OPEN ACCESS
IOP Publishing
Journal of Neural Engineering
J. Neural Eng. **15** (2018) 056011 (11pp)
https://doi.org/10.1088/1741-2552/aacfe4

Phase locking value revisited: teaching new tricks to an old dog

Ricardo Bruña[1,2,4], **Fernando Maestú**[1,2] and **Ernesto Pereda**[1,3]

[1] Laboratory for Cognitive and Computational Neuroscience, Canter for Biomedical Technology, Technical University of Madrid, Pozuelo de Alarcón, Madrid, Spain
[2] Departamento de Psicologia Experimental, Procesos Psicologicos y Logopedia, Faculty of Psychology, Complutense University of Madrid, Pozuelo de Alarcón, Madrid, Spain
[3] Electrical Engineering and Bioengineering Group, Departamento de Ingeniería Industrial e Instituto Universitario de Neurociencia (IUNE), University of La Laguna, San Cristobal de La Laguna, Tenerife, Spain

E-mail: ricardo.bruna@ctb.upm.es

Received 31 January 2018, revised 26 June 2018
Accepted for publication 28 June 2018
Published 27 July 2018

Abstract

Objective. Despite the increase in calculation power over the last few decades, the estimation of brain connectivity is still a tedious task. The high computational cost of the algorithms escalates with the square of the number of signals evaluated, usually in the range of thousands. In this work we propose a re-formulation of a widely used algorithm that allows the estimation of whole brain connectivity in much smaller times. *Approach*. We start from the original implementation of phase locking value (PLV) and re-formulated it in a computationally very efficient way. What is more, this formulation stresses its strong similarity with coherence, which we used to introduce two new metrics insensitive to zero lag synchronization: the imaginary part of PLV (iPLV) and its corrected counterpart (ciPLV). *Main results*. The new implementation of PLV avoids some highly CPU-expensive operations and achieves a 100-fold speedup over the original algorithm. The new derived metrics were highly robust in the presence of volume conduction. Moreover, ciPLV proved capable of ignoring zero-lag connectivity, while correctly estimating nonzero-lag connectivity. *Significance*. Our implementation of PLV makes it possible to calculate whole-brain connectivity in much shorter times. The results of the simulations using ciPLV suggest that this metric is ideal to measure synchronization in the presence of volume conduction or source leakage effects.

Keywords: synchronization, neuroscience, computational efficiency, brain connectivity

Supplementary material for this article is available online

(Some figures may appear in colour only in the online journal)

1. Introduction

The study of the brain caught the interest of researchers before neuroscience was a discipline of its own. Following the then-called phrenological doctrine, for a long time the brain was considered an organ, or rather a set of organs, with distinct parts taking care of distinct duties. However, in recent years some researchers have begun to departure from this traditional approach. Instead, they argue than cognition, thought and action are supported by the collective work of sets of brain areas (networks) (Friston 1994). In this paradigm, the brain areas form a network (the *connectome*), entwined by white matter tracts (anatomical connectome), in which different sub-networks communicate dynamically with each other to perform different functions (functional connectome).

[4] Author to whom any correspondence should be addressed.

Whereas the anatomical connectome is assessed from diffusion tensor imaging (DTI) (Le Bihan *et al* 2001), the functional connectome is constructed from neuroimaging techniques such as functional magnetic resonance imaging (fMRI) and electro- and magneto-physiology (Friston *et al* 2003, Brookes *et al* 2011). In the latter case, however, the description of the connectome is not straightforward, as the underlying mechanisms of the brain are unknown. Instead, we must trust on measures of brain activity at different recording sites. In this framework, functional connectivity (FC) is defined as the existence of a statistical dependence between the activities in two or more sites above chance level, a dependence that can be evaluated in different ways.

In the case of signals such as M/EEG, one of the most studied connectivity hypothesis is that of phase synchronization (PS) (Lachaux *et al* 1999, López *et al* 2014, Fries 2015, Garcés *et al* 2016). Under this scenario, two brain areas show the activity of different oscillators and, in the case of connectivity between different regions, their oscillation properties (i.e. phase or frequency) should be related. If this relation can be mathematically evaluated, we get an estimator of phase connectivity.

Of the many PS measures available in the literature, one of the most used, mainly due to its simplicity, is the Phase Locking Value (PLV, (Lachaux *et al* 1999)), sometimes called mean phase coherence (MPC) (Mormann *et al* 2000). This measure evaluates the instantaneous phase difference of the signals under the hypothesis that connected areas generate signals whose instantaneous phases evolve together. In this case, the phases of the signals are said to be 'locked', and their difference is therefore constant. However, real-world signals are inherently noisy and it is not always possible to be sure that the evaluated signal only comes from one oscillator.

The problem is solved by allowing some deviation from the condition of a constant phase difference. Thus, PLV evaluates the spread of the distribution of phase differences, and the connectivity estimation is linked to this spread. The narrower the distribution of the phase difference, the higher the PLV value, which ranges between zero (no phase dependence) and one (complete phase dependence). We will elaborate on the mathematics behind the estimation of the PLV in the next section. Here we will provide a brief description the main steps involved.

The most common method to calculate PLV is based on the instantaneous phase of the signals obtained using the Hilbert or the wavelet transforms. In both cases, the calculations are fast and reliable, which has undoubtedly contributed to its use. The calculation of the phase difference and its spread is also mathematically simple, and, in theory, not very time-consuming. However, even when the computation of PLV is fast and easy for a fair amount of data, the computational cost grows with the square of the number of signals. This is especially important in the case of M/EEG, distributed source-space analysis, where such number ranges from the thousands to the hundreds of thousands.

It is also noteworthy that despite the estimation of the PLV and other related indices from data has a long history, from time to time new studies appear (Ewald *et al* 2012, Kovach 2017) in which different aspects of the calculation procedures, and the indices themselves, are analyzed with a fresh eye to refine the existing methodologies. This paper intends to be one of such studies and to shed new light over PS estimation, interpretation and use. Thus, we start from the original PLV formulation (Lachaux *et al* 1999) and re-write it to obtain an equivalent expression that is much easier to compute, reducing the time required for the PLV calculation up to a factor of 100. Furthermore, we also show that this new formulation is closely related to that of coherency, thereby allowing an interpretation of this well-studied function in terms of the PLV (see also Kovach (2017)). In addition, such interpretation allows for the formulation of two PLV-based zero-lag-insensitive measures: the imaginary PLV (iPLV) and the corrected iPLV (ciPLV), both of which can be used as alternatives to the imaginary part of the coherency (Nolte *et al* 2004) and its corrected version (Ewald *et al* 2012), in the assessment of direct PS from M/EEG data.

2. Methods

2.1. Computational optimization

In their original paper, Lachaux *et al* (1999) defined the PLV as a time-dependent connectivity measured tailored to study evoked activity. The idea behind their definition is that the stimulus resets the phase of the neural oscillators so that signals connected in a given time should have a stable phase-difference along trials. Its mathematical formulation reads:

$$PLV_{i,j}(t) = \frac{1}{N}\left|\sum_{n=1}^{N} e^{-i(\varphi_i(t,n)-\varphi_j(t,n))}\right| \tag{1}$$

where N is the number of trials and $\varphi_i(t,n)$ is the instantaneous phase for signal i in trial n at time t.

This definition can be extended to resting-state data, by assessing phase locking as a stable phase-difference over time, thereby obtain the so-called MPC[5] (Mormann *et al* 2000):

$$PLV_{i,j} = \frac{1}{T}\left|\sum_{t=1}^{T} e^{-i(\varphi_i(t)-\varphi_j(t))}\right| \tag{2}$$

where T is the data length.

In either case, one must extract the instantaneous phase $\varphi(t)$ of each signal. Moreover, for the phase to be physically meaningful, it is necessary that only one oscillator is present in each signal. This is achieved, e.g. by means of a narrow-band pass filtering or, equivalently, the convolution with a narrow band complex wavelet such as that of Morlet (Bruns 2004).

After the filtering process, we obtain a band-pass version of the Hilbert analytical signal:

$$cBP\{x(t)\} = x_{BP,H}(t) = x_{BP}(t) + i\tilde{x}_{BP}(t) = A(t)\cdot e^{-i\varphi_i(t)} \tag{3}$$

where $\tilde{x}$ represent the Hilbert transform of x, and BP stands for band pass. The instantaneous phase is the angle between

[5] Henceforth, for the sake of simplicity, we will use $PLV_{i,j}$ to refer to both the time varying phase locking value as derived by Lachaux *et al*, and the mean phase coherence. In any case, both indices can be easily distinguished since the former one ($PLV_{i,j}(t)$) explicitly depends on time, whereas the latter one, averaged over the whole data segment, does not.

the real and the imaginary parts of the Hilbert analytical signal, or the angle between the original (band-pass) signal and its Hilbert transform.

The instantaneous phase is usually extracted from this analytical signal, the phase difference estimated and, finally, the exponentiation calculated to get the unit phase difference vector. However, these two operations (phase extraction and exponentiation) are computationally expensive, but, as we will show, they can be easily circumvented by using the properties of exponentials.

First, let us obtain the oscillatory part of the analytical signal by normalizing (3):

$$\dot{x}_{BP,H,i}(t) = \frac{x_{BP,H,i}(t)}{|x_{BP,H,i}(t)|} = \frac{A_i(t)\cdot e^{-i\varphi_i(t)}}{A_i(t)} = e^{-i\varphi_i(t)}. \tag{4}$$

From this, we easily derive the exponential of the phase difference:

$$\begin{aligned}\dot{x}_{BP,H,i}(t)\cdot(\dot{x}_{BP,H,j}(t))^* &= e^{-i\varphi_i(t)}\cdot\left(e^{-i\varphi_j(t)}\right)^* \\ &= e^{-i\varphi_i(t)}\cdot e^{i\varphi_j(t)} = e^{-i(\varphi_i(t)-\varphi_j(t))}\end{aligned} \tag{5}$$

where $(\cdot)^*$ represents complex conjugate. Thus, expression (2) can be rewritten as:

$$PLV_{i,j} = \frac{1}{T}\left|\sum_{t=1}^{T}\dot{x}_{BP,H,i}(t)\cdot(\dot{x}_{BP,H,j}(t))^*\right| \tag{6}$$

or, using vector algebra:

$$PLV_{i,j} = \frac{1}{T}\left|\dot{x}_i\cdot\dot{x}_j^T\right| \tag{7}$$

where $\dot{x}_i$ is a vector version of $\dot{x}_{BP,H,i}(t)$ and $\dot{x}^T$ the transpose conjugate of $\dot{x}$. This calculation is computationally very efficient and allows a considerable speed up of the estimation, with low memory penalization. The calculation efficiency is discussed in the *results* section.

The efficient formulation can be extended to Lachaux's $PLV_{i,j}(t)$ by constructing the vectors with the tth sample of each trial:

$$PLV_{i,j}(t) = \frac{1}{T}\left|\dot{x}_{i,t}\cdot\dot{x}_{j,t}^T\right|. \tag{8}$$

2.2. Relation to coherence

If we expand (1) using the oscillatory part of x, we get:

$$PLV_{i,j}(t) = \frac{1}{N}\left|\sum_{n=1}^{N}e^{-i\varphi_i(t,n)}\cdot\left(e^{-i\varphi_j(t,n)}\right)^*\right|. \tag{9}$$

Phase synchronization only makes sense for signals composed of a single oscillatory component. Yet, it is mathematically possible to calculate the PLV from both a broadband signal and an arbitrarily narrow band one, even when these calculations do not have, in principle, physical sense.

If we take the extreme case of a single oscillator whose spectrum is non-zero only at a given frequency, f_o, the signal can be written as an out-of-phase cosine, where the phase at the initial time is equal to the Fourier phase, and the phase at any other time is determined by the delay and the frequency.

$$\begin{aligned}x_f(t) = A\cdot\cos(2\pi f_o t+\varphi) = A\cdot\Re\left\{e^{-i(2\pi f_o t+\varphi)}\right\} \\ X_f(f_o) = \tfrac{A}{2}\cdot e^{-i\varphi}\end{aligned} \tag{10}$$

where $X(f)$ is the Fourier transform of $x(t)$ at f_o and $\Re\{x\}$ stands for the real part of x. In this case, Lachaux's definition of PLV is no longer time dependent, and its formulation can be simplified as:

$$\begin{aligned}PLV_{i,j} &= \frac{1}{N}\left|\sum_{n=1}^{N}e^{-i\varphi_i(n)}\cdot\left(e^{-i\varphi_j(n)}\right)^*\right| \\ &= \frac{1}{N}\left|\sum_{n=1}^{N}\frac{X_i(f,n)\cdot(X_j(f,n))^*}{|X_i(f,n)|\cdot|X_j(f,n)|}\right|.\end{aligned} \tag{11}$$

This formula closely resembles that of coherence (Nunez *et al* 1997). In fact, coherence can be rewritten as:

$$\begin{aligned}COH_{i,j}(f) &= \frac{\left|\sum_{n=1}^{N}X_i(f,n)\cdot(X_j(f,n))^*\right|}{\sqrt{\sum_{n=1}^{N}|X_i(f,n)|^2\cdot\sum_{n=1}^{N}|X_j(f,n)|^2}} \\ &= \frac{\left|\sum_{n=1}^{N}|X_i(f,n)|\cdot|X_j(f,n)|\cdot\frac{X_i(f,n)\cdot(X_j(f,n))^*}{|X_i(f,n)|\cdot|X_j(f,n)|}\right|}{\sqrt{\sum_{n=1}^{N}|X_i(f,n)|^2\cdot\sum_{n=1}^{N}|X_j(f,n)|^2}}.\end{aligned} \tag{12}$$

With this formulation, it becomes clear that coherence is a weighted average of the unit phase vectors, i.e. a version of PLV weighted by the joint amplitude of the signals at a given frequency. Alternatively, coherence can be a regarded as a version of PLV weighted by the signal-to-noise (SNR) ratio of each trial, because if the environmental conditions (noise) are stable, the amplitude of a signal is proportional to its SNR. However, this is only right if the level of noise is stable, as a sudden increase in the noise level could lead to higher overall amplitudes. In any case, coherence and PLV are tightly related and both are different formulations of the same principle.

It can be argued that PLV and coherence are inherently different, as the PLV must be calculated over a whole oscillator (i.e. a frequency band) and coherence is calculated independently for each frequency. However, following the previous logic, it is trivial to understand that PLV in a given band is related to the average coherence in the frequency range encompassed by the band weighted by their relative amplitude. Following this logic, it is possible to define a temporal version of coherence, where the phases in comparison are not frequency-related (i.e. Fourier or wavelet) phases, but temporal (i.e. Hilbert) phases:

$$hCOH_{i,j}(t) = \frac{\left|\sum_{t=1}^{T}x_{BP,H,i}(t)\cdot(x_{BP,H,j}(t))^*\right|}{\sqrt{\sum_{t=1}^{T}|x_{BP,H,i}(t)|^2\cdot\sum_{t=1}^{T}|x_{BP,H,j}(t)|^2}}. \tag{13}$$

We term this version of coherence the Hilbert coherence (hCOH). The only difference between hCOH and PLV is the weighting, as in PLV all the time samples has the same influence in the output, while in hCOH the influence of a given temporal sample is weighted by its amplitude.

We will elaborate further on the relationship between PLV and coherence in the *Results* section.

2.3. Zero-lag-insensitivity versions

Despite its popularity, PLV presents an important limitation when used to assess functional brain connectivity: its sensitivity to volume conduction (Stam *et al* 2007) and source-leakage effects. Indeed, in the case of M/EEG data at the sensor level, several sensors can simultaneously pick up the activity from the same source (volume conduction). In turn, at the source level and due to the low spatial resolution of the data, different neighboring sources may share some activity (source leakage).

Fortunately, due to the low capacitance of the tissues of the head for the physiological frequencies and the small distance that the currents have to travel, the propagation of the signals of interest can be considered instantaneous (Nolte *et al* 2004, Stam *et al* 2007). Under this assumption, volume conduction/source leakage occurs with zero-lag propagation. In other words, the phase difference of the part of the signals related to such spurious connectivity must be zero.

Following this logic, Stam *et al* (2007) and (Nolte *et al* (2004) came with two different PS metrics that discard zero-lag connectivity and therefore are insensitive to volume conduction. They are the phase lag index (PLI) and the imaginary part of coherency, respectively. Due to the tight relation of PLV with coherency, it is possible to extend the imaginary part of coherency to PLV to obtain a PLV-based measure insensitive to volume conduction effects, which, for symmetry, we will term imaginary PLV (iPLV):

$$iPLV_{i,j,t} = \frac{1}{T}\Im\left\{\dot{x}_{i,t}\cdot\dot{x}_{j,t}^{T}\right\} \tag{14}$$

where $\Im\{x\}$ stands for the imaginary part of x. This measure is insensitive to zero-lag effects, as it removes the contribution of the zero phase differences that due to the complex exponentiation gives real PLV values. However, (14) is not normalized, as its upper bound, corresponding to two signals with a phase difference $\varphi > 0$, is $\sin(\varphi)$. This can be corrected analogously to what (Ewald *et al* 2012) did for the imaginary part of coherency, to define a corrected imaginary PLV (ciPLV):

$$ciPLV_{i,j,t} = \frac{\frac{1}{T}\Im\left\{\dot{x}_{i,t}\cdot\dot{x}_{j,t}^{T}\right\}}{\sqrt{1-\left(\frac{1}{T}\Re\left\{\dot{x}_{i,t}\cdot\dot{x}_{j,t}^{T}\right\}\right)^{2}}}. \tag{15}$$

This definition of ciPLV is similar to the definition of lagged coherence as introduced by Pascual-Marqui and coworkers (Pascual-Marqui *et al* 2011).

3. Materials

3.1. Real data

In studying the behavior of the proposed formulation of the PLV algorithm, we calculated PLV using three different implementations for different data lengths and number of signals. For the sake of fidelity, we used real source-space data, even when the computation is completely deterministic and any data set with the same characteristics would produce the same results.

Test data consisted of five minutes of eyes-closed resting-state magnetoencephalographic (MEG) activity acquired using a 306-sensors Elekta Vectorview system (Elekta, AB, Stockholm, Sweden), located inside a magnetically shielded room (VacuumSchmelze GmbH, Hanau, Germany) at the Laboratory for Cognitive and Computational Neuroscience (Madrid, Spain). Data was acquired with a sampling rate of 1000 Hz (anti-alias band-pass filter of 0.1–330 Hz) and filtered using a spatiotemporal signal space separation method (Taulu and Simola 2006).

Data were segmented in 20 or 40 four-second artifact-free segments and band-pass filtered in the classical alpha band (8 to 12 Hz) using a 2000th order FIR filter in two passes with 2 s of real data as padding on both sides. The instantaneous phase of the signal was determined using Hilbert's analytical signal. To avoid edge effects, the analytical signal was calculated prior to the removal of the padding. Additionally, to get usable timescales for the original implementations, data was downsampled by a factor of ten.

We then placed 2459 dipoles inside the head of the subject, in a 1 cm homogeneous 3D grid. Source-space data was calculated using a realistic single shell as a forward model (Nolte 2003) and a beamformer as inverse method (van Veen *et al* 1997). The obtained spatial filter was applied to the sensor-space data to obtain up to 2459 time series in source-space. As no orientation constrain could be applied, the 3D source reconstruction was projected over the maximum power direction, obtaining a unique time series per source. Finally, we calculated the all-to-all PLV connectivity matrix for different number of sources, ranging between 500 and 2459. We calculated PLV using Matlab (The Mathworks Inc., Natick, Massachusetts) and the three functions whose codes are included in the appendix. Time and memory performances were measured using Matlab's tic/toc and FieldTrip's (Oostenveld *et al* 2011) memtic/memtoc functions, respectively.

To correctly evaluate the ability of PLV (and its derived metrics) and coherence to detect real connectivity, we used a set of non-linear coupled oscillators (described in the next subsection). However, as it is interesting to observe the differences between these metrics when working with real signals, the supplementary material (stacks.iop.org/JNE/15/056011/mmedia) shows the connectivity matrices obtained with PLV, coherence and their derived metrics for the MEG data described here.

3.2. Simulated data

To evaluate the behavior of coherence and PLV in signals with a known degree of coupling, we used a pair of coupled chaotic systems: a Rössler system and a Lorenz one (Quian-Quiroga *et al* 2000). In this setup, the Rössler system acts as driver and an oscillation frequency can be defined from it. The slave

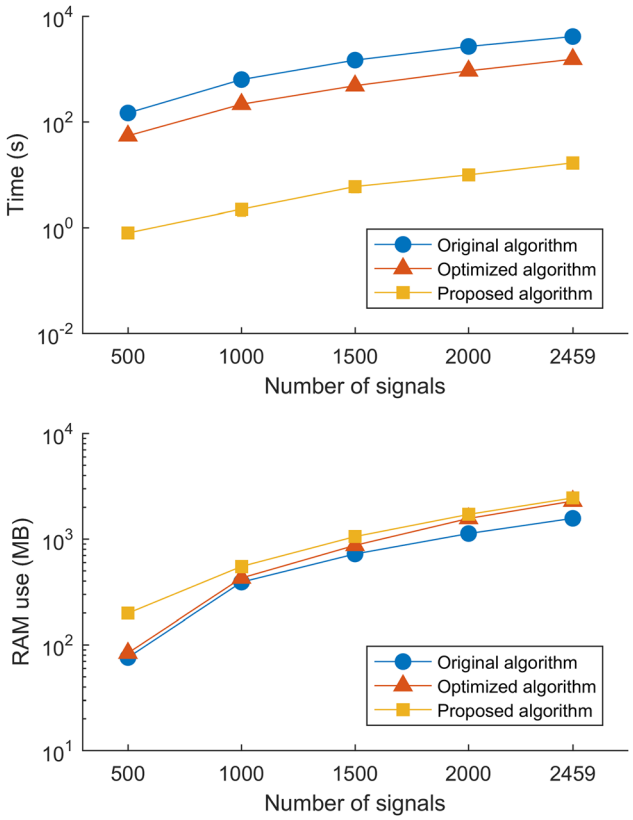

Figure 1. Execution time (up) and RAM use (down) of the three evaluated PLV algorithms for a different number of signals and 40 trials. The plotted values for each data point are mean and standard deviation of the execution time and RAM use over five executions. The proposed algorithm achieves a 100-fold speed-up with a small increase in memory use. Please note the logarithmic scale of the *y*-axis.

Lorenz system is driven by the Rössler to an extent determined by the coupling parameter C, ranging from zero (completely independent systems) to one. The mathematical definition of the coupled systems is:

$$\begin{cases} \dot{x}_1 = -a \cdot (y_1 + z_1) \\ \dot{y}_1 = a \cdot (x_1 + 0.2 \cdot y_1) \\ \dot{z}_1 = a \cdot (0.2 + z_1 \cdot x_1 - 5.7 \cdot z_1) \end{cases} \quad (16)$$
$$\begin{cases} \dot{x}_2 = 10 \cdot (y_2 - x_2) \\ \dot{y}_2 = 28 \cdot x_2 - y_2 - x_2 * z_2 + C \cdot y_1^2 \\ \dot{z}_2 = x_2 \cdot y_2 - {}^{8}\!/_{3} \cdot z_2 \end{cases}$$

where subscripts 1 and 2 refer to the Rössler and the Lorentz system, respectively, a is a scaling parameter determining the fundamental frequency of the Rössler oscillator and C is the coupling parameter. The scaling parameter was set to 10, establishing the oscillatory frequency at $0.585 \cdot \pi$ radians per sample, and the oscillatory band was set to $0.570 \cdot \pi$ to $0.600 \cdot \pi$ radians per sample. With this setup, we generated 50 pairs of signals of 20 000 samples for different values of C ranging between zero and one. PS between the systems was finally estimated using the first variable of each chaotic system x.

4. Results

4.1. Speedup obtained using the proposed algorithm

Firstly, we estimated the CPU time and RAM consumption for the calculation of PLV using the three evaluated algorithms. Figure 1 shows the results of the performance test using 40 trials and different numbers of signals. Every configuration was evaluated five times, and the figure shows the mean and the standard deviation. The original implementation is the slowest, but the most memory-efficient of the three evaluated algorithms. The optimized implementation achieves almost a 3-fold speedup with a slight increase in memory consumption. However, the proposed algorithm allows for a 100-fold speedup over the optimized implementation, with only a marginal (especially at a high number of signals) increase in memory use. Results using 20 trials of data showed a similar result, and are detailed in the supplementary material, along with a more detailed description of the speeding-up achieved by the proposed algorithm under different conditions. As the calculations with 20 trials takes half the time, it was possible to increase the number of repetitions from 5 to 50, giving as a result a more robust estimation of the CPU time and memory consumption.

These results show that the proposed algorithm can achieve a dramatic speed-up in the calculation of the PLV values. The algorithm even allows the calculation of PLV from non-down-sampled data, calculating the full connectivity matrix of 2459 time-series in an average of 113 s, using an average of around 8 GB of RAM memory, over 100 executions. Extrapolating the speeding-up results shown in the previous paragraph to this case, even with the optimized implementation this execution would take 3 h.

4.2. Comparison of PLV and coherence

As noted in the *methods* section, both PLV and coherence are phase-synchronization estimation algorithms. Both metrics measure similar properties of the data, but while coherence gives one result per frequency, PLV gives one result per frequency band. This difference arises from the fact that coherence estimates phase synchronization from Fourier's phase, whereas PLV estimates it from Hilbert's phase. In the extreme case of an infinitely narrow band, both phase definitions converge, but in a normal case where the band is finite, the results of coherence and PLV would diverge.

Figure 2 shows the value of synchronization estimated by PLV and coherence for different values of coupling. In an ideal scenario where the evolution of the synchronization with the coupling was linear, the expected graph would be a straight line from a value of zero when there is no coupling to a value of when there is perfect phase synchronization. However, the behavior of PLV is non-linear with the coupling. In general, the ideal case would be a monotonously increasing line, taking values close to zero for low values of coupling and close to one for high values of coupling.

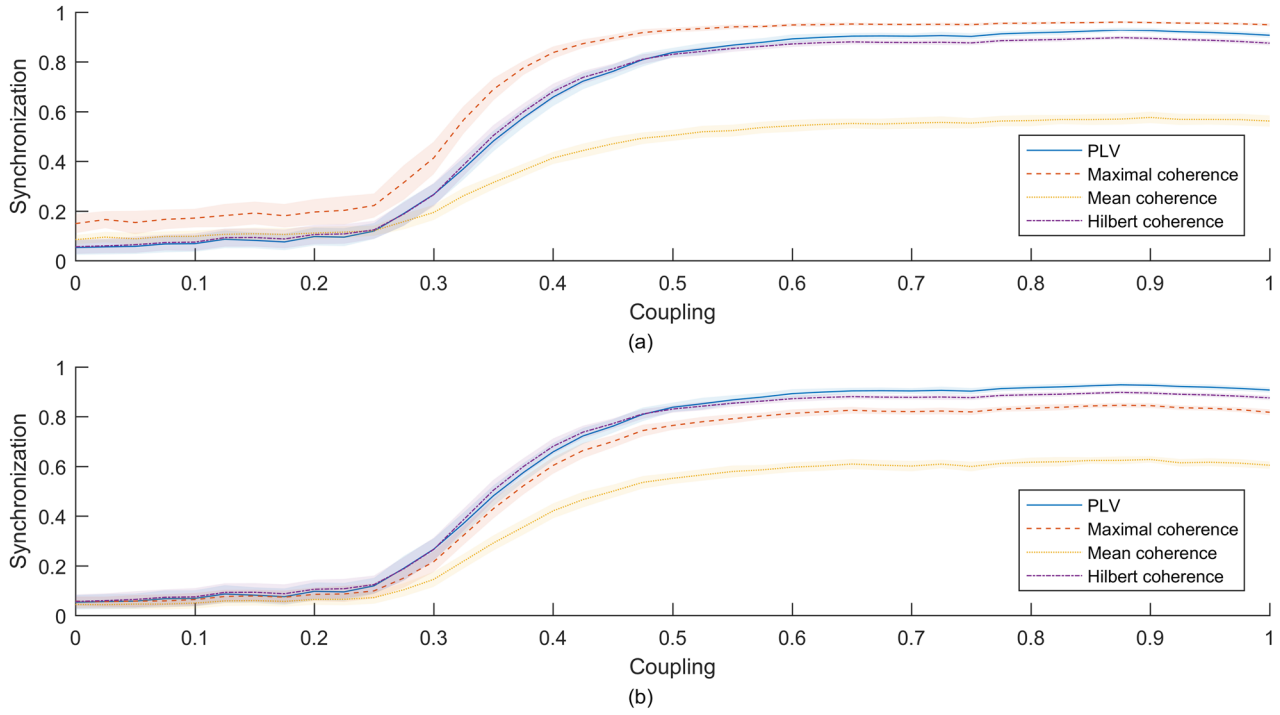

Figure 2. Values of the synchronization indices analyzed for different couplings for a pair of Rössler and Lorenz systems. The lines represent the mean value over 50 executions; the shadow areas represent the standard deviation. (a) PLV, maximal and average coherence for the band between $0.570 \cdot \pi$ and $0.600 \cdot \pi$ radians per sample. Coherence was calculated using a Hamming window of length 400 with 200 samples of overlapping. (b) The same that in (a), using a window length of 100 samples with 50 samples of overlapping. Note that only coherence-based metrics are affected by this change.

To obtain a band-wise coherence value, we adopted two different approaches using either the maximum or the mean value of coherence over the evaluated band. The values of PLV and the different coherence approaches are shown in figure 2(a). In either case, PLV and coherence evolve closely with increasing coupling, with coherence values slightly overestimated for low couplings. This overestimation is due to the lower number of phases used for the coherence calculation as compared to PLV or hCOH. Indeed, for the calculation of coherence data was split into 400 samples segments with 200 samples of overlapping, giving a total of 99 segments, and 99 independent phases. For the calculation of PLV, with a bandwidth of $0.030 \cdot \pi$ radians per sample, there is three effective sample for every 100 samples, giving a total of 600 effective samples, and thus 600 independent phases.

To address this problem, figure 2(b) shows the results using a smaller window for the coherence calculation. In this case, with 100 samples windows and 50 samples overlap, coherence is calculated using 399 phases. The overestimation for low coupling is visibly lower than in the previous case, but the maximal value of coherence drops to 0.85. Consequently, the reduction in the calculated synchronization can be explained from the frequency smoothing associated with a shorter time window. Due to the narrowness of the frequency peak of the Rössler oscillator, a larger amount of frequency smoothing forces the determination of coherence from phases coming from both the peak itself and the surrounding frequencies.

The effect is similar when evaluating the mean coherence over a band, as coupled frequencies are averaged together with non-coupled ones. However, PLV and hCOH do not show this effect, since the instantaneous phase of the signal is calculated over the whole band and not from an average of frequencies. We might conclude that in this instance, PLV is superior to coherence, as the position and narrowness of the oscillator over the observed band are not relevant. PLV could even handle an oscillator with varying frequency, given that the oscillatory frequency never goes outside of the observed band. On the other hand, maximal coherence would only return the coupling at the most common frequency, ignoring the coupling at other frequencies, and mean coherence would dramatically underestimate the connectivity.

Note that PLV and hCOH produce very similar results, even though both formulations are slightly different. In the special case in study the amplitude of the synthetic data is approximately constant, so the formulations given in (6) and (13) converge. This would not be true in a general case, where the different normalization would give rise to differences in the estimated connectivity. In the supplementary material we provide a more general case where PLV and hCOH results diverge.

4.3. Effects of volume conduction

As previously noted, one of the main criticisms to PLV and coherence when applied to brain signals is their sensitivity to volume conduction. It can be modeled as an instantaneous projection of one signal onto the other, giving rise to zero-lag synchronization. As shown in the methods section, Guido Nolte's proposed imaginary part of coherency (Nolte *et al* 2004, Ewald *et al* 2012) can be extended to the PLV, giving a set of zero-lag insensitive measures. Thus, to evaluate the

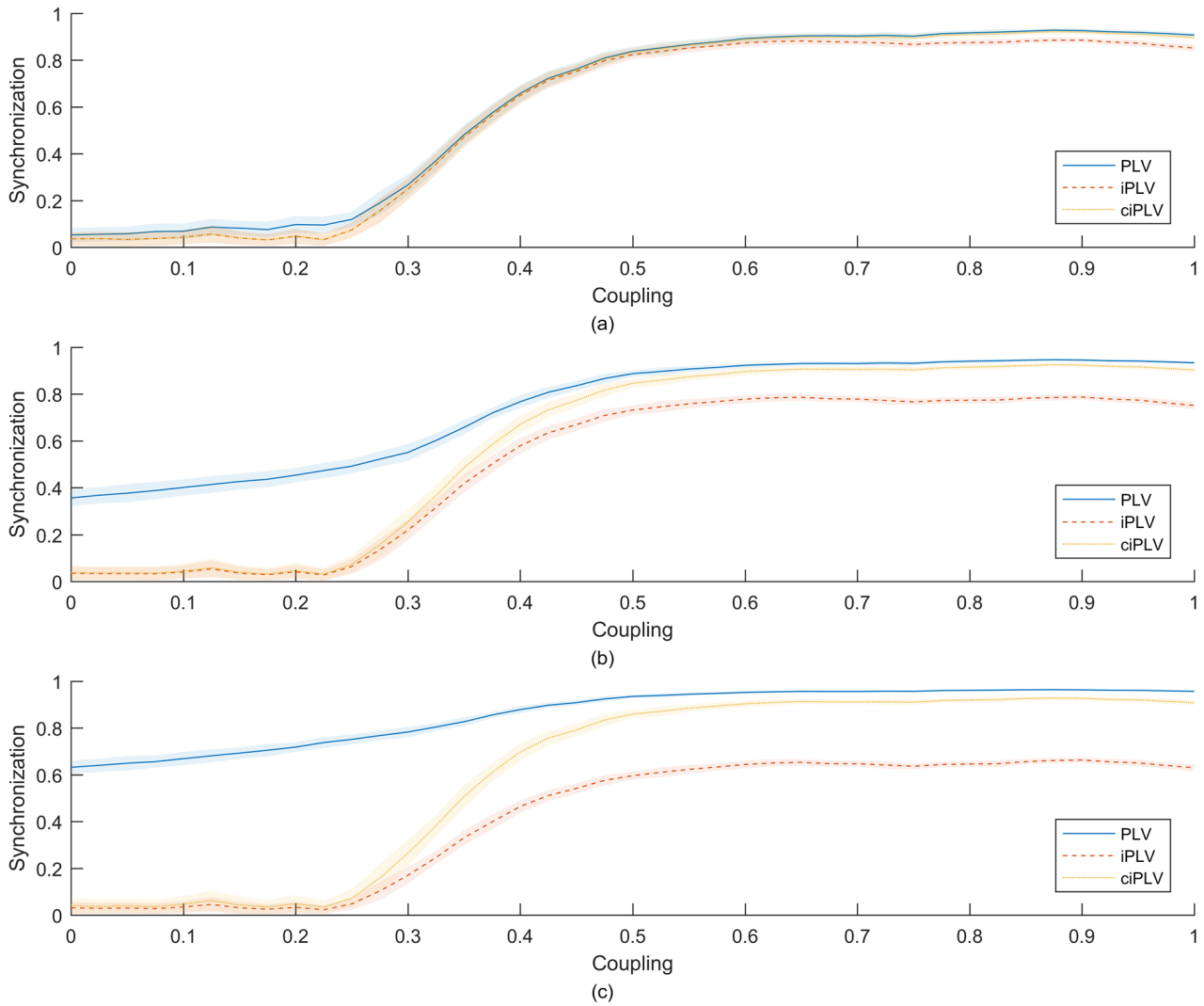

Figure 3. Values of the synchronization indices analyzed for different couplings for one Rössler and one Lorenz system. The lines represent the mean value over 50 executions, whereas the shadowed areas represent the standard deviation. (a) Synchronization estimated using PLV, iPLV and ciPLV. (b) The same that in (a), after adding a 10% of linear mixing to the signals. (c) Idem after adding a 20% of linear mixing.

behavior of these PLV derived metrics, we used the same pair of chaotic systems defined in the previous section. The systems show nonzero-lag phase synchronization, and the volume conduction can be introduced using instantaneous linear mixing (Haufe *et al* 2013, Porz *et al* 2014):

$$\begin{aligned} \tilde{x} &= x + V \cdot y \\ \tilde{y} &= y + V \cdot x \end{aligned} \tag{17}$$

where V is a parameter determining the amount of mixing. Figure 3 shows the synchronization estimated between the signals described in the previous section using the PLV, iPLV and ciPLV, and values of V of 0, 0.1 and 0.2.

Figure 3 shows the values of synchronization estimated using PLV and its derived metrics for different amounts of linear mixing, simulating volume conduction. In an ideal scenario, where the estimated synchronization is independent of the level of volume conduction, the synchronization estimated should be similar to that in figure 2. Thus, the effect of volume conduction can be observed as a deviation from this ideal behavior.

Figure 3(a) shows the estimated synchronization for $V = 0$ (without zero lag mixing), measured with PLV, iPLV and ciPLV. Values of PLV and ciPLV are almost identical, indicating that the synchronization of the pair of chaotic oscillator is nonzero lag. However, iPLV values are slightly lower, especially for high couplings, therefore indicating that the lag between both signals is almost, but not exactly, one-quarter of a cycle, so that the PLV complex vector has a small real component.

Figure 3(b), corresponding to $V = 0.1$, shows the effect of 10% volume conduction between the signals, introducing low spurious zero-lag synchronization. PLV values increase especially at lower couplings, whereas iPLV and ciPLV values remain unaffected. iPLV shows lower synchronization values that ciPLV, which remains almost identical to the $V = 0$ case. This is because the complex PLV vector now has an important real component that is completely ignored by iPLV. This shows that ciPLV is superior, as it uses this real component to normalize the imaginary one. Figure 3(c) shows the effect of 20% volume conduction, with similar results, but exacerbating

the estimation errors in PLV and iPLV. ciPLV, however, continues to extract the correct synchronization value.

From the results above, it might be deduced that, while iPLV is under-estimated in the presence of volume conduction, ciPLV is insensitive to it. To test this hypothesis, we calculated the values of PLV, iPLV and ciPLV for three fixed values of coupling ($C = 0.0$, $C = 0.4$ and $C = 0.6$) and varying values of volume conduction, from 0 to 0.999 99. Note that, as follows from equation (17), at a value of volume conduction of 1 both signals are exactly equal, so the numerator and the denominator in (15) would become zero, raising an indetermination.

The results for these simulations are shown in figure 4. For the three values of coupling studied, PLV and iPLV rapidly deviate from the ideal value, the first one increasing, and the second one decreasing. This is, in fact, the expected behavior: PLV is sensitive to volume conduction, thus the PLV value must increase with the amount of linear mixing; on the other hand, iPLV removes from the result the part due to volume conduction, strongly underestimating synchronization, especially for high values of linear mixing. ciPLV is also affected by the level of volume conduction, slightly overestimating the synchronization with increasing amounts of it, but the deviation is much smaller than in the other two cases. It is not possible to state that ciPLV is completely insensitive to volume conduction, but the error introduced by its presence is clearly much smaller than in the other two metrics.

5. Discussion

The formulation of PLV described in this paper allows for a speedup of two orders of magnitude in its calculation. This result is particularly relevant for the study of EEG/MEG source-space connectivity. Source-space models based in volumes, as the one used in the *results* section above, typically include around 2500 sources for the whole brain and around 1500 for gray matter. In the case of sources based in surfaces, which are normally used in minimum norm estimates (Pascual-Marqui *et al* 1994, Gramfort *et al* 2013), the number of sources ranges between 1000 and 10 000 per hemisphere. In both cases, even with subsampling, a whole brain connectivity study would take up to several hours per subject (and band).

The usual approach in those cases is the parcellation of the brain and the extraction of a representative time course per parcel (Korhonen *et al* 2014, Farahibozorg *et al* 2017). This reduces the number of time courses to the order of the hundreds, thereby allowing the estimation of whole-brain connectivity in a reasonable amount of time. However, this method forces the definition of homogeneous parcels, characterized by one unique time series, thus discarding some information associated to intra-area variability, which may have a significant impact on the connectivity results (Dimitriadis *et al* 2018). With the proposed algorithm, the calculation of source-to-source PLV would take only a few minutes per subject (and band), allowing the estimation of whole-brain connectivity for a whole study in a matter of hours.

In addition, our reformulation of the original PLV definition stressed the similarities between PLV and coherence. The comparison of both metrics, as applied to a pair of chaotic systems with different coupling levels, showed that PLV and coherence values are closely related, with PLV showing a better behavior in the (very realistic) situation where the coupling takes place over a whole (possibly narrow) band instead of at a single frequency. However, this is not a flaw of coherence, but the direct effect of its definition. PLV seems best to evaluate synchronization over a whole band, whereas coherence does to evaluate it at fixed frequencies.

Finally, we evaluated the behavior of PLV in the relevant case of volume conduction/source leakage as simulated by instantaneous linear mixing of the signals corresponding to the two systems. Analogously to coherency (Nolte *et al* 2004, Ewald *et al* 2012), we introduced the imaginary part of PLV and its corrected version. Both algorithms proved insensitive to volume conduction, but iPLV was flawed by the effect of a real component of the complex PLV vector. As for the imaginary part of coherency, iPLV only reaches the maximal value when the phase difference of the two signals is exactly $\pi/2$ radians, whereas it drops to almost zero when the phase difference is small but consistently nonzero. ciPLV corrects this behavior in a similar way to the lagged coherence or corrected imaginary part of coherency (Pascual-Marqui *et al* 2011, Ewald *et al* 2012), being both unbiased and insensitive to volume conduction/source leakage effects. This is not the first time that a metric based on the imaginary part of PLV is used (Dimitriadis *et al* 2017, 2018, Palva *et al* 2018, Wang *et al* 2018), but, to our knowledge, the expected behavior had not been thoroughly tested yet.

These two metrics mimic the corresponding metrics derived from the coherence and inherit their advantages. However, in the same way that PLV, the use of Hilbert's phases instead of the Fourier's ones allows the direct estimation of the synchronization over a whole frequency band, instead of a single frequency. This is especially important in cases where the frequency of a single oscillator varies within a certain frequency band, as is the case of high frequency bands of brain activity (high beta or gamma).

Both sensor and source space EEG/MEG connectivity must deal with the burden of zero-lag connectivity. Zero lag connectivity between two sensors is usually interpreted as the effect of a common source in MEG and an effect of volume conduction in EEG. At the source level, it is usually interpreted as an effect of source leakage. Both interpretations are based on the idea that brain signals cannot travel instantaneously between different parts of the brain, and any real connectivity entails the existence of a delay (Stam *et al* 2007). Even when some scenarios allow zero-lag connectivity in the brain (Kovach 2017), it is not possible to distinguish real zero-lag connectivity from volume conduction or source leakage.

Several measures have been developed to deal with this problem. The most notable of these measures is the imaginary part of coherence and its corrected variant. However, even when coherence is effectively a phase-synchronization measure, the results do not completely fit in the phase synchronization hypothesis in the brain. PLI is another measure developed to estimate phase connectivity ignoring the contribution of zero lag (Stam *et al* 2007), but the metric has

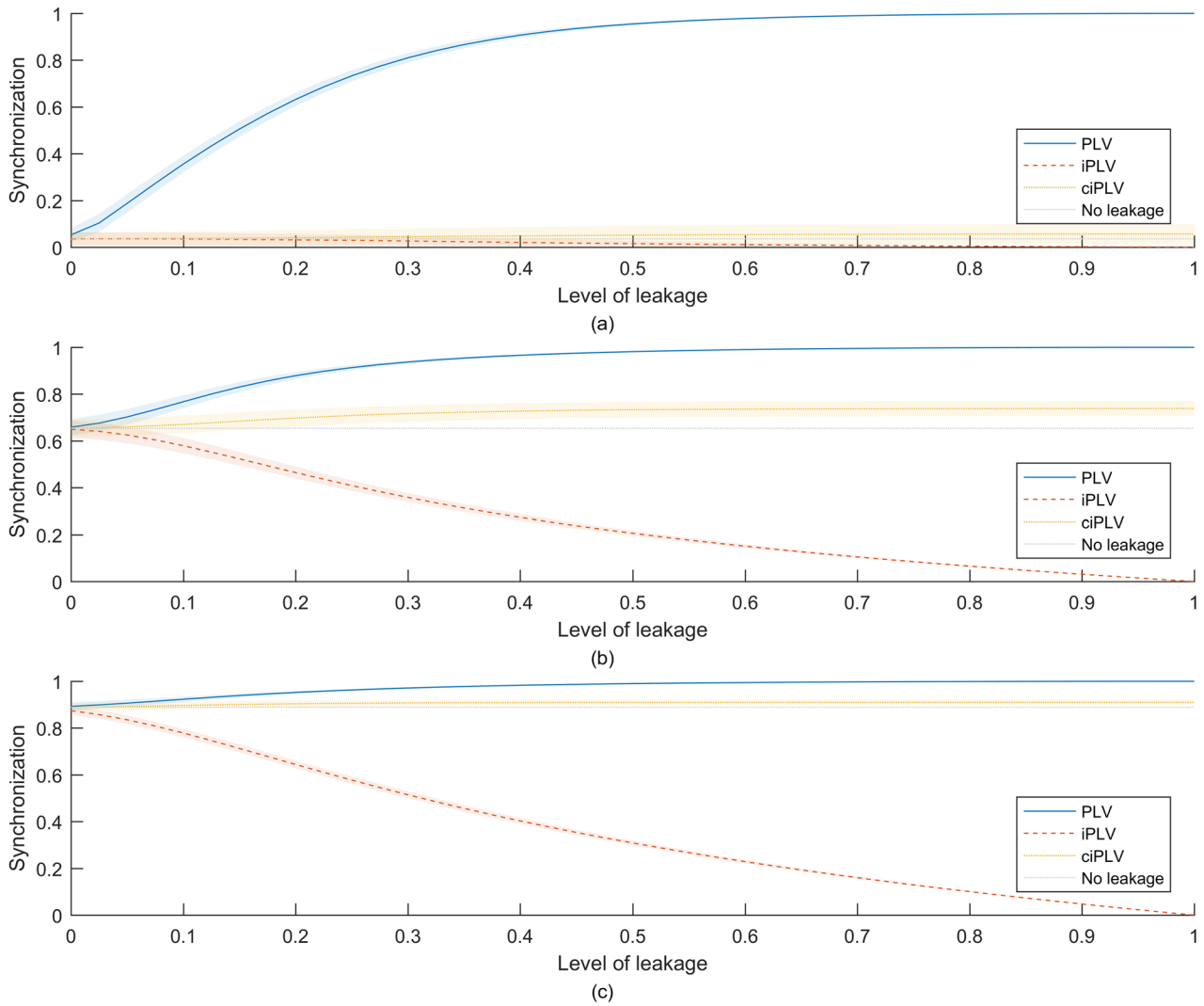

Figure 4. Values of the synchronization indices analyzed for different levels of volume conduction. The dark line represents the mean value over 50 executions, and the shadow line represents the standard deviation. The grey line represents the ideal value, i.e. the PLV value obtained with no volume conduction. (a) Synchronization estimated using PLV, iPLV and ciPLV when the oscillators present a value of coupling of 0. (b) The same that in (a), for a value of coupling of 0.4. (c) Idem for a value of coupling of 0.6.

showed a low test-retest reliability (Colclough *et al* 2016, Garcés *et al* 2016).

The variations of PLV proposed in this paper mimic their coherence counterparts and have proved effective to remove zero-lag connectivity while keeping intact nonzero-lag connectivity. This allows the study of sensor or source space EEG/MEG connectivity evaluating only real synchronizations, and completely ignoring volume conduction or source leakage ghost synchronizations. As discussed, this method, is not perfect (Kovach 2017), but in the same fashion that imaginary coherency (Nolte *et al* 2004) it allows for the evaluation of real connectivity without the influence of zero-lag interference.

Nonetheless, according to the latest results in the literature (Palva *et al* 2018, Wang *et al* 2018), no bivariate FC index (whether zero lag insensitive or not) is free from the effect of spurious detection of connectivity due to source leakage. Yet, when it comes to the detection of true connectivity, we believe that both iPLV and especially ciPLV are excellent choices that can be very efficiently estimated using the algorithms we introduced here.

6. Conclusion

In this paper, we propose a new formulation of the PLV algorithm that allows for fast computation by bypassing CPU-demanding calculations. The direct use of this new algorithm in Matlab makes possible a speedup by a factor of 100 from a vector-optimized implementation. Even though this implementation can be improved, it is unlikely that a similar speedup can be achieved using the original formulation.

Crucially, the optimization is only based in a mathematical reformulation of the original definition of PLV, in which the phase extraction and the exponentiation necessary for its calculation are replaced by a simple matrix multiplication. What is more, the algorithm can be implemented in any language, and the code provided in the appendix, created in Matlab, is completely machine-independent. Thus, differently to a toolbox we recently released (García-Prieto *et al* 2017), there is no need to compile the source code in C language, which makes the present formulation much easier to use. It must also be

noted that this paper only deals with the mathematical definition and as a consequence there may be room for computational improvement.

In conclusion, we have shown that despite its longevity, it is still possible to refine further the estimation of the (allegedly) most popular method of PS, the PLV index. The new formulation not only does allow for a much faster estimation of the index, but by highlighting the similarities between PLV and coherence, it also prompts the definition of two new measures insensitive to zero lag. These measures, derived from the vector product of the real and imaginary part of PLV, are analog to the (corrected) imaginary part of coherence and are therefore robust against volume conduction and source leakage effects. We hope these new developments will encourage further the use PLV-related measures in the analysis of neurophysiological data at the sensor and the source level alike.

Acknowledgment

The authors would like to thank X Bello for her review of the style of this manuscript. This study was partially supported by one project from the Spanish Ministry of Economy and Competitiveness (TEC2016-80063-C3-2-R) to EP and one pre-doctoral fellowship from the Spanish Ministry of Education (FPU13/06009) to RB.

Appendix

In the results section, we tested three different implementations for the calculation of PLV. In this appendix, we show the three different codes developed in Matlab to evaluate the behavior of each algorithm. In the codes provided, all three implementations are fed with the Hilbert analytic signal, with shape signals per samples per trials.

The first implementation includes two nested loops, so every pair of signals is evaluated independently. This option is the most memory conservative but also the slowest one, as only two signals are evaluated in each interaction. The code is:

```
[nc, ns, nt] = size(data);
phs = angle(data);

plv = zeros(nc, nc, nt);
for t = 1: nt
    for c1 = 1: nc
        for c2 = c1 + 1: nc
            dphs = phs(c1,:, t) - phs(c2,:, t);
            plv(c1, c2, t) = abs(mean(exp(1i * dphs)));
        end
    end
end
```

The second implementation is a vectorized and memory efficient version of the algorithm. In this implementation every signal is compared against all the other signals, avoiding two loops. This option uses slightly more memory but achieves a speedup of factor 2.5 over the original one. The code is:

```
[nc, ns, nt] = size(data);
phs = angle(data);

plv = zeros(nc, nc, nt);
for t = 1: nt
    tplv = complex(zeros(nc));
    for s = 1: ns
        dphs = bsxfun(@minus, phs(:, s, t), phs(:, s, t)');
        tplv = tplv + exp(1i * dphs);
    end
      plv(:,:, t) = abs(tplv / ns);
end
```

The last implementation uses the proposed algorithm. This algorithm allows the direct calculation of all source pairs at once, with negligible memory increase. In addition to avoiding the angle and exp functions, the implementation removes the need to use a loop over the signals. The code is:

```
[nc, ns, nt] = size(data);
ndat = data ./ abs(data);

plv = zeros(nc, nc, nt);
for t = 1: nt
    plv(:,:, t) = abs(ndat(:,:, t) * ndat(:,:, t)') / ns;
end
```

ORCID iDs

Ricardo Bruña https://orcid.org/0000-0003-1007-900X
Fernando Maestú https://orcid.org/0000-0002-3195-0071
Ernesto Pereda https://orcid.org/0000-0001-5965-164X

References

Brookes M J *et al* 2011 Measuring functional connectivity using MEG: methodology and comparison with fcMRI *NeuroImage* **56** 1082–104
Bruns A 2004 Fourier-, hilbert- and wavelet-based signal analysis: are they really different approaches? *J. Neurosci. Methods* **137** 321–32
Colclough G L *et al* 2016 How reliable are MEG resting-state connectivity metrics? *NeuroImage* **138** 284–93
Dimitriadis S I, Salis C, Tarnanas I and Linden D E 2017 Topological filtering of dynamic functional brain networks unfolds informative chronnectomics: a novel data-driven thresholding scheme based on orthogonal minimal spanning trees (OMSTs) *Front. Neuroinform.* **11** 28
Dimitriadis S I, López M E, Bruña R, Cuesta P, Marcos A, Maestú F and Pereda E 2018 How to build a functional connectomic biomarker for mild cognitive impairment from source reconstructed MEG resting-state activity: the combination of ROI representation and connectivity estimator matters *Front. Neurosci.* **12** 306
Ewald A *et al* 2012 Estimating true brain connectivity from EEG/MEG data invariant to linear and static transformations in sensor space *NeuroImage* **60** 476–88
Farahibozorg S-R, Henson R N and Hauk O 2017 Adaptive cortical parcellations for source reconstructed EEG/MEG connectomes *NeuroImage* **169** 23–45
Fries P 2015 Rhythms for cognition: communication through coherence *Neuron* **88** 220–35
Friston K J 1994 Functional and effective connectivity in neuroimaging: a synthesis *Hum. Brain Mapp.* **2** 56–78
Friston K J, Harrison L and Penny W 2003 Dynamic causal modelling *NeuroImage* **19** 1273–302
Garcés P, Martín-Buro M C and Maestú F 2016 Quantifying the test-retest reliability of magnetoencephalography resting-state functional connectivity *Brain Connect.* **6** 448–60
García-Prieto J, Bajo R and Pereda E 2017 Efficient computation of functional brain networks: towards real-time functional connectivity *Front. Neuroinform.* **11** 8
Gramfort A, Luessi M, Larson E, Engemann D A, Strohmeier D, Brodbeck C, Goj R, Jas M, Brooks T, Parkkonen L and Hämäläinen M 2013 MEG and EEG data analysis with MNE-Python *Front. Neurosci.* **7** 267
Haufe S *et al* 2013 A critical assessment of connectivity measures for EEG data: a simulation study *NeuroImage* **64** 120–33
Korhonen O, Palva S and Palva J M 2014 Sparse weightings for collapsing inverse solutions to cortical parcellations optimize M/EEG source reconstruction accuracy *J. Neurosci. Methods* **226** 147–60
Kovach C K 2017 A biased look at phase locking: brief critical review and proposed remedy *IEEE Trans. Signal Process.* **65** 4468–80
Lachaux J P *et al* 1999 Measuring phase synchrony in brain signals *Hum. Brain Mapp.* **8** 194–208
Le Bihan D *et al* 2001 Diffusion tensor imaging: concepts and applications *J. Magn. Reson. Imaging* **13** 534–46
López M E *et al* 2014 Alpha-band hypersynchronization in progressive mild cognitive impairment: a magnetoencephalography study *J. Neurosci.* **34** 14551–9
Mormann F *et al* 2000 Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients *Physica* D **144** 358–69
Nolte G 2003 The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors *Phys. Med. Biol.* **48** 3637–52
Nolte G *et al* 2004 Identifying true brain interaction from EEG data using the imaginary part of coherency *Clin. Neurophysiol.* **115** 2292–307
Nunez P L *et al* 1997 EEG coherency *Electroencephalogr. Clin. Neurophysiol.* **103** 499–515
Oostenveld R *et al* 2011 FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data *Comput. Intell. Neurosci.* **2011** 156869
Palva J M *et al* 2018 Ghost interactions in MEG/EEG source space: a note of caution on inter-areal coupling measures *NeuroImage* **173** 632–43
Pascual-Marqui R D *et al* 2011 Assessing interactions in the brain with exact low-resolution electromagnetic tomography *Phil. Trans. R. Soc.* A **369** 3768–84
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** 49–65
Porz S, Kiel M and Lehnertz K 2014 Can spurious indications for phase synchronization due to superimposed signals be avoided? *Chaos* **24** 033112
Quian-Quiroga R, Arnhold J and Grassberger P 2000 Learning driver-response relationships from synchronization patterns *Phys. Rev.* E **61** 5142–8
Stam C J, Nolte G and Daffertshofer A 2007 Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources *Hum. Brain Mapp.* **28** 1178–93
Taulu S and Simola J 2006 Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements *Phys. Med. Biol.* **51** 1759–68
van Veen B D *et al* 1997 Localization of brain electrical activity via linearly constrained minimum variance spatial filtering *IEEE Trans. Bio-Med. Eng.* **44** 867–80
Wang S H, Lobier M, Siebenhühner F, Puoliväli T, Palva S and Palva J M 2018 Hyperedge bundling: a practical solution to spurious interactions in MEG/EEG source connectivity analyses *NeuroImage* **173** 610–22