A Novel Methodology to Study Synchrony, Causality and Delay in EEG Data

Roberto Cruz Oropesa[1], Oscar Dalmau[1], José L. Marroquin[1], Thalia Harmony[2]

[1] Center for Research in Mathematics,
Mexico

[2] Universidad Nacional Autónoma de México,
Instituto de Neurobiologia,
Mexico

{robertoc, dalmau, jlm}@cimat.mx

Abstract. Synchrony and causality analyses performed in EEG data have improved the understanding of complex interactions within the human brain. However, few attempts have been conducted for using the delay magnitude as a feature of synchrony events. A new methodology for studying synchrony in EEG data is presented here. It includes synchrony detection and a novel mechanism to estimate delay magnitude and sign - hence causality - between narrow bandwidth signals. Synchrony detection and delay estimation are separated in two steps: first, significant synchrony is detected using a measure based on phase differences, then, delay is estimated by analyzing the dispersion of measure maxima in the space spanned by time, frequency and delays. Synthetic EEG data is used to validate the methodology using a synchrony spectral model with controlled bandwidth and multivariate autoregressive models (MVAR). The proposed methodology achieves a superior performance in causality estimation than state-of-the-art techniques in accuracy and robustness to noise. We also present an analysis of data from a psychophysiological experiment of figure categorization. This methodology provides a reliable method to estimate the delay magnitude of synchrony events and it is a better alternative for studying causality than the state-of-the-art techniques employed here.

Keywords. Delay estimation for synchrony events, synchrony detection, causality analysis, time-frequency analysis.

1 Introduction

The study of interactions among brain areas has been conducted for achieving a better understanding of brain function. EEG has been used in these studies because it allows one to observe a dynamic evolution of complex connectivity networks, based on the analysis of electric potential series recorded at several electrode sites. Besides, an increasing interest to establish causality or information flow associated to a synchrony relation has existed since the latter part of the previous century [31]. These studies still are conducted in recent neurophysiological research [25, 34, 6].

Often, a synchrony event is represented by its time and/or frequency localization, extracted from event detection, and its causality. The use of the associated delay magnitude as an additional element to characterize a synchrony relation allows a more complete representation of an event, since it allows for a better discrimination among synchrony relations that occur in a same region of the time-frequency plane and can be used to find evidence of a single event in several electrode pairs and thus, to construct a compact representation that could lead to a better understanding of a relation, allowing one to observe its associated connectivity network in a summarized way.

A new methodology for synchrony and causality analysis in EEG data is presented in this work,

that provides a complete characterization of a synchrony relation. The methodology is based on the detection of significant synchrony events using a measure based on phase differences between the narrow-band signals that come from a time-frequency decomposition, and on the subsequent analysis of the dispersion of maxima of this measure in the space spanned by time, frequency and possible delays to estimate the associated delay magnitude. We show that this methodology constitutes a better alternative for causality estimation than state-of-the-art methods, besides the fact that it provides more information than those measures (i.e., the delay magnitude).

2 Related Work

Cross-correlation function [30], h^2 index [22], Granger causality [11] and mutual information [8] have been used to detect synchrony in the time domain. Time localization and causality are determined by the first three measures, while cross-correlation [17, 39] and h^2 [39] have been used to estimate delay magnitude. However, localizing a synchrony event in frequency is not possible with any of these time domain measures.

Other measures are designed to study synchrony in the frequency domain, like: coherence [9], coherence corrected for zero-lag connections[26, 27], the imaginary part of normalized cross-spectrum [23], partial coherence [29], Noise Contribution Ratio (NCR) [1], Directed Transfer Function (DTF) [18], direct Directed Transfer Function (dDTF) [20], Partial Directed Coherence (PDC) [3], generalized Partial Directed Coherence (gPDC) [4] and isolated effective coherence (iCoh) [28]. Frequencies involved in a synchrony event are estimated by all these measures and the causality estimation is provided by all but coherence, the imaginary part of normalized cross-spectrum and partial coherence. However, these frequency domain techniques are not able to estimate time localization or delay magnitude of a synchrony relation.

Phase Slope Index (PSI) is proposed for studying synchrony and causality in the frequency domain [24], by analyzing the slope of the phase of normalized cross-spectrum. PSI is defined by:

$$PSI_{xy}(\Omega) = \mathsf{Im}\left(\sum_{\omega \in \Omega} C_{xy}^*(\omega) C_{xy}(\omega + \Delta\omega)\right),$$

where ω is expressed in radians per sample, $\mathsf{Im}(.)$ is the imaginary part of a complex number, the operator $(.)^*$ indicates the complex conjugate, $C_{xy}(\omega)$ represents the normalized cross-spectrum between signals x and y, Ω is a set of frequencies and $\Delta\omega$ refers to the distance between adjacent frequencies expressed in radians per sample. If PSI has a positive value then causality is $x \rightarrow y$ for frequencies in Ω, while if PSI is negative then $x \leftarrow y$. If a synchrony relation does not exist or a synchrony event has zero phase difference, then PSI gets close to zero. However, PSI cannot estimate the time interval or the associated delay magnitude for a detected relation.

Other techniques like time-frequency coherence [2], Phase Lock Value (PLV) [21], Magnitude of Phase Difference (MPD) [2], Phase Lag Index (PLI) [32] and Weighted Phase Lag Index (WPLI) [38] have been proposed for the time-frequency plane. In [2], normalized cross-spectrum estimation is performed over a sliding time window to calculate time-frequency coherence. Time and frequency localization is performed by time-frequency coherence, PLV and MPD while PLI and WPLI provide frequency localization only. On the other hand, MPD is designed to detect relations with zero phase difference, hence an event cannot be detected if its associated phase difference is distant from zero. These measures do not estimate causality or delay magnitude.

The possibility to estimate causality based on the asymmetry of the phase difference sign distribution is discussed in [32] and it is the basic concept of directed Phase Lag Index (dPLI) [33], a measure defined for each frequency ω as:

$$dPLI(\omega) = \frac{1}{N} \sum_{t=1}^{N} H(\mathsf{sgn}\left(\Delta\phi_{xy}(t, \omega)\right)),$$

where $\Delta\phi_{xy}(t, \omega)$ is the instantaneous phase difference, wrapped into the interval $(-\pi, \pi]$, between the signals that are obtained from a

Table 1. Measures employed to study synchrony and causality in EEG data. The table presents a summary of the information each technique provides regarding a synchrony event. Horizontal lines delimit measures proposed in the same domain i.e., time domain, frequency domain and time-frequency domain, from top to bottom. The legend corresponding to the Event Localization column is: t (time localization), f (frequency localization), tf (time and frequency localization) and n (none)

Measure	Event localization	Causality	Delay
cross-correlation	t	yes	yes
h^2	t	yes	yes
granger causality	t	yes	no
mutual information	n	no	no
coherence	f	no	no
Im Coherency	f	no	no
PSI	f	yes	no
partial coherence	f	no	no
NCR	f	yes	no
DTF	f	yes	no
dDTF	f	yes	no
PDC	f	yes	no
gPDC	f	yes	no
iCoh	f	yes	no
time-frequency coherence	tf	no	no
PLV	tf	no	no
MPD	tf	no	no
PLI	f	no	no
WPLI	f	no	no
dPLI	f	yes	no

time-frequency decomposition of signals x and y for time t and frequency ω; $H(.)$ corresponds to Heaviside step function. dPLI ranges from 0 to 1 and its reference value to estimate causality is 0.5. If dPLI has a value significantly lower than 0.5 then information flow is $x \leftarrow y$, while if dPLI is significantly greater than 0.5 causality is $x \rightarrow y$. dPLI provides the frequency localization of a relation, but does not estimate its time localization or delay magnitude.

As mentioned before, none of the above measures is able to extract complete information about a synchrony event, as can be seen in Table 1.

3 Methods

We call the proposed methodology *Simultaneous Synchrony and Delay Estimation* (SSDE) because it detects synchrony and estimates its associated delay, although the detection stage is isolated from the delay estimation procedure. SSDE is designed to analyze pairs of electrode series with several trials $x(t, r)$ and $y(t, r)$ using their time-frequency decompositions $X(t, \omega, r)$ and $Y(t, \omega, r)$. The definition of a plausible delay set is required by SSDE and it is defined as the interval $\Gamma = [-a, a]$, where $a \in \mathbb{Z}$ is measured in samples. A proper selection of a depends on data sampling frequency because delays in Γ should be reasonable when they are expressed in milliseconds.

A time-frequency decomposition for each electrode time series trial is obtained passing each trial signal through a quadrature filter bank where each filter is centered at a given frequency, defining decomposition frequency bands [12]. The output of each filter is a complex signal that has non-zero amplitude in a frequency interval (ω_0, ω_2) where filter center frequency ω_1 satisfies $\omega_0 < \omega_1 < \omega_2$ (see [12] for details).

In the synchrony detection step, first, we calculate a synchrony measure in the time-frequency domain; then, significant synchrony is detected by performing multiple hypothesis tests and finally, we group significant synchrony points into synchrony regions in the time-frequency plane. In the delay estimation procedure, we compute the delay associated to each detected synchrony event.

3.1 Synchrony Analysis

Let $\rho(t, \omega)$ be a synchrony measure. SSDE main idea for synchrony detection is to apply a delay $\tau \in \Gamma$ to a time-frequency decomposition signal $Y(t, \omega, r)$, trying to maximize the synchrony between $X(t, \omega, r)$ and $Y(t + \tau, \omega, r)$, according to ρ. The selected measure ρ is MPD, which for $X(t, \omega, r)$ and $Y(t + \tau, \omega, r)$ is defined by [2]:

$$\rho(t, \omega, \tau, r) = 1 - \frac{1}{\pi} \left| \mathsf{wrap}\left(\phi_x(t, \omega, r) - \phi_y(t + \tau, \omega, r)\right) \right|, \quad (1)$$

where ϕ_x and ϕ_y indicates instantaneous phase at time t, frequency ω and trial r for each signal, while the wrap(.) operator wraps phase difference into the interval $(-\pi, \pi]$.
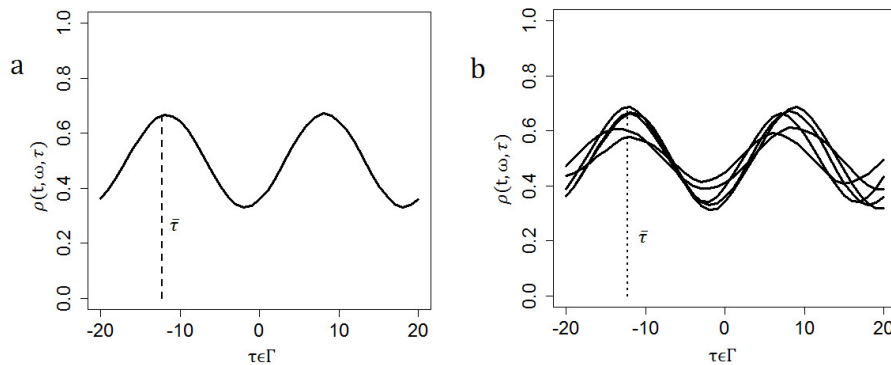
Fig. 1. Panel a: Curve of $\rho(t_0, \omega_0, \tau)$ (see equation 2) vs $\tau \in \Gamma$ for a given (t_0, ω_0) with significant synchrony. One observes two local maximum, where $bar\tau$ represents the real delay. Panel b: Curves $\rho(t_0, \omega_i, \tau)$ vs $\tau \in \Gamma$ for a given t_0 and frequency bands ω_i around ω_0. Discontinuous lines indicate the true delay $\bar{\tau}$. Local maximum are aligned around the real delay

An oscillatory behavior with respect to $\tau \in \Gamma$ is expected in $\rho(t, \omega, \tau, r)$ because of the periodic nature of the wrapped phase difference.

Then, the mean over trials is calculated for each triplet (t, ω, τ) as:

$$\rho(t, \omega, \tau) = \frac{1}{R} \sum_{r=1}^{R} \rho(t, \omega, \tau, r). \tag{2}$$

Finally, the proposed synchrony measure ρ^* for the time-frequency plane is defined by:

$$\rho^*(t, \omega) = \max_{\tau} \rho(t, \omega, \tau). \tag{3}$$

For each (t, ω) in the time-frequency plane, the maximum degree of phase synchronization between $X(t, \omega, r)$ and $Y(t, \omega, r)$ is quantified by ρ^*. The detection of significant synchrony requires to perform multiple hypothesis tests where the test statistic is ρ^*. Note that MPD in [2] only detects synchrony with zero delay. Thus, ρ^* may be considered as an extension of MPD that detects synchrony events with any associated phase difference.

Once significant synchrony points (t, ω) in the time-frequency plane are detected, then synchrony connected regions are built using a region growing algorithm considering an 8-point neighborhood. It is assumed that a single synchrony event is represented by each connected grown region.

Hence, the time and frequency localization of a synchrony event are given by the localization of the associated region in the time-frequency plane.

3.2 Delay Estimation

Synchrony measure $\rho(t, \omega, \tau, r)$ exhibits an oscillatory behavior with respect to the delays $\tau \in \Gamma$, because of the circular nature of the wrapped phase difference. Therefore, a delay estimation based on the delay that maximizes ρ^* is not reliable. An example of this oscillatory behavior is observed in Fig. 1 Panel a.

In formal terms, for a given frequency ω_1 could exist delays $\tau_1, \tau_2 \in \Gamma$ that satisfy:

$$\exists k \in \mathbb{Z} : \omega_1 \tau_1 = \omega_1 \tau_2 + 2\pi k.$$

We call τ_1 and τ_2 periodic delays for frequency ω_1. It follows from equation (1) that:

$$\rho(t, \omega_1, \tau_1) = \rho(t, \omega_1, \tau_2).$$

The distance $\Delta \tau$ between periodic delays τ_1 and τ_2 for a frequency ω is given by:

$$\Delta \tau(\omega) = |\tau_1 - \tau_2| = \frac{2\pi}{\omega_1} |k|.$$

If a set of frequencies $\{\omega_i\}$, where a synchrony relation exists for a given t_0, is inspected then it is observed that local maxima of $\rho(t_0, \omega_i, \tau)$ (see

equation 2) with respect to $\tau \in \Gamma$ are found close to the real delay $\bar{\tau}$ for all frequencies. However, local maxima that correspond to periodic delays with respect to $\bar{\tau}$ are going to be more dispersed because $\Delta\tau(\omega)$ depends on frequency. This phenomenon is shown in Fig. 1 Panel b, where it is observed that local maxima located at the left portion of the X axis have a better alignment than local maxima at the right portion.

Our delay estimation procedure is based on the previous observation and studies the delay concentration over a connected region C in the time-frequency plane, that represents a synchrony event: first, possible delays for C are estimated, e.g., as local maxima of $\rho(t, \omega, \tau)$, considering only Γ. Then, the most concentrated delay is selected as estimate for the real delay. Note that this procedure requires that a non-zero frequency bandwidth is involved in each synchrony event.

Let ω and $\omega + \Delta\omega$ be frequencies with $\Delta\omega > 0$, τ_1 be a delay, τ_2 and τ_3 be periodic delays with respect to τ_1 for ω and $\omega + \Delta\omega$, respectively. The distance between τ_2 and τ_3 is referred to as the *movement of periodic delays*, it is denoted by $\Delta\tau_{2\pi}(\omega, \omega + \Delta\omega)$ and it is defined as:

$$\Delta\tau_{2\pi}(\omega, \omega + \Delta\omega) = \Delta\tau(\omega) - \Delta\tau(\omega + \Delta\omega).$$

Direct calculation shows that $\Delta\tau_{2\pi}(\omega, \omega + \Delta\omega)$ for periodic delays with a same $k \in \mathbb{Z}$ is:

$$\Delta\tau_{2\pi}(\omega, \omega + \Delta\omega) = \frac{2\pi\Delta\omega\,|k|}{\omega(\omega + \Delta\omega)}. \qquad (4)$$

Dispersion between periodic delays at two different frequencies is quantified by expression (4), that indicates a decrease in spread as frequencies increase. This implies that delay estimation will be more difficult as frequencies increase.

In practice, it is not necessary to construct the curves $\rho(t, \omega, \tau)$ vs τ and then find the local maxima. Rather, if the wrapped phase difference $\Delta\phi_{xy}^e$ is estimated, one can find the candidate delays for each (t, ω) as:

$$\tau_k(t, \omega) = \frac{\Delta\phi_{xy}^e + 2\pi k}{\omega}, \qquad (5)$$

where the integers k are selected so that all candidate delays are within the interval Γ. In particular, one may take $k \in [k_{\min}(t, \omega), k_{\max}(t, \omega)]$ where:

$$k_{\min}(t, \omega) = \text{Int}\left[\frac{(-a - \tau_0(t, \omega))\,\omega}{2\pi}\right], \qquad (6)$$

$$k_{\max}(t, \omega) = \text{Int}\left[\frac{(a + \tau_0(t, \omega))\,\omega}{2\pi}\right], \qquad (7)$$

where the Int[.] operator takes the integer part of its argument. The mean direction $\overline{\Delta\phi_{xy}^e}$ may be estimated by averaging the wrapped phase difference direction over trials:

$$\overline{\Delta\phi_{xy}^e}(t, \omega) = \arctan\left(\frac{\sum\limits_{r=1}^{R} \sin \Delta\phi_{xy}^e(t_0, \omega_0, r)}{\sum\limits_{r=1}^{R} \cos \Delta\phi_{xy}^e(t_0, \omega_0, r)}\right). \qquad (8)$$

By repeating this procedure over all $(t, \omega) \in C$, one obtains groups G_k of candidate delays -one for each value of k- and the representative of the group with the least dispersion may be selected as the estimated delay τ^*.

These representatives $m(k)$ may be found as modes of a kernel estimator for their distribution P_k:

$$P_k(\tau) = \sum_{(t, \omega) \in C} K_h(\tau - \tau_k(t, \omega)), \qquad (9)$$

where K_h is a kernel with bandwidth h. The modes are:

$$m(k) = \arg\max_\tau P_k(\tau),$$

and may be found by direct search over a discretization of Γ. Note that the height of these local modes represents a robust estimate for their associated dispersions -as the height goes lower then greater is the dispersion- , so that the highest mode may be selected as the estimate τ^*. A synchrony event delay magnitude is estimated by $|\tau^*|$, while τ^* sign indicates causality. If $\tau^* > 0$ then $x \rightarrow y$, while if $\tau^* < 0$ then $x \leftarrow y$.

The correct implementation of this procedure requires an appropriate estimation for the kernel bandwidth h. For instance, if one chooses a Gaussian kernel:

$$K_h(x) = \exp\left(-0.5x^2/h^2\right),$$

then h may be obtained using Silverman's rule:

$$h = 1.06\hat{\sigma}n^{-1/5}, \tag{10}$$

where $\hat{\sigma}$ is calculated as the standard deviation of the delay group G_0 and n is its cardinality. The complete procedure is summarized in Algorithm 1.

The results are improved if one uses a resampling procedure over trials to produce a robust delay estimator. To do this, one applies Algorithm 1 N_c times using a random subset of N_r trials each time and then one obtains the final estimator as the median of these N_c results.

Algorithm 1 Delay estimation over a connected synchrony region

---

Require: Connected region with significant synchrony C, Γ.
Ensure: Delay estimator τ^* for the synchrony relation represented by C.
 1: **for all** $(t_0, \omega_0) \in C$ **do**
 2:     Find the mean direction of the wrapped phase difference $\overline{\Delta\phi^e_{xy}}(t, \omega)$ using Eq. (8).
 3:     Compute $k \in [k_{\min}(t, \omega), k_{\max}(t, \omega)]$ using Eqs. (6), (7).
 4:     Calculate $\tau_k(t, \omega)$ for all $k \in [k_{\min}(t, \omega), k_{\max}(t, \omega)]$ using Eq. (5) with $\overline{\Delta\phi^e_{xy}}(t, \omega)$.
 5: **end for**
 6: Compute the kernel bandwidth h using Eq. (10).
 7: Find $\bar{k}_{\min} = \min\limits_{(t,\omega)\in C} k_{\min}(t, \omega)$ and $\bar{k}_{\max} = \max\limits_{(t,\omega)\in C} k_{\max}(t, \omega)$.
 8: **for all** integer $k \in [\bar{k}_{\min}, \bar{k}_{\max}]$ **do**
 9:     Compute $P_k(\tau)$ for τ in a suitable discretization of Γ.
10:     Calculate the mode $m(k)$ and $H(k) = P_k(m_k)$.
11: **end for**
12: Find $k^* = \arg\max\limits_k H(k)$.
13: Set $\tau^* = m(k^*)$.

3.3 Experiments

In this section we describe a set of experiments conducted to characterize the performance of SSDE and compare it with other techniques. PSI and dPLI are selected as reference measures for synchrony detection and causality estimation because they are causality techniques that could be extended to give time-frequency information in a simple way, facilitating a fair comparison with the proposed methodology.

Time-frequency coherence also is selected as reference measure because is widely used for detecting synchrony in EEG data. A method based on MVAR models is also employed as a reference measure because the most used techniques to estimate causality are based on MVAR models [1, 11, 18, 20, 3, 4, 28]. In PSI extension, the time-frequency estimation of the normalized cross-spectrum $C_{xy}(t, \omega)$ is performed over trials, and the asymmetry of the phase difference sign distribution is also computed over trials. Time-frequency PSI extension is defined as:

$$PSI(t, \omega) = \sum_{\omega' \in \Omega(\omega)} \mathsf{Im}\{C^*_{xy}(t, \omega')C_{xy}(t, \omega' + \Delta\omega)\}, \tag{11}$$

where $\Omega(\omega)$ is a 2Hz frequency window, centered at ω. Time-frequency dPLI extension is:

$$dPLI(t, \omega) = \frac{1}{R}\sum_r H(\Delta\phi_{xy}(t, \omega, r)). \tag{12}$$

Time-frequency coherence is calculated by:

$$c(t, \omega) = |C_{xy}(t, \omega)|, \tag{13}$$

where $C_{xy}(t, \omega)$ estimate is calculated over trials. Since coherence detects synchrony but does not estimate causality then it is used only for studying detection capability. MVAR models are defined by the general expression:

$$\vec{x}(t) = \sum_{l=1}^{p} A(t, l)\vec{x}(t - l) + \vec{\xi}(t), \tag{14}$$

where $\vec{x}(t) = [x_1(t), x_2(t), \ldots, x_d(t)]^T$ is a vector whose elements are time series corresponding to electrodes or selected regions of interest, $A(t, l)$ are the coefficient matrices of the model, p indicates the model order and $\vec{\xi}(t)$ represents an uncorrelated noise vector.

In the MVAR method considered here (denoted in what follows as Time Dependent MVAR or

TDMVAR), coefficients matrices $A(t, l)$ for a given instant t_0 are estimated by solving a least squares problem over trials and over a time window $v(t)$ centered at t_0:

$$\bar{A}(t_0, l) = \arg\min_{A(l)} \sum_{r=1}^{R} \sum_{t \in v(t)} D(t, r, A(l)),$$

$$D(t, r, A(l)) = \left\| \vec{x}(t, r) - \sum_{k=1}^{p} A(l)\vec{x}(t - l, r) \right\|_2^2,$$

where r denotes the trial number, and it is assumed that the time series are organized in R trials with the same amount of samples. Note that since one wishes to detect synchrony events that are localized at unknown time locations, frequency domain MVAR methods are not adequate. We employ two models for synchrony simulation: a spectral model that produces narrow-band synchronization and an MVAR model.

3.3.1 Generation of Simulated EEG Data

Synchrony relations are simulated with two kinds of generators: an MVAR model and a novel synchrony spectral model that keeps synchrony bandwidth properly controlled. The last one is employed to study synchrony events with narrow bandwidth.

Data generation with MVAR models is performed using equation (14), with time invariant coefficient matrices $\{A(l)\}_{l=1}^{p}$ and a Gaussian white noise vector $\vec{\xi} \sim N_d(\vec{0}, I_{d \times d})$. During the generation process, the first p values of each time series are generated as white noise while the remaining data are obtained from equation (14). The first 100 values of each time series are not included and a total of 102400 valid instants are generated for every time series, that are organized in 200 trials consisting of 512 samples each.

3.3.2 Proposed Synchrony Spectral Model with Controlled Synchrony Bandwidth

Signal generation with this proposed model is based on the Fourier representation of a signal. The generation process is a modification of the Surrogate Data algorithm [35] and it allows to model power spectra variations over trials. Let Ω be a frequency set formed by the DC and the positive frequencies of the discrete Fourier transform, $A(\omega)$ be a mean amplitude spectrum calculated over real EEG data and $B_x(\omega)$ be a random positive variable that satisfies:

$$E[B_x(\omega)] = A(\omega).$$

Let $\phi_x(\omega)$ be a random variable with uniform distribution over $[0, 2\pi)$. The synthetic EEG signal $\tilde{x}(t)$ is obtained by:

$$\tilde{x}(t) = \sum_{\omega \in \Omega} B_x(\omega) \cos(\omega t + \phi_x(\omega)) + \eta(t), \quad (15)$$

where $\eta(t)$ represents additive white Gaussian sensor noise.

Let $\tilde{x}_1(t)$ and $\tilde{x}_2(t)$ be signals generated by model (15) using phase spectrum $\phi_1(\omega)$ and $\phi_2(\omega)$, respectively, $\Omega_1 \subset \Omega$ be a set of frequencies where the synchrony relation is introduced, $v(t)$ be a window function that specifies synchrony time localization, $\bar{\tau}$ be the associated delay and $\gamma \in [0, 1]$ be a parameter controlling synchrony strength. An instantaneous phase function $\phi_{2s}(t, \omega)$ for a signal synchronized with $\tilde{x}_1(t)$ is computed as:

$$\phi_{2s}(t, \omega) = \begin{cases} \begin{aligned} &v(t)(\phi_1(\omega) + \omega\bar{\tau}) \\ &+ (1 - v(t))\phi_2(\omega) \\ &+ I_{[v(t)>0]}\psi(t), \end{aligned} & \omega \in \Omega_1, \\ \qquad\qquad \phi_2(\omega), & \omega \notin \Omega_1, \end{cases}$$

where $I_{[\cdot]}$ is the indicator function and $\psi(t)$ represents phase noise in a synchrony relation. A signal $\tilde{x}_{2s}(t)$ synchronized with $\tilde{x}_1(t)$ is calculated by:

$$\begin{aligned} \tilde{x}_{2s}(t) = \ &\sum_{\omega \in \Omega} B_2(\omega)(1 - \gamma) \cos(\omega t + \phi_2(\omega)) \\ &+ \sum_{\omega \in \Omega} B_2(\omega)\gamma \cos(\omega t + \phi_{2s}(t, \omega)) \\ &+ \eta(t), \end{aligned} \quad (16)$$

where $\eta(t)$ indicates additive white Gaussian sensor noise. A synchrony event is introduced between $\tilde{x}_{2s}(t)$ and $\tilde{x}_1(t)$ with time localization $v(t)$ and frequency localization Ω_1, hence synchrony bandwidth is properly controlled. If $\bar{\tau} > 0$, then $\tilde{x}_{2s}(t)$ precedes $\tilde{x}_1(t)$, while if $\bar{\tau} < 0$ then the

synchrony event causality is $\tilde{x}_{2s} \leftarrow \tilde{x}_1$. When $\gamma = 0$ a synchrony relation does not exist while if $\gamma = 1$ then synchrony strength is maximum.

3.3.3 Performance Indicators

Hypothesis testing is required for detecting synchrony, where the null hypothesis represents the non existence of synchrony. If the null hypothesis is rejected then synchrony is detected. To test this hypothesis, in all cases we estimate the null distribution by a permutation procedure in which the repetitions corresponding to one of the considered electrodes are permuted while those of the other electrode are kept fixed and the statistics of interest are collected [21]. Selected statistics for time-frequency PSI, time-frequency dPLI and time-frequency coherence are the measure values themselves, while off-diagonal coefficients of $A(t, l)$ are the selected statistic for TDMVAR since these coefficients are zero under the null hypothesis. The detection statistic for SSDE is the value of ρ^*.

Causality is determined by comparing a measure value with a reference value (generally 0 although this value is 0.5 for dPLI). Causality estimation with a TMAVR is based on the matrix position of a significant coefficient value.

SSDE estimates an associated delay τ^* over a synchrony region C, and therefore SSDE causality estimation is not a punctual estimator, while time-frequency PSI (equation (11)) and time-frequency dPLI (expression (12)) provide a causality estimation for each point (t, ω). Here, causality determination with time-frequency PSI and time-frequency dPLI is performed by selecting the predominant causality over the modelled synchrony region. If a tie occurs then it is considered as a bidirectional relation that corresponds to delays $\bar{\tau} = 0$, as is discussed in [2][7]. If causality estimation coincides with the modelled synchrony event causality then the causality estimation is considered a success.

For SSDE, if delay estimator τ^* computed over region C satisfies that:

$$|\text{round}(\bar{\tau}) - \text{round}(\tau^*)| \leq 2,$$

where $\bar{\tau}$ is the true delay, then delay estimation is regarded as a success.

Employed performance indicators are:

— True Positives Rate (TPR) to quantify synchrony detection capability.

— Causality Estimation Success Rate (CESR).

— Delay Estimation Success Rate (DESR) that it is used only for SSDE, since it is the only technique with this feature.

A comparison between SSDE and TDMVAR needs to consider that synchrony relations are characterized by SSDE using time-frequency information while TDMVAR provides time information only. If there is a frequency ω where synchrony in (t, ω) is detected by the SSDE methodology, then we say that SSDE detects synchrony in t. A voting over detected frequencies is required to assign a causality for t using SSDE, and the most common causality is selected. The following performance indicators are employed to compare SSDE methodology and TDMVAR:

— Global Temporal Detection Rate (GTDR):

$$\frac{\text{detected instants}}{\text{instants where synchrony exists}}.$$

— Temporal Causality Estimation Success Rate (TCESR):

$$\frac{\text{instants with correct causality estimation}}{\text{instants where synchrony exists}}.$$

— Instantaneous Detection Rate (IDR):

$$\frac{\text{examples where synchrony is detected for an instant } t}{\text{total of examples}}.$$

The first two indicators are calculated as mean rates over the total of examples and their values correspond to detection and causality estimation over the complete time course. A time curve is provided for the third indicator, allowing an observation of detection rates for each instant t when synchrony is restricted to a time interval.

3.3.4 Experiments with Simulated EEG Data

Experiments conducted with simulated EEG data are described in this section. Simulated data is generated using the proposed synchrony spectral model with controlled synchrony bandwidth and MVAR models. Sampling frequency is 200 samples per second and each trial is formed by 512 samples. Every example contains 200 trials and synchrony is located between the 201st and the 311st sample. Associated delays $\bar{\tau}$ for synchrony events are selected randomly between 0 and 13 samples ($\bar{\tau} \in \mathbb{R}$) with causality $x \leftarrow y$.

Time-frequency decompositions are obtained using a bank of quadrature filters tuned at frequencies between 0.5 and 30Hz, where contiguous frequency bands are separated by 0.5Hz and each filter has a frequency support of 2Hz. Correction for multiple hypotheses testing is done by selecting a confidence level α low enough, so that the expected number of false positives is kept under control.

In particular, we chose $\alpha = 5 \times 10^{-5}$, which corresponds to expecting less than one false positive on a time frequency field with 512 time samples and 60 frequency samples. The frequency localization of a synchrony event is described using the central frequency and the synchrony bandwidth. The interval of possible delays is $\Gamma = [-15, 15]$ and its discretization used to find modes $m(k)$ is built using delay groups G_k. When the resampling procedure is used to estimate τ^*, the total of random trial subsets was $N_c = 100$ and each trial subset contains $N_r = R/2$ trials that are selected using bootstrap, where R is the total of trial in each example.

When simulated data are generated by the synchrony spectral model, generated signals are not contaminated with sensor or phase noises and the synchrony strength parameter is $\gamma = 1$. The window size for estimating the temporal MVAR model coefficients is computed as the size of a rectangular window, associated to a uniform distribution with equal variance than the low-pass filter temporal response. This is the general configuration of experiments, unless some different configurations are explicitly described for any parameter of the analysis.

3.3.5 SSDE Reliability Study

In a first experiment, 24 datasets are employed, each one contains 1000 examples synchronized using the synchrony spectral model of equation (16). Central frequencies ranges from 5 to 28Hz and the synchrony bandwidth is 2Hz. In this experiment, reliability of SSDE in relation to delay and causality estimations is studied, including the use of the resampling strategy. Delay estimation is assessed by DESR and causality estimation by CESR.

A second experiment studies the behavior of DESR as the modelled synchrony bandwidth is increased. Generated datasets contain 1000 examples where central frequencies ranges from 5 to 28Hz and synchrony bandwidth varies from 2 to 6Hz.

A third experiment explores the performance of SSDE for different values of the central frequency of the synchrony event as the number of trials is varied. Generated datasets contain 500 examples where synchrony is located at central frequencies 5, 10, 15, 20 and 25Hz with a bandwidth of 2Hz. The number of trials is chosen between 50, 100, 150, 200 and 300 trials.

A fourth experiment compares SSDE and TDMVAR regarding their detection capabilities, using IDR as performance indicator, over data synchronized with the proposed synchrony spectral model of equation (16), when the synchrony bandwidth is narrow (2Hz). Two cases are considered: when synchrony is over the complete time course and when it is limited to a time interval between the 200th and the 310th sample number. A total of 300 examples are synchronized with synchrony's frequency localization between 14 and 16Hz. The associated delays $\bar{\tau}$ for synchrony events are selected randomly between 0 and 13 samples ($\bar{\tau} \in \mathbb{Z}$) with causality $x \leftarrow y$.

3.3.6 SSDE Sensibility to Data Contamination

The noise sensibility study considers two contamination models:

— Noise sensor model, varying the variance of the noise process $\eta(t)$.

— Spurious trials model, where some trials do not have synchrony relations.

For each contamination model, a dataset of 500 examples is built where synchrony events are introduced by the proposed synchrony spectral model (see equation (16)) with central frequencies 10, 15 and 20Hz and 2Hz synchrony bandwidth. Let f_c be a central frequency expressed in Hz where a modelled synchrony event is localized, then the time-frequency decomposition calculates frequency bands in the interval $[f_c - 3.5\text{Hz}, f_c + 3.5\text{Hz}]$. The synchrony measures included in the noise sensibility analysis are SSDE, time-frequency PSI, time-frequency dPLI and time-frequency coherence (see equation (13)); time-frequency coherence is excluded from the causality comparison because it cannot estimate causality. SNR levels considered for the noise sensor model ranges from 1 to 6, and the percentage of spurious trials goes from 0 to 70 for the spurious trials model.

3.3.7 SSDE Performance on Data Generated from MVAR Models

In these experiments, time-frequency decompositions are obtained using a bank of quadrature filters with frequency support of 6Hz. SSDE and TDMVAR are included in this comparative study, where SSDE delay estimation procedure does not include the resampling strategy.

Detection and causality estimation when data are generated from an MVAR model with varying synchrony strength, are studied in the first experiment, where the following MVAR model is employed for generating simulated data:

$$
\begin{aligned}
x_1(t) &= 1.2x_1(t-1) - 0.9x_1(t-2) \\
&\quad +\beta x_2(t-\bar{\tau}) + \xi_1(t), \\
x_2(t) &= 1.45x_2(t-1) - 0.85x_2(t-2) + \xi_2(t),
\end{aligned}
$$

where the true delay $\bar{\tau} \in \mathbb{Z}$ satisfies $1 \leq \bar{\tau} \leq 5$ and parameter β controls synchrony strength. Values of β on interval $[0.01, 0.09]$ are analyzed in this experiment. A set of 500 examples is generated for each value studied of β, selecting randomly the true delay $\bar{\tau}$ for each example. GTDR is used as

detection indicator and TCESR is employed for the causality comparison.

A second experiment explores the detection capabilities of both techniques when synchrony is confined to a certain temporal window, employing IDR as performance indicator. The data for this experiment are generated using the following MVAR model:

$$
\begin{aligned}
x_1(t,r) &= 1.42x_1(t-1,r) - 0.85x_1(t-2,r) \\
&\quad +v(t)x_2(t-5,r) + \xi_1(t,r), \\
x_2(t,r) &= 1.8x_2(t-1,r) - 0.96x_2(t-2,r) \\
&\quad +\xi_2(t,r),
\end{aligned}
$$

where $v(t)$ is defined by:

$$
v(t) = \left\{ \begin{array}{ll} \beta, & t \in [200, 310], \\ 0, & t < 200 \vee t > 310, \end{array} \right.
$$

for every trial r. Two values are considered for parameter β: 0.2 and 0.07. The synchrony relation $x_2(t) \rightarrow x_1(t)$ is presented in the time interval $[200, 310]$, instead of over the complete time course as in the previous experiment. We employ a different base MVAR model because the peaks of energy spectrums are located in lower frequencies than the previous MVAR model, allowing to study synchrony in lower frequencies.

3.3.8 Experiments with Real EEG Data

A figure categorization experiment [14] is studied using SSDE. A series of figures representing an animal or other object are presented to subjects (18 children, 9 females, 8 to 10 years old, right-handed and with a normal neurological examination), that are instructed to press a button if a figure represents an animal that its name starts with a vowel. If the name of a represented animal begins with a consonant, then subjects have to press another button. In the case when a figure represents another kind of object, subjects do not press a button.

Activity is recorded from 1000 milliseconds before the presentation of stimulus to 1560 milliseconds after its onset using a MEDICID 3E system with a sampling frequency of 200 samples per second. The inter-stimulus interval was 5 seconds. Subjects are seated in a

comfortable chair in front of the videomonitor. Stimuli are delivered by a MINDTRACER system synchronized to the MEDICID 3E acquisition system. The EEG signals are registered on 20 electrode sites corresponding to the 10/20 system with reference to linked ears: Fp1, Fp2, F7, Fz, F3, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, Oz and O2. EOG is obtained from a supraorbital electrode and from another electrode located on the external canthus of the right eye.

The amplifier bandwidth is selected between 0.5 and 30Hz. Each trial is visually edited and only those corresponding to correct responses and without artifacts are analyzed.
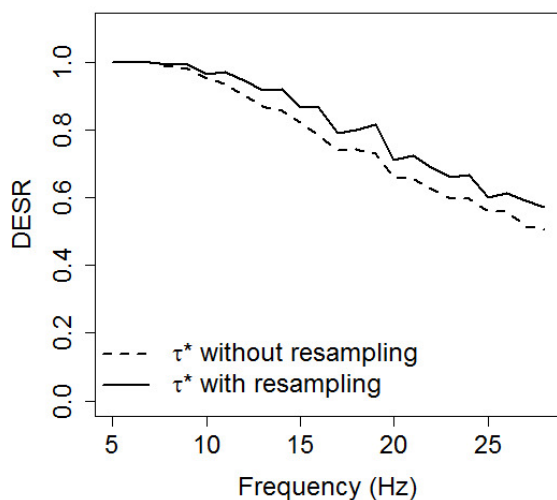
Fig. 2. Curves of Delay Estimation Success Rate (DESR) vs central frequency of synchrony events, calculated for a relation with 2Hz bandwidth. The curves correspond to delay estimation with and without resampling

The experiment is divided in 4 experimental conditions (animal starting with a vowel, animal starting with a consonant, not animal starting with a vowel, not animal starting with a consonant) and we present here results corresponding to the condition where the stimuli correspond to animals starting with a consonant, that has a total of 270 trials. We perform two separate analysis: one

corresponding to the pre-stimulus – where the valid time window is 430-995 milliseconds – and one studying the post-stimulus, where the valid time window is 1005-2125 milliseconds or 5-1125 milliseconds after the stimulus onset.

The time-frequency decomposition for each signal is computed using a bank of quadrature filters centered at frequencies between 0.5 and 30Hz, where contiguous frequency bands are separated by 0.5Hz and each filter has a frequency support of 2Hz. To correct for multiple hypotheses testing, false positives are restricted to a level $\alpha = 10^{-7}$, corresponding to expecting 1 false positive pixel for the complete time-frequency plane over all possible electrode pairs in the analysis.

Delay estimation τ^* is obtained combining the resampling strategy and the other algorithm (Algorithm 1), where the resampling strategy is employed when synchrony events have a bandwidth lower than 4Hz and the delay estimation without resampling is used when synchrony relations are extensive in frequency (bandwidth of or greater than 4Hz). This combined delay estimation technique is employed because the resampling procedure takes too much time to estimate delays when regions are too extensive within the time-frequency decomposition, like volume conduction regions. When the resampling strategy is applied, we employ 100 samplings with replacement, each one conformed by 135 trials, to estimate the associated delay. A set of possible delays $\Gamma = [-15, 15]$ is employed, that corresponds to consider delays with a maximum magnitude of 75 milliseconds.

4 Results

Results obtained with simulated and real EEG data are presented in this section.

4.1 SSDE Reliability Study

A plot of DESR vs synchrony's central frequency for delay estimation with and without resampling is shown in Fig. 2. A decreasing behavior is observed in both curves when the central frequency of synchrony events is increased. Note that the delay estimation procedure with resampling improves

performance for all frequencies. If the synchrony bandwidth is 2Hz then a reliable delay estimator is provided by SSDE when synchrony is located in the delta, theta and alpha bands, having a DESR greater than 80% for all frequencies lower than 18Hz.

In Fig. 3, a comparative between SSDE with re-sampling, time-frequency PSI and time-frequency dPLI, for synchrony events with bandwidth of 2Hz is presented, regarding causality estimation. Note that the best performance was achieved by SSDE when central frequencies are lower than 22Hz, while SSDE and PSI had a similar performance for higher frequencies. In both cases CESR was greater than 85% for all frequencies.
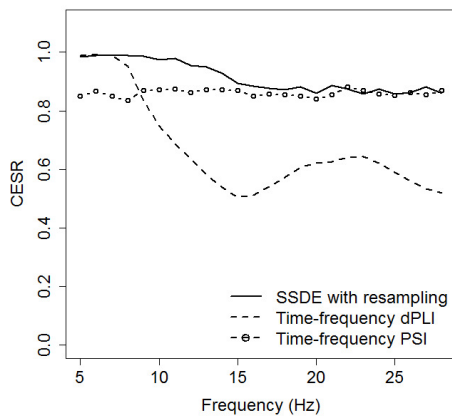
Fig. 3. Curves of Causality Estimation Success Rate vs central frequency of synchrony events, calculated for synchrony relations with 2Hz bandwidth. The curves correspond to SSDE with resampling, time-frequency PSI and time-frequency dPLI

Fig. 4 displays the values of DESR for SSDE with resampling, calculated when the synchrony bandwidth was varied from 2 to 6Hz. Note that delay estimations were correct in more than 95% of the examples when the synchrony bandwidth is equal or higher than 4Hz for all frequencies. The worst case corresponded to a synchrony bandwidth of 2Hz, therefore this bandwidth is used in all experiments as representatives of the worst case for delay and causality estimations.
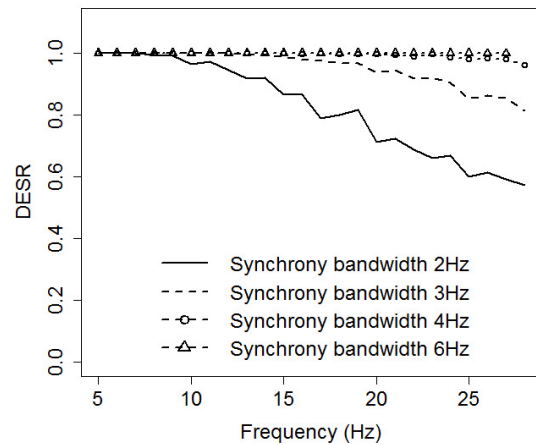
Fig. 4. Curves of DESR vs central frequency of synchrony events, computed for synchrony relations with varying bandwidth. Results were achieved by SSDE with resampling

Results calculated over examples with different number of trials are shown in Fig. 5. Reliability of delay estimation increased as the total number of trials was increased, though the difference in performance became smaller when the examples contained 200 or more trials. If a synchrony event was located at a central frequency equals or lower than 15Hz then a good delay estimate (over 83% of correct delay estimations) was achieved using 150 trials or more. For all studied frequencies, good delay estimations were obtained using 200 trials, therefore this value is employed in the next experiments.

In Fig. 6, IDR curves corresponding to SSDE and TDMVAR are shown, when synchrony events were introduced with a synchrony spectral model, where synchrony was located between 14 and 16Hz (central frequency 15Hz and a synchrony bandwidth of 2Hz). Note that SSDE detection was correct when synchrony was located over the complete time course (Panel a) or it was limited to some time window (Panel b), though detection slightly decreased around the time window boundaries. TDMVAR, however, failed to detect synchrony in both cases.
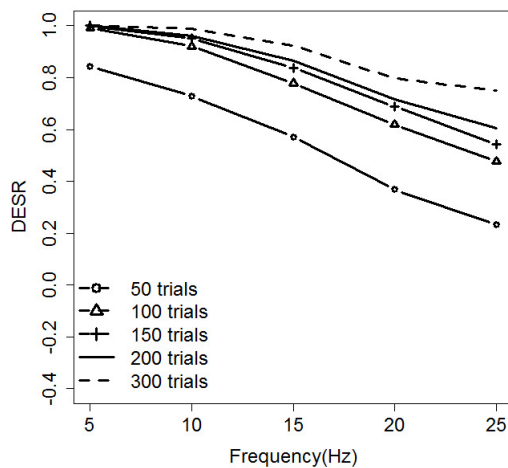
Fig. 5. Curves of DESR vs central frequency of synchrony events calculated for synchrony events with 2Hz bandwidth, located in several central frequencies, when the number of trials was varied. Results were obtained by SSDE with resamping

4.2 Study of Sensibility to Data Contamination

TPR vs SNR curves obtained for the sensor noise model are presented in Fig. 7. The best results for synchrony detection were achieved by time-frequency coherence, although competitive results were provided by SSDE, while time-frequency PSI and time-frequency dPLI gave worse results.

Causality estimation performance (CESR vs SNR curves) for the sensor noise model are presented in Fig. 8. The best performance was achieved by SSDE while time-frequency PSI became more competitive for high central frequencies.

Synchrony detection performances (TPR) calculated for synchrony events contaminated with the spurious trials model is displayed in Fig. 9. The best detection power was achieved by time-frequency coherence and SSDE had competitive results. For all measures, performance did not seen to depend on frequency.

Causality estimation performances computed for varying levels of spurious trials contamination is presented in Fig. 10. SSDE obtained the best performance of all compared measures, for all

central frequencies. SSDE performance deteriorated for high frequencies, while time-frequency PSI performance was less affected when central frequencies were higher.

## 4.3 SSDE Performance on Data Generated from MVAR Models

The results in Fig. 11 were obtained in the study performed on data generated with an MVAR model when the synchrony strength was varied (parameter β) and it was localized over the complete time course. The detection capability was measured by GTDR and the quality of causality estimates was established by TCESR. Regarding synchrony detection (Fig. 11 Panel a), SSDE was better than TDMVAR when $\beta \leq 0.07$ and their performances were similar when $\beta > 0.07$. In the case of causality estimation, SSDE had a superior performance than TDMVAR and it achieved good results when $\beta \geq 0.03$.

Results from an experiment conducted for data generated with an MVAR model that had strong or weak synchrony relations, where the synchrony was located in a time window, are presented in Fig. 12. When the synchrony was strong ($\beta = 0.2$, Fig. 12 Panel a), SSDE and TDMVAR obtained good results over the synchrony time window, considering that spurious detections outside the synchronization interval were caused by border effects. However, synchrony detections diminished for TDMVAR when the synchrony was weak ($\beta = 0.07$, Fig. 12 Panel b), while SSDE maintained its good performance.

4.4 Results Obtained by SSDE with Real EEG Data

Several synchrony relations are detected by SSDE when it is applied to the real EEG dataset described above.

To summarize the results and facilitate interpretation we performed a clustering between electrodes involved in significant synchrony relations, so that clusters are formed with electrodes connected with small delays (with delay magnitude smaller than 7.5 ms) synchronized with other clusters with some relatively large common
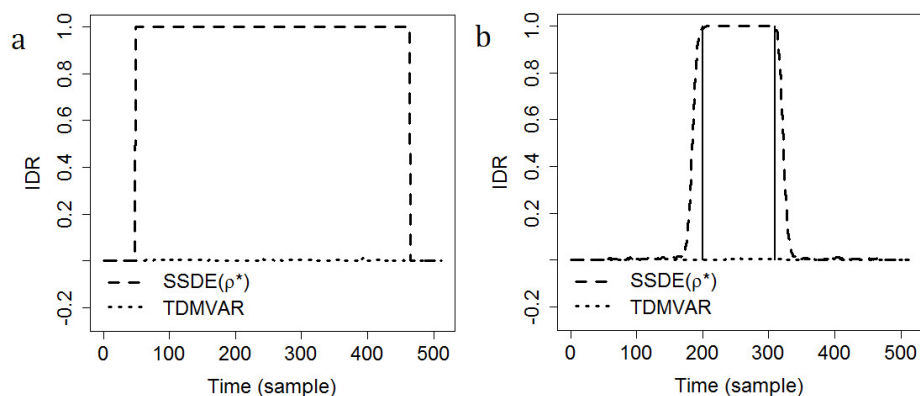
Fig. 6. Curves of IDR vs time calculated for synchrony events with central frequency 15Hz and a synchrony bandwidth of 2Hz. The curves correspond to SSDE and temporal MVAR models. Panel a: Results obtained when the synchrony existed over the complete time course. Panel b: Results achieved when synchrony was limited to a time window. Vertical continuous lines delimit the time window where synchrony was introduced

delay. In particular we use the following clustering rules:

1) Two electrodes in the same cluster must be synchronized with delay magnitude smaller than 7.5 ms.

2) For each electrode in a cluster, there must exist a significant synchrony relation with an electrode in other cluster such that the delay magnitude and sign is the same for all related electrodes in both clusters. In this way, each cluster represents a network of electrodes interconnected with 0-delay relations and connected with another similar network with a common delay. The 0-delay relations may be due to volume conduction, or to some other more complex phenomenon, such as Volume Transmission [13], but the differentiation between these cases is beyond the scope of this work. An example is presented in Fig. 13 for the synchronization in the alpha band in the pre-stimulus, and the complete results in Tables 2 and 3, corresponding to pre-stimulus and post-stimulus, respectively (note that in some cases the clusters involve a single electrode). As explained above, the significance level used for synchrony detection corresponds to an expected false positive ratio of less than one false positive for a time-frequency field of 512×60 pixels and for

all 190 possible electrode pairings.

Table 2. Information of synchrony events detected by SSDE for the real EEG dataset in the pre-stimulus segment

Source nodes	Sink nodes	Time interval (ms)	Frequency interval (Hz)	$|\tau*|$ (ms)
O1,O2,T6	F7,Fp1,Fp2,F8, F3	430-995	9.5-11.5	45
Cz,C4,Pz	T3,F7	430-995	1.5-4	15
Fp1	Cz	430-885	23.5-25	40
T4	Cz	430-950	23.5-26.5	35

Table 3. Information of synchrony events detected by SSDE for the real EEG dataset after the presentation of the stimulus. The time interval reported is relative to the stimulus onset

Source nodes	Sink nodes	Time interval (ms)	Frequency interval (Hz)	$|\tau*|$ (ms)
C3,Cz,P3	T4,T6	5-1125	1.5-4	15
Cz,C4,Pz	T3,F7	5-1125	1.5-4	20
P3	T4	330-1125	5-8	10
P4	T3	265-1125	5-8	10
Pz	Fp2	5-905	2-4	10
F3	T4	5-905	2-4	20

We confirm the causality estimation provided by SSDE for each detected synchrony event, performing a time-frequency PSI analysis, where the causality of a detected synchrony relation (connected region at time-frequency plane) is estimated by means of a voting and the predominant causality in the region is assigned as
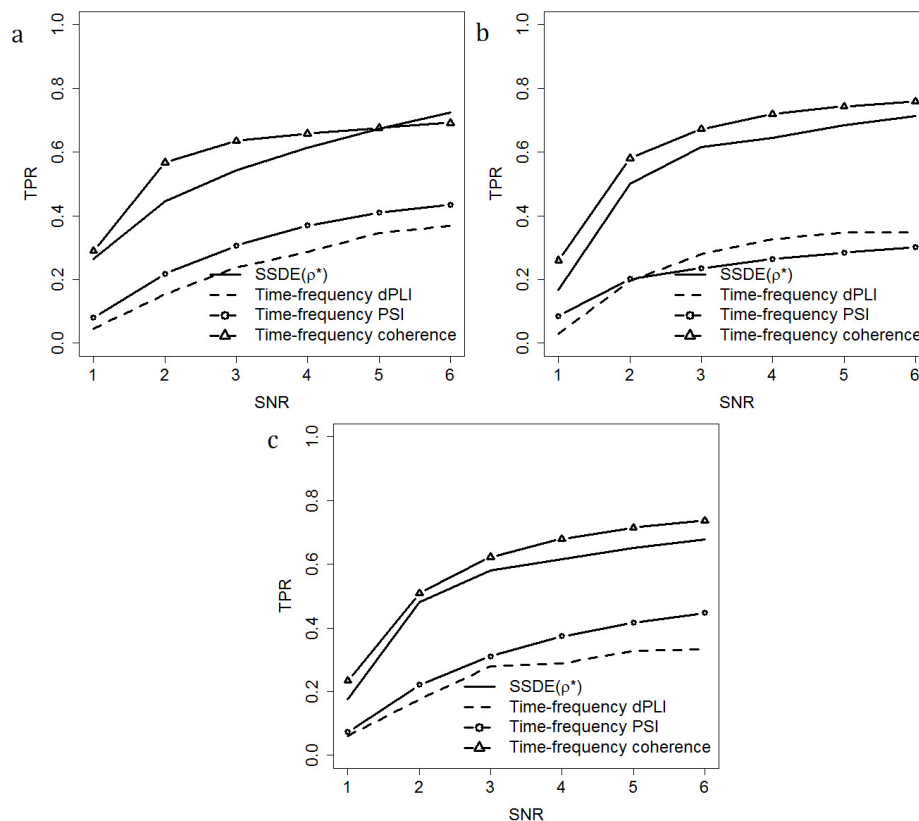
Fig. 7. Synchrony detection performance curves (TPR vs SNR) computed for time-frequency coherence, SSDE, time-frequency PSI and time-frequency dPLI for synchrony events with a bandwidth of 2Hz and different central frequencies. Panel a: Central frequency 10Hz. Panel b: Central frequency 15Hz. Panel c: Central frequency 20Hz

its causality.

These results confirm SSDE causality estimations for every connectivity network and synchrony relation presented in Tables 2 and 3.

We also performed the detection of significant synchrony using time-frequency coherence (Eq. (13)) and found that detected significant regions coincide with those found with our method.

5 Discussion

The proposed SSDE methodology separates the detection phase from the delay estimation phase. This general strategy for causality analysis allows one to use any other synchrony measure, e.g., time-frequency coherence, which may have better performance regarding detection. Points in the time-frequency plane that have a significant degree of synchrony are found in the detection phase. Then, connected regions of synchrony are constructed in the time-frequency plane, so that a synchrony event is represented by a single region.

Delay estimation is based on the analysis of the dispersion of the maxima of the proposed synchrony measure in the space spanned by time, frequency and possible delays, so the delay that gives the least dispersion of these maximums is selected as the delay estimate τ^* corresponding to the complete synchrony region. SSDE is able to separate "apparent" synchrony events caused by volume conduction because of its delay estimation feature, and a compact representation of a synchrony relation, based on groups of electrodes, may be constructed employing
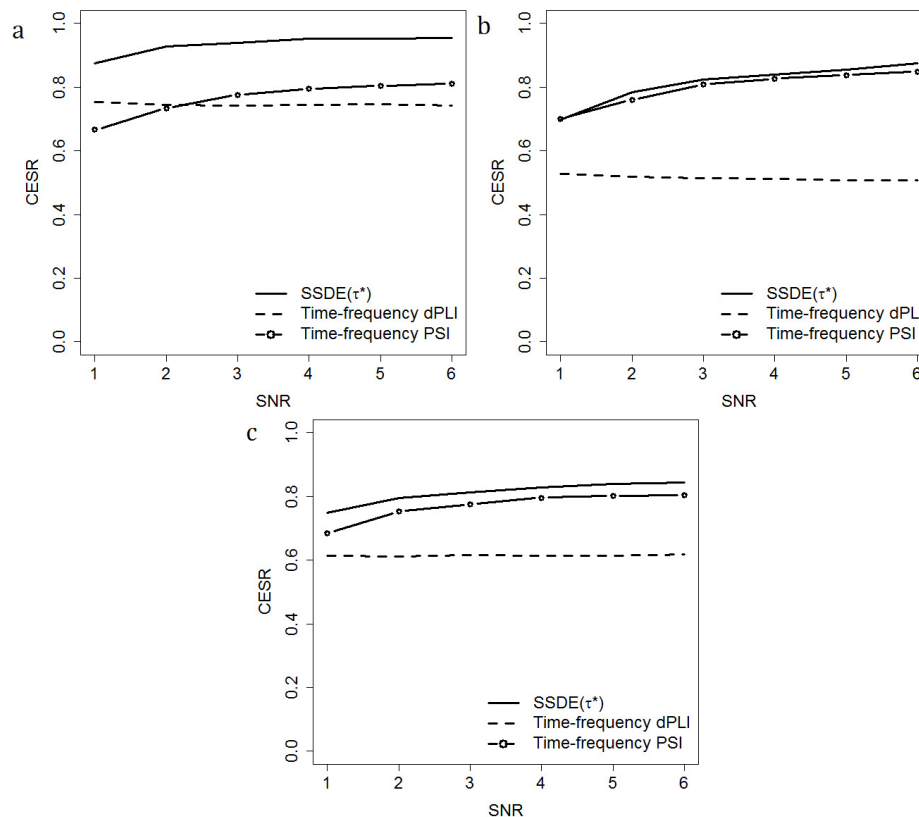
Fig. 8. Causality estimation performance curves (CESR vs SNR) for SSDE, time-frequency PSI and time-frequency dPLI, calculated for synchrony relations with a bandwidth of 2Hz, located at different central frequencies. Panel a: Central frequency 10Hz. Panel b: Central frequency 15Hz. Panel c: Central frequency 20Hz

simultaneously the evidence of the synchrony relation in several electrode pairs and zero-lag synchrony information, that could be cause by volume conduction or other communication mechanisms within the brain.

It is found that the performance of the delay estimation procedure may be improved by employing a resampling strategy over trials. This, however, has a significant impact on the computational cost. For example, in a computer with a processor Intel Xeon with cores at 2.2GHz, running a Linux-based operating system and using a C++ implementation of SSDE without parallelization techniques, for a single electrode pair – 200 trials, $\Gamma = [-15, 15]$, and a time-frequency decomposition of 60 frequency bands and 512 samples per trial – synchrony measure calculation takes about

25 seconds, and the delay estimation step lasts around 2 seconds, when detected regions have an estimated bandwidth between 1 and 2Hz, and the modelled synchrony interval has a length of 110 samples.

If one uses the delay estimation procedure with 100 subsets of 100 trials each -selecting each subset employing bootstrap-, then the delay estimation step takes approximately 43 seconds for each pair, representing almost 22 times the execution time without resampling.

One way for keeping the computation time manageable is to use the resampling strategy only for detected regions with relatively small bandwidth (e.g., less than 4 Hz), since for regions with larger bandwidth the procedure without resampling gives practically the same results.
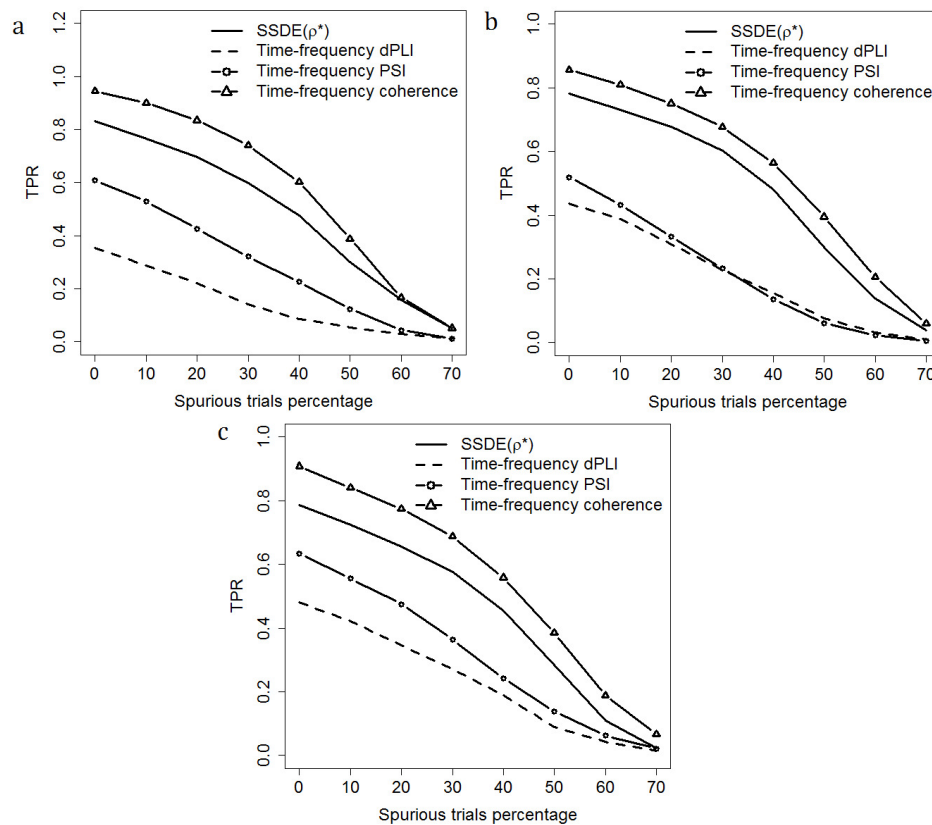
Fig. 9. Synchrony detection performance curves (TPR vs spurious trials percentage) for time-frequency coherence, SSDE, time-frequency PSI and time-frequency dPLI, computed for synchrony events with a 2Hz bandwidth, localized at different central frequencies. Panel a: Central frequency 10Hz. Panel b: Central frequency 15Hz. Panel c: Central frequency 20Hz

Thus, in the case of the real EEG dataset -190 possible electrode pairs, 270 trials and the rest of the parameters set as before- ρ^* calculations take 1 hours and 23 minutes, and the delay estimation stage, employing the resampling strategy for synchrony events with bandwidth lower than 4Hz, lasts 12 hour and 39 minutes.

Note that the use of parallelization techniques will improve the execution time, diminishing significantly the increase in computational cost caused by the resampling strategy.

Delay estimation is more reliable for low frequencies (below 18Hz), although the reliability becomes greater as the synchrony bandwidth increases, so that a synchrony bandwidth of 3Hz is sufficient to analyze synchrony events located at frequencies lower than 25Hz. When the analyzed data contains at least 200 trials, SSDE delay estimations can be trusted and they are more accurate as the number of trials increases, for all frequencies.

Our experiments show that SSDE has better performance than time-frequency PSI and time-frequency dPLI for synchrony detection and causality estimation, and its performance is comparable to that of time-frequency coherence for detecting synchrony events. SSDE has better detection capability than TDMVAR, when synchrony relations are weak ($\beta \geq 0.07$). In general, it seems to be more robust for synchrony detection and causality estimation than TDMVAR even when data is generated with a
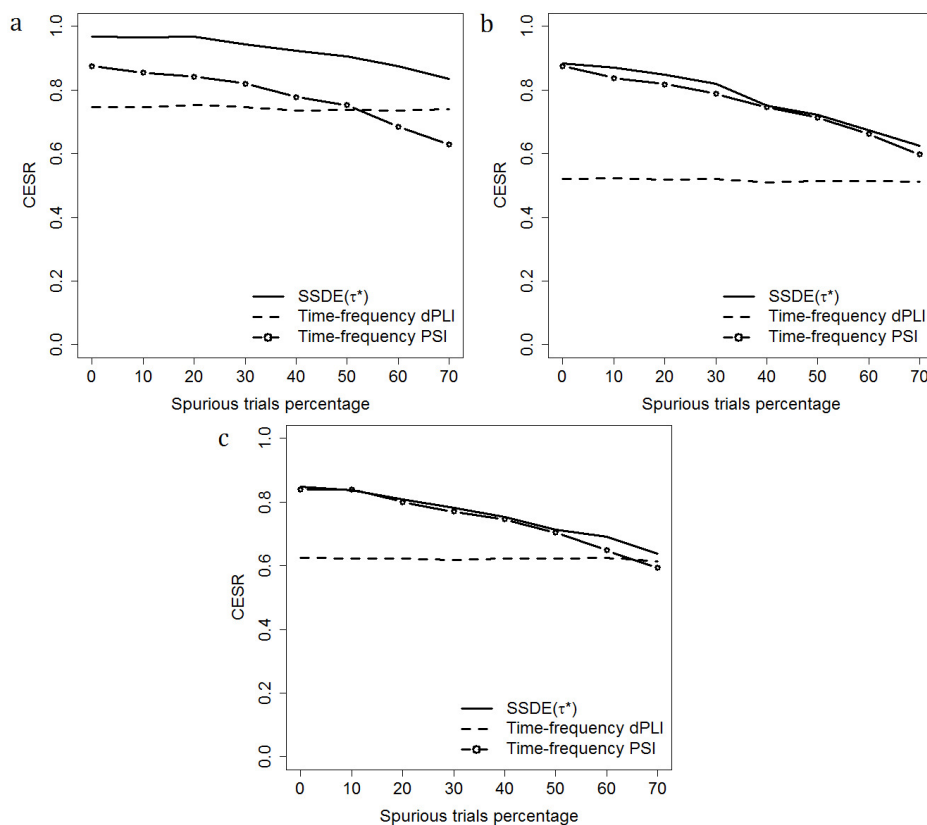
Fig. 10. Causality estimation performance curves (CESR vs spurious trials percentage) for SSDE, time-frequency PSI and time-frequency dPLI, calculated using synchrony events with a bandwidth of 2Hz, localized at several central frequencies. Panel a: Central frequency 10Hz. Panel b: Central frequency 15Hz. Panel c: Central frequency 20Hz
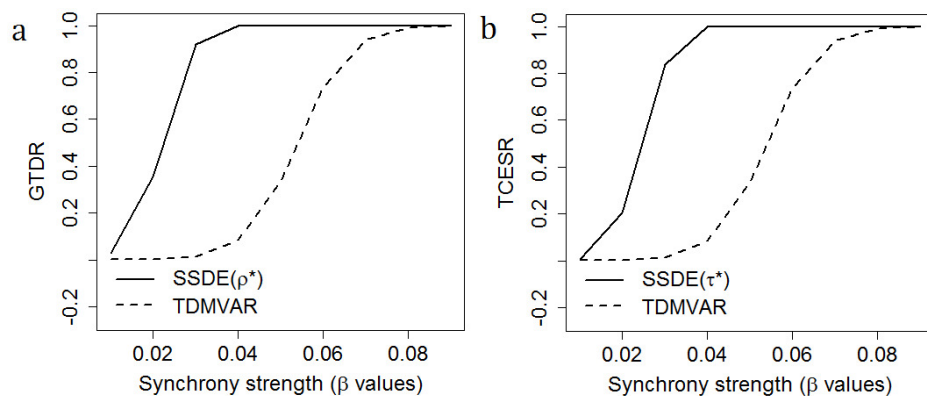
Fig. 11. Results obtained for SSDE and TDMVAR for data generated with an MVAR model, when the synchrony strength varied (parameter β). Panel a: Synchrony detection performance. Panel b: Causality estimation results
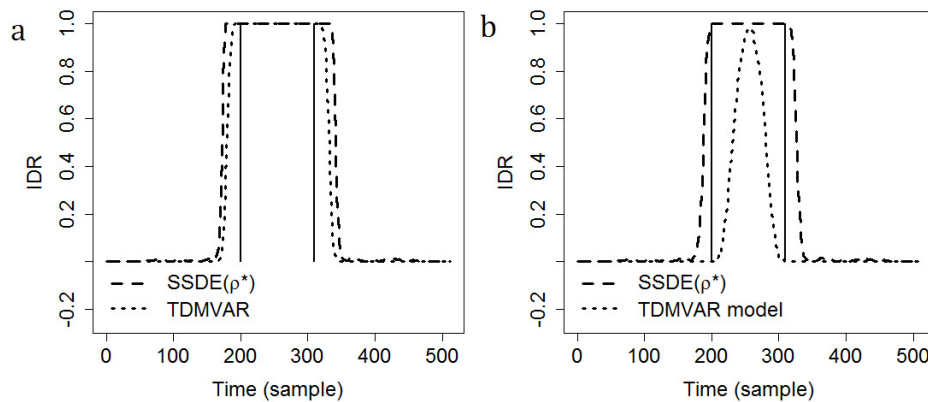
Fig. 12. Synchrony detection results (IDR) obtained for SSDE and TDMVAR over data generated with an MVAR model, when synchrony strength was varied and the synchrony was limited to a time interval. Vertical continuous lines indicate the time interval limits. Panel a: Results for strong synchrony relations ($\beta = 0.2$). Panel b: Performance for weak synchrony relations ($\beta = 0.07$)
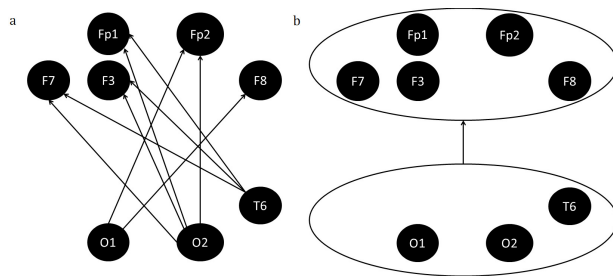
Fig. 13. Results for the synchrony relation detected at the alpha band by SSDE. Panel a: Electrode pairs where the synchrony relation is observed. Arrows direction indicates causality for each pair, which goes from the occipital to the frontal region. Panel b: Representation with electrode clusters, localized in the frontal area (higher cluster) and the occipital region (lower cluster). Arrow direction represents causality

MVAR model, if the data is contaminated with non-synchronized trials.

However, MVAR based methods are able to discriminate direct and indirect synchrony relations, allowing for an accurate analysis of complex connectivity networks, though synchrony with a wide bandwidth is required.

The results in Tables 2 and 3 clearly differentiate the different states between the samples of EEG previous to the stimulus and those after.

From the psychophysiological point of view, previous to the stimulation is generally observed synchronic activity within the alpha band which has been related to a state of rest. This alpha activity has its origin in the occipital regions and propagates to the anterior areas [36, 37].

Several psychophysiological experiments have shown that during the performance of the task (post-stimuli segment) frequencies within the delta (1.5-3.5 Hz) and theta (3.6-7.0 Hz) bands had higher power than the prestimulus segments where frequencies within the alpha band are characteristic. The increases in the slower bands has been related to attention [5, 10, 19, 16, 15].

The complete characterization of narrow-band synchrony events provided by SSDE, particularly the delay magnitude estimation, is unique to this method, and it allows one to build a representation based on electrode clusters that allows us a better interpretation of a synchrony relation.

6 Conclusions

In this work, we introduced a novel methodology (SSDE) for analyzing synchrony and causality in EEG data, based completely on a single synchrony paradigm (phase differences). The SSDE allows us to analyze narrow-band signals, which provides a reliable mechanism, for estimating

quantitatively the associated delay in detected synchrony relations, providing a more complete characterization of them.

However, the separation of detection and delay estimation phases, allows one to use others synchrony measures in the detection step.

Synchrony detection within a time-frequency decomposition provides an accurate time and frequency location for synchrony event, allowing us to estimate the bandwidth associated to a synchrony relation, based on the idea that a synchrony relation is represented by a connected region of the time-frequency decomposition.

The estimation of the delay is better for low frequencies, increasing its reliability when increasing the synchrony bandwidth, although is reliable when synchrony is found in a narrow bandwidth. SSDE proved to be a better alternative than state-of-the-art techniques used here for studying causality. This is based on the observation that SSDE was more robust to the influence of sensor noise and spurious trials in data and its results were better when synchrony was confined within a narrow bandwidth.

A spectral synchrony model was proposed. This model allowed us to control the bandwidth of a synchrony relation. The cluster-based representation introduced here allows us a better comprehension of the electrode network associated to a synchrony event.

Acknowledgements

References

1. **Akaike, H. (1968).** On the use of a linear model for the identification of feedback systems. *Annals of the Institute of Statistical Mathematics*, Vol. 20, No. 1, pp. 425–439.

2. **Alba, A., Marroquin, J. L., Peña, J., Harmony, T., & Gonzalez-Frankenberger, B. (2007).** Exploration of event-induced EEG phase synchronization patterns in cognitive tasks using a time-frequency-topography visualization system. *Journal of Neuroscience Methods*, Vol. 161, No. 1, pp. 166–182.

3. **Baccalá, L. & Sameshima, K. (2001).** Partial directed coherence: a new concept in neural structure determination. *Biological Cibernetics*, Vol. 84, No. 6, pp. 463–474.

4. **Baccalá, L. A., Sameshima, K., & Takahashi, D. Y. (2007).** Generalized Partial Directed Coherence. *2007 15th International Conference on Digital Signal Processing*, pp. 163–166.

5. **Başar, E., Başar-Eroğlu, C., Karakaş, S., & Schürmann, M. (2000).** Brain oscillations in perception and memory. *International Journal of Psychophysiology*, Vol. 35, No. 2-3, pp. 95–124.

6. **Courellis, H., Mullen, T., Poizner, H., Cauwenberghs, G., & Iversen, J. (2017).** EEG-based quantification of cortical current density and dynamic causal network connectivity generalized across subjects performing BCI-monitored cognitive tasks. *Frontiers in Neuroscience*, Vol. 11, pp. 180.

7. **David, O. & Friston, K. J. (2003).** A neural mass model for MEG/EEG: coupling and neural dynamics. *NeuroImage*, Vol. 20, No. 3, pp. 1743–1755.

8. **Dawels, J., Vialatte, F., Musha, T., & Cichocki, A. (2010).** A comparative study of synchrony measures for the early diagnosis of Alzheimer's desease based on EEG. *Neuroimage*, Vol. 49, No. 1, pp. 668–693.

9. **Gardner, W. A. (1992).** A unifying view of coherence in signal processing. *Signal Processing*, Vol. 29, No. 2, pp. 113–140.

10. **Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997).** High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral cortex*, Vol. 7, No. 4, pp. 374–385.

11. **Granger, C. W. J. (1969).** Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, Vol. 37, No. 3, pp. 424–438.

12. **Guerrero, J. A., Marroquin, J. L., Rivera, M., & Quiroga, J. A. (2005).** Adaptive monogenic filtering and normalization of ESPI fringe patterns. *Optics Letters*, Vol. 30, No. 22, pp. 3018–3020.

13. **Guidolin, D., Marcoli, M., Maura, G., & Agnati, L. F. (2016).** New dimensions of connectomics and network plasticity in the central nervous system. *Reviews in Neurosciences*, Vol. 28, No. 2, pp. 113–132.

14. **Harmony, T., Fernández, T., Fernández-Bouzas, A., Silva-Pereyra, J., Bosch, J., Díaz-Comas, L., & Galán, L. (2001).** EEG changes during word and figure categorization. *Clinical Neurophysiology*, Vol. 112, No. 8, pp. 1486–1498.

15. **Harmony, T., Fernández, T., Gersenowies, J., Galán, L., Fernández-Bouzas, A., Aubert, E., & Díaz-Comas, L. (2004).** Specific EEG frequencies signal general common cognitive processes as well as specific task processes in man. *International Journal of Psychophysiology*, Vol. 53, No. 3, pp. 207–216.

16. **Harmony, T., Fernandez, T., Silva, J., Bernal, J., Díaz-Comas, L., Reyes, A., Marosi, E., Rodríguez-Camacho, M., & Rodríguez, M. (1996).** EEG delta activity: an indicator of attention to internal processing during performance of mental tasks. *International Journal of Psychophysiology*, Vol. 24, No. 1-2, pp. 161–171.

17. **Hebert, R., Lehmann, D., Tan, G., Travis, F., & Arenander, A. (2005).** Enhanced EEG alpha time-domain phase synchrony during transcendental meditation: implications for cortical integration theory. *Signal Processing*, Vol. 85, No. 11, pp. 2213–2232.

18. **Kamiński, M. & Blinowska, K. (1991).** A new method of the description of the information flow in the brain structures. *Biological Cibernetics*, Vol. 65, No. 3, pp. 203–210.

19. **Klimesch, W. (1999).** EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews*, Vol. 29, No. 2-3, pp. 169–195.

20. **Korzeniewska, A., Mańczak, M., Kamiński, M., Blinowska, K., & Kasicki, S. (2003).** Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method. *Journal of Neuroscience Methods*, Vol. 125, No. 1-2, pp. 195–207.

21. **Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999).** Measuring phase synchrony in brain signals. *Human Brain Mapping*, Vol. 8, No. 4, pp. 194–208.

22. **Lopez da Silva, F., Pijn, J. P., & Boeijinga, P. (1989).** Interdependence of EEG signals: lineal vs nonlinear associations and the significance of time delays and phase shifts. *Brain Topography*, Vol. 2, No. 1, pp. 9–18.

23. **Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., & Hallett, M. (2004).** Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical Neurophysiology*, Vol. 115, No. 10, pp. 2292–2307.

24. **Nolte, G., Ziehe, A., Nokulin, V. V., Schlögl, A., Krämer, N., Brismar, T., & Müller, K. (2008).** Robustly estimating the flow direction of information in complex physical systems. *Physical Review Letters*, Vol. 100, No. 23, pp. 234101.

25. **Numan, T., Slooter, A. J. C., van der Kooi, A. W., Hoekman, A. M. L., Suyker, W. J. L., Stam, C. J., & van Dellen, E. (2017).** Functional connectivity and network analysis during hypoactive delirium and recovery from anesthesia. *Clinical Neurophysiology*, Vol. 128, No. 6, pp. 914–924.

26. **Pascual-Marqui, R. D. (2007).** Coherence and phase synchronization: generalization to pairs of multivariate time series, and removal of zero-lag contributions.

27. **Pascual-Marqui, R. D. (2007).** Instantaneous and lagged measurements of linear and nonlinear dependence between groups of multivariate time series: frequency decomposition.

28. **Pascual-Marqui, R. D., Biscay, R. J., Bosch-Bayard, J., Lehmann, D., Kochi, K., Yamada, N., Kinoshita, T., & Sadato, N. (2014).** Isolated effective coherence (iCoh): causal information flow excluding indirect paths.

29. **Pascual-Marqui, R. D., Biscay, R. J., Valdes-Sosa, P. A., Bosch-Bayard, J., & Riera-Diaz, J. J. (2011).** Cortical current source connectivity by means of partial coherence fields.

30. **Pfurtscheller, G. (1972).** *Invited Discussion: Some results of the analysis of epileptic seizure patterns by correlation-methods*, chapter 18. Springer Vienna, Vienna, pp. 285–290.

31. **Sakkalis, V. (2011).** Review of advance techniques for the estimation of brain connectivity measured with EEG/MEG. *Computers in Biology and Medicine*, Vol. 41, No. 12, pp. 1110–1117.

32. **Stam, C. J., Nolte, G., & Daffertshofer, A. (2007).** Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Human Brain Mapping*, Vol. 28, No. 11, pp. 1178–1193.

33. **Stam, C. J. & van Straaten, E. C. W. (2012).** Go with the flow: Use of a directed phase lag index (dPLI) to characterize patterns of phase relations in a large-scale model of brain dynamics. *Neuroimage*, Vol. 62, No. 3, pp. 1415–1428.

34. **Thatcher, R. W., Palmero-Soler, E., North, D. M., & Biver, C. J. (2016).** Intelligence and EEG measures of information flow: efficiency and homeostatic neuroplasticity. *Scientific Reports*, Vol. 6, pp. 38890.

35. **Thelier, J., Eubank, S., Longtin, A., Galdrikian, B., & Farmer, J. D. (1992).** Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, Vol. 58, No. 1, pp. 77–94.

36. **Valdés, P., Bosch, J., Grave, R., Hernandez, J., Riera, J., Pascual, R., & Biscay, R. (1992).** Frequency domain models of the EEG. *Brain Topography*, Vol. 4, No. 4, pp. 309–319.

37. **Valdes-Sosa, P. A., Sanchez-Bornot, J. M., Sotero, R. C., Iturria-Medina, Y., Aleman-Gomez, Y., Bosch-Bayard, J., Carbonell, F., & Ozaki, T. (2009).** Model driven EEG/fMRI fusion of brain oscillations. *Human Brain Mapping*, Vol. 30, No. 9, pp. 2701–2721.

38. **Vinck, M., Oostenveld, R., van Wingerden, M., Battaglia, F., & Pennartz, C. M. (2011).** An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *Neuroimage*, Vol. 55, No. 4, pp. 1548–1565.

39. **Wang, H. E., Bénar, C. G., Quilichini, P. P., Friston, K. J., Jirsa, V. K., & Bernard, C. (2014).** A systematic framework for functional connectivity. *Frontiers in Neuroscience*, Vol. 8, pp. 405.