions is higher than TE, according to its definition in section 2.2.3. However, from the control experiment, we cannot conclude that the smaller TE values are due to its definition or due to the strong synchrony in the signals.

Therefore, we introduced a transient decrease of the noise variance in the cause signals ($x(t)$). The logic of designing this transient change is the following: the level of synchronization will increase with weaker noise, but the system and input magnitude remain the same because the contribution of the noise change to the signal amplitude is negligible; thus if TE is insensitive to the level of synchronization of signals, its values are expected to stay constant. However, as the results show in Figure 4B, there is a transient decrease of TE during the interval where noise variance is decreased, suggesting that TE performs poorly in the cases where the cause and effect signals are strongly synchronized. As such strong synchronized oscillations are common phenomena in the context of transient neural events, one would need to pay extra attention when using TE (as a widely-applied causality measure) to investigate the direction of causation during these transient phenomena.

3.2 The case of deterministic perturbations

In this section, we directly address the benefits of rDCS over TE and DCS when applied to signals driven by deterministic perturbations. To illustrate this specific property, we designed some simple transient events perturbing the innovation parameters of a stationary SVAR process with uni-directional coupling. The events are generated by

feeding the cause signal with innovations with non-zero time-varying means, such that both signals will exhibit temporal oscillations. We refer to these events as *perturbation events* in the following sections. These perturbations intrinsically define a hidden state that parametrizes the ground truth distribution of peri-event data. We exploit the hidden state and demonstrate that the proposed alignment method in Section 2.5 is efficient for recovering the time-varying causal direction between the two variables.

3.2.1 Simulation procedure

We simulated a non-stationary uni-directionally-coupled autoregressive system defined in Eq. 2a and Supplementary Eq. 2b. The causal direction is $X^2 \rightarrow X^1$. The system is designed as a bivariate SVAR(4) process with a time-invariant coefficient matrix: $\mathbf{a}^\top = [-0.55, -0.45, -0.55, -0.85]$, $\mathbf{b}^\top = [1.4, -0.3, 1.5, 1.7]$, $\mathbf{c}^\top = [0, 0, 0, 0]$ and $\mathbf{d}^\top = [0.9, -0.25, 0, 0.25]$. These coefficients were randomly generated and kept after checking the stability of the SVAR(4) system. Uni-directional interactions are ensured by setting the autoregressive coefficients associated to interactions in the opposite direction (i.e., $\mathbf{c}$) to zero for all lags.

We enforce non-stationarity of η_t^2, the innovations of the ground truth cause process $\{X_t^2\}$. Both innovations η_t^1 and η_t^2 are drawn from a Gaussian distribution with unit variance (with no correlation in between, i.e., $\mathrm{Cov}[\eta_t^1, \eta_t^2] = 0$); the difference is that $\mathbb{E}[\eta_t^1] = k_t^1 = 0$ while $\mathbb{E}[\eta_t^2] = k_t^2$ is non-zero and time-varying. We designed the time-varying profile of k_t^2 as a Morlet-shaped waveform to mimic the oscillatory properties of neural event signals: $k_t^2 = H \exp(-(\alpha x)^2/2) \cos(5\alpha x)$, where $\alpha = 2/25$ is a constant controlling the event duration, and $H = 4$ is the amplitude of the highest peak in the center of the event. The total duration of the Morlet-shaped waveform is 101 ms. The innovation's mean designed for X^2 is shown in Figure 5B (top left panel).

We generated this bi-variate SVAR(4) process for 1300*s* consisting of 5,000 trials of perturbation events by transiently varying η_t^2, detecting event occurrence based on the cause X_t^2, as illustrated in Figure 5A. The central peaks of these Morlet events are used as the ground-truth reference points for which peri-event time $t' = 0$, and used to extract a dataset of multi-trial events ensemble with a 200-*ms* peri-event window such that t' ranges, from −99 *ms* to +100 *ms* (i.e., there is no alignment procedure that could lead to selection bias, see Section 2.5). The event waveforms of the cause variable X_t^2 and the effect variable X_t^1 are illustrated in Figure 5B (bottom left, middle left). The whole process is repeated 100 times to obtain variabilities plotted in the figure.

3.2.2 Effect of trial selection and alignment on model estimation and causality measures

The designed deterministic innovation (i.e., identical across trials), can be seen as imposing a hidden state evolving across the peri-event interval. The event ensembles obtained by this ground truth model define a dataset where no event selection and alignment is needed. We can compare the SVAR model estimation and causality measures resulting from this dataset to the outcomes obtained by selecting and aligning events based on either variable X^2 or X^1, as discussed in Section 2.5.

To validate the recoverability theory in the presence of selection bias due to the event detection procedure, we test the *single-time selection* setting (see Section 2.5) where sub-threshold trials are removed from the ground truth peri-event dataset (as illustrated in Figure 3D(Top right)), thus preserving the ground-truth hidden states (Figure 5B(Top middle)). The peri-event trials having reference point values higher than a threshold $d_0 = 3SD$ for the chosen variable are selected, where the standard deviation is computed from the whole signal. The selected event ensembles are shown in Figure 5A(middle center) for thresholding based on X_t^1 and Figure 5A(middle right) for thresholding based on X_t^2. Notably, this kind of selection is only feasible when the hidden state in known, which is not realistic practically for real data.

Next we demonstrate the appropriateness of the approach performed on real data (i.e., selection and smoothed alignment based on putative cause), we set d_0 as a threshold and performed smoothed alignment over the original signal itself. We obtain an event ensemble by selecting local peaks for points over d_0 as new reference points, which is shown in Figure 5B(middle right and bottom right). This can be seen as a smoothed version of the ground-truth dynamics, which is also confirmed by checking the aligned hidden states (Figure 5B(Top right)).

While inferring SVAR model parameters of the event ensembles according to Shao et al. (2022), the true model order 4) can be recovered for all five ensembles. Figure 5C demonstrates the recoverability of conditional probabilities for ensembles aligned by the putative cause. Coupling strengths from the putative cause to the putative effect are plotted in red. As described in the simulation procedure in Section 3.2.1, the coupling strength is constant over time, which is reflected in Figure 5C(left). Consistent with the theory in Section 2.5, biased selection of event trials on the samples at $t' = 0$ leads to unbiased estimation of the coupling strength $X_t^2 \rightarrow X_t^1$ aligned by the cause X^2 (denoted also as "$|X^2$" in Figure 5C(middle)). By comparison, the coupling strength in the other direction is slightly biased at negative peri-event times (t') but still relatively close to its true value 0). This contrast holds for alignment with local peaks over threshold, as seen in Figure 5C(right).

Figure 5D (Top, middle, bottom) shows the corresponding results of how causality measures perform in the three alignment scenarios. For clearer visualization of TE and DCS with a zoomed vertical scale, see Supplementary Figure S8. During the periods where no transient events occur, all three measures are able to infer a time-invariant causal effect in the ground-truth direction ($X^2 \rightarrow X^1$) compared to the opposite direction. Besides, in line with theoretical predictions, DCS is higher than TE and is equal to rDCS. During the perturbation events, in the ground truth direction TE and DCS remain constant and rDCS exhibit a rhythmic pattern. These results match the theoretical predictions: TE and DCS measures the connectivity strength, which does not change, while rDCS measures the combined causal effect related to the connectivity and the event-based changes at the cause, yielding larger variations transmitted to the effect node.

In the transient time scale, thresholding leads to selection bias in estimating causality measures. In the case where event ensembles are aligned by *single-time selection* of the cause X_t^2, TE and DCS of the ground truth direction is underestimated while rDCS is slightly overestimated around $t' = 0$. A bias appears in the opposite direction while aligned by the effect, but the direction of causation is detected correctly. The case of local peak alignment shows similar results, except the peak amplitude of the smoothed rDCS is less amplified. Notably, in the smoothed case, a transient increase is observed in both TE and DCS, resembling the envelope of the perturbation. This is likely an effect of the smoothing procedure but is quickly interrupted by a negative bias due to thresholding,

making the results unreliable in detecting transient changes. We also showed a negative example with putative effect alignment in Supplementary Figure S5, where the coupling strength in the causal direction is much weaker than the model in Figure 5. The coupling strength of the causal direction undergoes a sharp decrease at peri-event time $t' = 0$, leading to a transient underestimation of TE, DCS and rDCS for both single-time and local peak scenarios. The close-to-zero value of rDCS is misleading for the inference of transient causal interactions, thus illustrating the unreliability of putative effect alignment.

Thus, this simulation experiment of perturbation events demonstrates the effectiveness of rDCS in reflecting the causal influence when the cause is perturbed by a deterministic exogenous input compared to TE and DCS, validating that rDCS is a better measure to address event-based causal interactions. More importantly, we highlight here the trial alignment problem when dealing with event-based data, especially when events occur spontaneously. Supplementary Figure S7 is a clear example showing the impact of alignment on information-theoretic measures: aligning on the actual effect could reverse the detected direction of causation. Thus, by contrasting the different impacts of alignment on information-theoretic measures, we show that in practice, selection via thresholding and aligning the event ensemble with the local peaks of the putative cause is a good way to assess the ground truth event-based causality given uni-directional connections. This approach will be further applied to real data in the next section.

3.3 Validation on SWRs-based causality between CA3 and CA1 regions

Sharp Wave-Ripple (SWR) events, hypothesized as a key element in implementing memory consolidation in the brain, have been reported in the electrophysiological recordings within the hippocampus of both macaques and rodents. In this section we detect SWRs in an experimental dataset to investigate the behavior of TE, DCS and rDCS in a neuroscientific context where the event-hosting brain regions are uni-directionally coupled, i.e., in a situation where the causal direction is known *a priori*.

SWRs are primarily generated in the CA1 area of the hippocampus. The somas of CA1 pyramidal cells are located in the pyramidal layer ('pl') while their dendritic trees are rooted in the stratum radiatum ('sr'). It is hypothesized that the dendritic trees receive strong excitatory inputs from the pyramidal cells in CA3 which generate post-synaptic activities in the dendritic trees. This results in LFP activities at low frequencies (0–30Hz, due to the sharp-wave) and in the gamma band (30–80Hz, due to CA3 oscillations). Then the dendritic activities propagate to the soma, where recurrent interactions between inhibitory and excitatory cells generate a fast oscillation, the ripples (80–250 Hz).

We applied the event-based causality analysis to an open source dataset where electrophysiological recordings in the CA3 and CA1 regions of rodent hippocampus have been performed with 4 shanks of 8 channels simultaneously in each region (Mizuseki et al., 2014). In agreement with the SWR generation mechanism explained in the above paragraph, anatomical studies (Csicsvari et al., 2000) support uni-directional anatomical coupling between these two regions within the hippocampal formation, i.e., the ground truth direction is known to be CA3 → CA1. The analysis is based on two Local Field Potential (LFP) data sessions recorded from the rat named 'vvp01' with a sampling rate of 1252 Hz. An example trace of a channel pair of both CA3 and CA1 regions is shown in Figure 6A. As SWRs are more challenging to observe during behavioral sessions, we perform our analysis only on a session of sleep which lasts 4943.588*s*.

Following Mizuseki et al. (2009), we detect SWRs by applying an 49-ordered FIR filter in the frequency band [140, 230]Hz to each channel of signals in both regions. The detailed detection procedure has been elaborated in Section 2.6 for the reference of readers and is similar to what is performed in Section 3.2. We set a threshold over the mean of the filtered signals (5 SD) to locate the events and align them according to the local peak time points over threshold.

Figure 6A(Bottom) shows the case aligned on the CA3 signals. The peri-event window for display has been chosen to be [-79.9, 79.9]*ms*, while VAR model estimation and the BIC-based model order selection are performed according to Shao et al. (2022). For each channel pair, we obtain two bi-variate event ensembles, corresponding to the two alignment conditions; thus, in total, we extract 2*1024 event ensembles (1024 channel pairs and 2 alignment conditions). The event waveforms and statistics of an example channel pair for different alignments are illustrated in Figure 6B.

SWR-based causality measures shown in Figure 6C compare the alignment by the putative cause and by the putative effect. The reference states used for estimating rDCS are the averaged states over the first 16 *ms* time points in the window. The standard deviation plotted in the figure originates from 100 times bootstrapped ensembles and the variability is averaged over 1024 channel pairs. In line with the theoretical predictions, the ground truth direction (CA3 → CA1) is well recovered when using an alignment by the putative cause, but not when aligning by the putative effect. TE, DCS and rDCS in the opposite of the truth direction are not significantly different from zero, which is consistent with the uni-directionality of anatomical connections posited by anatomical studies. Significantly stronger causal influences in the ground truth direction are shown by TE, DCS and rDCS before the alignment point ($t' = 0$), matching the hypothesized SWR generating mechanism that the CA3 region drives the SWR interactions in CA1 region. The lack of difference between the two directions in more stationary states might be explained by the ineffectivity of causal measures based on linear VAR models to capture non-linearity (Shajarisales et al., 2015). The transient increase in the non-ground truth direction when using alignment on the putative cause might be explained by the selection bias elaborated on in Section 2.5.

4 Discussion

In summary, we have discussed the benefits and shortcomings of two time-varying causality measures (TE and DCS) in characterizing causal interactions based on peri-event data. To address their insensitivity to deterministic perturbations, we proposed a novel measure, rDCS, justified within the SCM framework. We compared the performance of these causality measures on perturbation events with innovations having time-varying means and electrophysiological recordings of hippocampal SWRs. The benefits of rDCS are supported by the perturbation events presented in Section 3.2. As causality analysis of transient events

aims at uncovering the network mechanisms underlying these phenomena (e.g., addressing whether one event *drives* the other), we argue for the use of rDCS as it provably captures causal influences due to event-related changes in the cause that propagate to target regions through anatomical connections, even if these changes have little variability across trials. The outcome of rDCS is further illustrated on *in vivo* recordings of SWRs events in two hippocampal subfields.

Transient events are nonstationary signals that likely occur when the brain undergoes a transition from one state to another. Studying the "local" properties of the underlying non-equilibrium dynamics in regions of the state space might provide insights into the mechanism driving this transition. Earlier methods investigating such local dynamic properties include the local Lyapunov exponent (Pikovsky, 1993), while other common methods characterize local interactions between state variables within a short sliding time window, e.g., the local cross correlation (Buchner et al., 2009) or piecewise Granger causality (Ding et al., 2006). Our approach, although focused on the meaningful quantification of causal strength, is in line with the latter idea, where the time-varying SVAR model finds a 1-step local linear mapping in the trajectory formed by event trajectories, thus enabling to reveal transient causal interactions at a fast time scale, which may differ from the results obtained at equilibrium. As the measures are based on SVAR models, they can also be easily extended to a spectral form in order to capture the rich spectral properties in transient dynamics.

Contrasting the three measures of causal strength, TE is designed to assess conditional dependencies in observational data, while DCS and rDCS exploit this information to infer the impact of performing interventions of the SCM. In theory and as shown in the experiment of Section 3.1, TE can lead to counterintuitive outcomes applied to strongly synchronized events (a widely observed nonlinear phenomenon). While support has been provided for DCS and rDCS to be more appropriate measures of causal strength, they still require, like TE, certain assumption to be met (see also Section 2.1). A major concern is unobserved confounding, which might bias the estimated causal directions (e.g., the Simpson's paradox in Pearl (2000)). Confounding effects can be corrected for by including activities from other regions, and there are also a few theoretical approaches to account for unobserved confounding under strong assumptions (Geiger et al., 2015; Mastakouri et al., 2021).

Selection bias is a fundamental issue for analyzing spontaneous neural activities, especially in case of any unsupervised detection or analysis. In this study we have demonstrated its impact on the alignment of the detected transient events and the resulting bias in causal inference. However, our proposal of *putative cause alignment* to estimate causal effect is theoretically supported only in the case of uni-directional coupling. Future work should assess the effect of selection bias in the case of bidirectional interactions and establish a framework to correct for such bias, not only in the context of causal strength inference but more generally for recovering the underlying event dynamics.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://github.com/KaidiShao/event_causality_frontiers, https://crcns.org/data-sets/hc/hc-3/about-hc-3.

Author contributions

Conceptualization, MB and NL; Methodology, MB and KS; Software, KS; Validation, KS; Formal Analysis, KS and MB; Resources, NL; Writing—Original Draft Preparation, KS and MB; Writing—Review and Editing, KS, and MB; Supervision, MB. All authors contributed to the article and approved the submitted version.

Funding

All authors were supported by the Max Planck Society. NL and KS acknowledge the support from the Shanghai Municipal Science and Technology Major Project (Grant No. 2019SHZDZX02). MB was also supported by the German Federal Ministry of Education and Research (BMBF): Tübingen AI Center, FKZ: 01IS18039B.

Acknowledgments

MB and KS would like to thank Philipp Geiger and Mingyu Yang for useful discussions.

Conflict of interest

The author NL declared that he was an editorial board member of Frontiers, at the time of submission.

The remaining 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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnetp.2023.1085347/full#supplementary-material

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *J. Econ. Literature* 59, 391–425. doi:10.1257/jel.20191450

Ay, N., and Polani, D. (2008). Information flows in causal networks. *Adv. complex Syst.* 11, 17–41. doi:10.1142/S0219525908001465

Bareinboim, E., and Pearl, J. (2012). Controlling selection bias in causal inference. *Artif. Intell. Statistics* 22, 100–108.

Bareinboim, E., Tian, J., and Pearl, J. (2014). Recovering from selection bias in causal and statistical inference. *AAAI* 28, 2410–2416. doi:10.1609/aaai.v28i1.9074

Barnett, L., Barrett, A. B., and Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for Gaussian variables. *Phys. Rev. Lett.* 103, 238701. doi:10.1103/PhysRevLett.103.238701

Barnett, L., and Seth, A. K. (2015). Granger causality for state-space models. *Phys. Rev. E Stat. Nonlin Soft Matter Phys.* 91, 040101. doi:10.1103/PhysRevE.91.040101

Barnett, L., and Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *J. Neurosci. methods* 223, 50–68. doi:10.1016/j.jneumeth.2013.10.018

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

Bauer, A., Kantelhardt, J. W., Bunde, A., Barthel, P., Schneider, R., Malik, M., et al. (2006). Phase-rectified signal averaging detects quasi-periodicities in non-stationary data. *Phys. A Stat. Mech. its Appl.* 364, 423–434. doi:10.1016/j.physa.2005.08.080

Besserve, M., Lowe, S. C., Logothetis, N. K., Schölkopf, B., and Panzeri, S. (2015). Shifts of gamma phase across primary visual cortical sites reflect dynamic stimulus-modulated information transfer. *PLoS Biol.* 13, e1002257. doi:10.1371/journal.pbio.1002257

Besserve, M., Schölkopf, B., Logothetis, N. K., and Panzeri, S. (2010). Causal relationships between frequency bands of extracellular signals in visual cortex revealed by an information theoretic analysis. *J. Comput. Neurosci.* 29, 547–566. doi:10.1007/s10827-010-0236-5

Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., and Scott, S. L. (2015). Inferring causal impact using Bayesian structural time-series models. *Ann. Appl. Statistics* 9, 247–274. doi:10.1214/14-aoas788

Buchner, T., Zebrowski, J. J., Gielerak, G., and Grzeda, M. (2009). "Direction of information flow between heart rate, blood pressure and breathing," in *Complex dynamics in physiological systems: From heart to brain* (Berlin, Germany: Springer), 33–47.

Cekic, S., Grandjean, D., and Renaud, O. (2018). Time, frequency, and time-varying Granger-causality measures in neuroscience. *Statistics Med.* 37, 1910–1931. doi:10.1002/sim.7621

Correa, J., and Bareinboim, E. (2020). A calculus for stochastic interventions: Causal effect identification and surrogate experiments. *Proc. AAAI Conf. Artif. Intell.* 34, 10093–10100. doi:10.1609/aaai.v34i06.6567

Csicsvari, J., Hirase, H., Mamiya, A., and Buzsáki, G. (2000). Ensemble patterns of hippocampal CA3-CA1 neurons during sharp wave–associated population events. *Neuron* 28, 585–594. doi:10.1016/S0896-6273(00)00135-5

Diba, K., and Buzsaki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nat. Neurosci.* 10, 1241–1242. doi:10.1038/nn1961

Diks, C., and Wolski, M. (2016). Nonlinear Granger causality: Guidelines for multivariate analysis. *J. Appl. Econ.* 31, 1333–1351. doi:10.1002/jae.2495

Ding, M., Chen, Y., and Bressler, S. L. (2006). Granger causality: Basic theory and application to neuroscience. *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications* 2006, 437–460.

Ego-Stengel, V., and Wilson, M. A. (2010). Disruption of ripple-associated hippocampal activity during rest impairs spatial learning in the rat. *Hippocampus* 20, 1–10. doi:10.1002/hipo.20707

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* 83, 051112. doi:10.1103/PhysRevE.83.051112

Fries, P. (2015). Rhythms for cognition: Communication through coherence. *Neuron* 88, 220–235. doi:10.1016/j.neuron.2015.09.034

Geiger, P., Zhang, K., Schoelkopf, B., Gong, M., and Janzing, D. (2015). Causal inference by identification of vector autoregressive processes with hidden components. *International Conference on Machine Learning (PMLR)* (Lille, France). 1917–1925.

Geweke, J. F. (1984). Measures of conditional linear dependence and feedback between time series. *J. Am. Stat. Assoc.* 79, 907–915. doi:10.1080/01621459.1984.10477110

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424. doi:10.2307/1912791

Janzing, D., Balduzzi, D., Grosse-Wentrup, M., and Schölkopf, B. (2013). Quantifying causal influences. *Ann. Stat.* 41, 2324–2358. doi:10.1214/13-aos1145

Lee, A. K., and Wilson, M. A. (2002). Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron* 36, 1183–1194. doi:10.1016/s0896-6273(02)01096-6

Lizier, J. T. (2014). JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. *Front. Robotics AI* 1, 11. doi:10.3389/frobt.2014.00011

Marinazzo, D., Liao, W., Chen, H., and Stramaglia, S. (2011a). Nonlinear connectivity by Granger causality. *Neuroimage* 58, 330–338. doi:10.1016/j.neuroimage.2010.01.099

Marinazzo, D., Liao, W., Chen, H., and Stramaglia, S. (2011b). Nonlinear connectivity by Granger causality. *Neuroimage* 58, 330–338. doi:10.1016/j.neuroimage.2010.01.099

Marinazzo, D., Pellicoro, M., and Stramaglia, S. (2008). Kernel method for nonlinear Granger causality. *Phys. Rev. Lett.* 100, 144103. doi:10.1103/PhysRevLett.100.144103

Mastakouri, A. A., Schölkopf, B., and Janzing, D. (2021). "Necessary and sufficient conditions for causal feature selection in time series with latent common causes," in *International Conference on Machine Learning (PMLR)*. July 18 - 24, 2021 7502–7511.

McAllester, D., and Stratos, K. (2020). "Formal limitations on the measurement of mutual information," in International Conference on Artificial Intelligence and Statistics (PMLR). 875–884.

Mizuseki, K., Diba, K., Pastalkova, E., Teeters, J., Sirota, A., and Buzsáki, G. (2014). Neurosharing: Large-scale data sets (spike, LFP) recorded from the hippocampal-entorhinal system in behaving rats. *F1000Research* 3, 98. doi:10.12688/f1000research.3895.1

Mizuseki, K., Sirota, A., Pastalkova, E., and Buzsáki, G. (2009). Theta oscillations provide temporal windows for local circuit computation in the entorhinal-hippocampal loop. *Neuron* 64, 267–280. doi:10.1016/j.neuron.2009.08.037

Moneta, A., Chlass, N., Entner, D., and Hoyer, P. (2011). "Causal search in structural vector autoregressive models," in *Proceedings of the neural information processing systems mini-symposium on causality in time series*. Editors F. Popescu and I. Guyon (Vancouver, Canada: PMLR).

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

Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge: Cambridge Univ Press. 29.

Peters, J., Janzing, D., and Schölkopf, B. (2017). *Elements of causal inference – foundations and learning algorithms*. Cambridge: MIT Press.

Pikovsky, A. S. (1993). Local Lyapunov exponents for spatiotemporal chaos. *Chaos Interdiscip. J. Nonlinear Sci.* 3, 225–232. doi:10.1063/1.165987

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* 66, 688–701. doi:10.1037/h0037350

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

Schreiber, T. (2000). Measuring information transfer. *Phys. Rev. Lett.* 85, 461–464. doi:10.1103/PhysRevLett.85.461

Shajarisales, N., Janzing, D., Schölkopf, B., and Besserve, M. (2015). Telling cause from effect in deterministic linear dynamical systems. *International Conference on Machine Learning (PMLR)*, 285–294.

Shao, K., Logothetis, N. K., and Besserve, M. (2022). Bayesian information criterion for event-based multi-trial ensemble data. *arXiv preprint arXiv:2204.14096*. doi:10.48550/ARXIV.2204.14096

Shojaie, A., and Fox, E. B. (2022). Granger causality: A review and recent advances. *Annu. Rev. Statistics Its Appl.* 9, 289–319. doi:10.1146/annurev-statistics-040120-010930

Shorten, D. P., Spinney, R. E., and Lizier, J. T. (2021). Estimating transfer entropy in continuous time between neural spike trains or other event-based data. *PLoS Comput. Biol.* 17, e1008054. doi:10.1371/journal.pcbi.1008054

Shpitser, I., and Pearl, J. (2008). Complete identification methods for the causal hierarchy. *J. Mach. Learn. Res.* 9, 1941–1979.

Solo, V. (2016). State-space analysis of Granger-Geweke causality measures with application to fMRI. *Neural comput.* 28, 914–949. doi:10.1162/NECO_a_00828

Stokes, P. A., and Purdon, P. L. (2017). A study of problems encountered in Granger causality analysis from a neuroscience perspective. *Proc. Natl. Acad. Sci. U. S. A.* 114, E7063–E7072. doi:10.1073/pnas.1704663114

Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends cognitive Sci.* 3, 151–162. doi:10.1016/s1364-6613(99)01299-1

Vicente, R., Wibral, M., Lindner, M., and Pipa, G. (2011). Transfer entropy—A model-free measure of effective connectivity for the neurosciences. *J. Comput. Neurosci.* 30, 45–67. doi:10.1007/s10827-010-0262-3

Wen, X., Rangarajan, G., and Ding, M. (2013). Multivariate granger causality: An estimation framework based on factorization of the spectral density matrix. *Philosophical Trans. R. Soc. A Math. Phys. Eng. Sci.* 371, 20110610. doi:10.1098/rsta.2011.0610

Wibral, M., Pampu, N., Priesemann, V., Siebenhühner, F., Seiwert, H., Lindner, M., et al. (2013). Measuring information-transfer delays. *PLoS ONE* 8, e55809. doi:10.1371/journal.pone.0055809

Wibral, M., Vicente, R., and Lindner, M. (2014). Transfer entropy in neuroscience. *Directed information measures in neuroscience* (Berlin, Heidelberg: Springer).

Wismüller, A., Dsouza, A. M., Vosoughi, M. A., and Abidin, A. (2021). Large-scale nonlinear granger causality for inferring directed dependence from short multivariate time-series data. *Sci. Rep.* 11, 7817. doi:10.1038/s41598-021-87316-6

Wollstadt, P., Lizier, J. T., Vicente, R., Finn, C., Martinez-Zarzuela, M., Mediano, P., et al. (2019). IDTxl: The information dynamics toolkit xl: A python package for the efficient analysis of multivariate information dynamics in networks. *J. Open Source Softw.* 4, 1081. doi:10.21105/joss.01081

Woodward, J. (2001). *Causation and manipulability*. Stanford: Stanford Encyclopedia Of Philosophy.

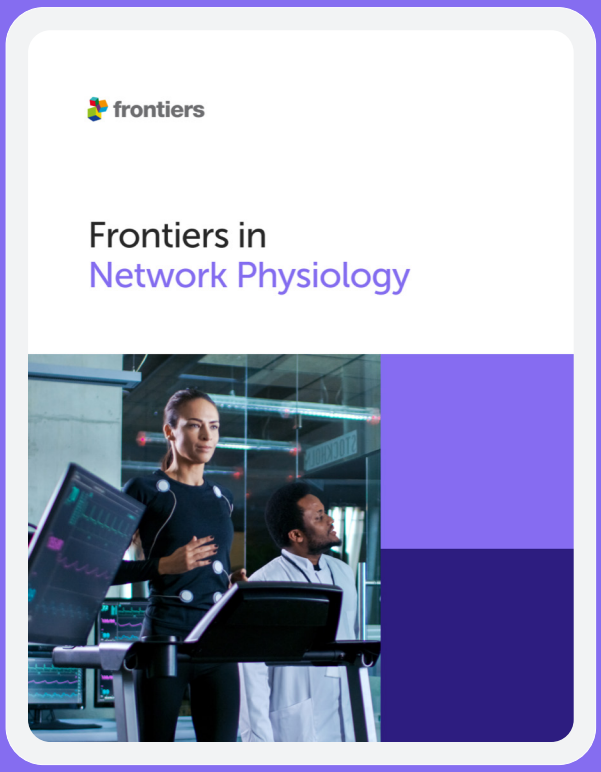
Frontiers in
Network Physiology

Explores how diverse physiological systems and organs interact

The first journal to focus on the mechanisms through which systems and organs interact and integrate to generate a variety of physiologic states.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact