Analysis of the Phase Locking Index for Measuring of Interdependency of Cortical Signals Recorded in the EEG

Sazonov, Andrei V.; Chin, Keong Ho; Bergmans, Jan W.M.; Arends, Johan B.A.M.; Griep, Paul A.M.; Verbitskiy, Evgeny A.; Cluitmans, Pierre J.M.; Boon, Paul A.J.M.

Published in:
EPRINTS-BOOK-TITLE

Publication date:
2007

Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

Analysis of the Phase Locking Index for Measuring of Interdependency of Cortical Signals Recorded in the EEG

Andrei V. Sazonov, Chin Keong Ho, *Student Member, IEEE*, Jan W. M. Bergmans, *Senior Member, IEEE*, Johan B. A. M. Arends, Paul A. M. Griep, Evgeny A. Verbitskiy, *Member, IEEE*, Pierre J. M. Cluitmans, *Member, IEEE*, Paul A. J. M. Boon.

Abstract—**The phase locking index (PLI) was introduced to quantify in a statistical sense the phase synchronization of two signals. It has been commonly used to process biosignals. In this paper, we analyze the PLI for measuring the interdependency of cortical source signals (CSSs) recorded in the Electroencephalogram (EEG). The main focus of the analysis is the probability density function, which describes the sensitivity of the PLI to the joint noise ensemble in the CSSs. Since this function is mathematically intractable, we derive approximations and analyze them for a simple analytical model of the CSS mixture in the EEG. The accuracies of the approximate probability density functions (APDFs) are evaluated using simulations for the model. The APDFs are found sufficiently accurate and thus are applicable for practical intents and purposes. They can hence be used to determine the confidence intervals and significance levels for detection methods for interdependencies, e.g., between cortical signals recorded in the EEG.**

I. Introduction

THE Electroencephalogram (EEG) results as a mapping of brain signals into several channels. These channels are recorded by electrodes located on the scalp or inside the brain. The EEG is widely used for brain monitoring. To date, EEG analysis is mainly based on visual inspection by human experts, since available signal-processing methods are not completely satisfactory for automated detection and diagnostics. Nevertheless, signal-processing methods can substantially complement visual inspection and help to make EEG analysis objective [1].

In this article we investigate a measure called the phase locking index (PLI), which belongs to nonlinear measures. The PLI emerged from theoretical studies of oscillating (chaotic) systems with couplings. It was developed to quantify in a statistical sense the phase synchronization of such systems from experimental data and, thereby, to characterize their coupling [2], [3].

The PLI (and its modifications) was already used for many types of biosignals such as: MEG and EMG [2], ECG, fMRI and EEG. For the EEG, it was mainly used in relation to epilepsy [4]-[6]. Furthermore, the PLI was used to obtain insights about anesthesia and migraine, and to assess differences in perception to music. A rationale of its use for EEG analysis is the experimental evidence suggesting that the brain network is partly oscillatory [7], [8]. The PLI perfectly fits to this 'oscillatory' view of the brain, since it is designed for such systems.

Although the PLI is widely used in EEG analysis, it has to our knowledge not been investigated thoroughly. Firstly, its sensitivity to noise and artifacts has not been shown analytically. However, the sensitivity has been assessed using simulations [9], [10]. Secondly, the influence of crosstalk between the sources through biological tissues of the head has not been evaluated. In addition, the practical use of the PLI is often based on ad hoc approaches. To date, use of the PLI in detection and classification methods requires empirically determined thresholds which are typically obtained from surrogate data analysis [10]. The accuracy and applicability of such methods is limited [11]. Furthermore, practical use of the PLI typically involves filtering and windowing of sampled signals. These operations may significantly affect the PLI and lead to misinterpretations of the EEG [12].

In this paper, we address the aforementioned issues through an analysis of the PLI as a measure of interdependency of cortical source signals (CSSs) using the EEG. The analysis pertains to a simple analytical model for the source mixture in the EEG. The model has two sources of signals with mutual crosstalk controlled by a parameter. The sensitivity of the PLI to noise and the number of samples in the signals, and to the amount of crosstalk between them can be described by the probability density function (PDF) of the (measured) PLI. The mean and variance, as well as all other statistics of the PLI can be computed from the PDF. Since the PDF for the PLI is not tractable analytically, we derive approximate probability distribution functions (APDFs) for the PLI.

Using APDFs, we analyze the mean and variance of the PLI for the model. The mean is associated with the

Manuscript received April 2, 2007.

A. V. Sazonov, C. K. Ho, J. W. M. S. Bergmans, and P. J. M. Cluitmans are with the Dept. of Electrical Engineering, Eindhoven University of Technology, PO box 513, 5600MB, Eindhoven, The Netherlands (phone: +31 40 2473288; fax: +31 40 2466508; e-mail: a.sazonov@tue.nl; c.k.ho@tue.nl; j.w.m.s.bergmans@tue.nl; p.j.m.cluitmans@tue.nl).

P. A. M. Griep, B. A. M. Arends, P. A. J. M. Boon are with the Dept. of Clinical Neurophysiology, Epilepsy Center Kempenhaeghe, Heeze, The Netherlands. P. A. J. M. Boon is also with Dept. of Neurology, Ghent University Hospital, Ghent, Belgium (e-mails: griepp@kempenhaeghe.nl; arendsj@kempenhaeghe.nl; boonp@kempenhaeghe.nl).

E. A. Verbitskiy is with the Applied Mathematics & Statistics/DSP group, Philips Research Laboratories, Eindhoven, The Netherlands

interdependency of the source signals and the variance characterizes statistical uncertainty of each single measurement. We evaluate the accuracy of the analytically obtained mean and variance by comparison with the mean and variance computed numerically using Monte Carlo simulations. The simulations show that the APDFs are sufficiently accurate for practical intents and purposes.

The APDFs, being analytically tractable, clearly expose the relation between the amount of noise, crosstalk and number of samples in the source signals and the measured PLI. The APDFs can be used to compute the significance levels for interdependencies and likelihoods for interdependency detection methods. The formulas for the APDFs are compact and readily applicable for practical use.

II. Model

A. Physiological Considerations

We investigate the PLI using a model for the CSS mixture in the EEG. To build the model, we use the following physiological considerations.

We assume that within each cortical area q, neurons P_q can be partitioned into disjoint 2 subsets: $P_q = P_{q1} \bigcup P_{q2}$. Neurons P_{q1} function independently and generate spontaneous background signals that are different for different areas. We assume that the background signals for different areas are mutually independent and have the same power. Neurons P_{q2} are involved in oscillatory coupling, e.g. through reciprocal connections with the thalamus, under control of the brain stem and forebrain modulatory systems [7]. The amount of neurons P_{q2} for an area q depends on the coupling and is zero for uncoupled areas. Coupled neurons P_{q2} are synchronized in their discharging time. Since the coupling is oscillatory, their joint signal is oscillatory as well and has a prominent spectral peak corresponding to fundamental frequency $\tilde{f}_0$. As with all characteristics of biological systems, $\tilde{f}_0$ may fluctuate in time. However, we assume that it remains within a subband of width Ω for a given time period. Furthermore, we assume that the power spectrum of the background signal of P_{q1} can be treated as approximately flat in this subband.

In order to extract signals corresponding to P_{q2} and to analyze the interdependencies between areas q, the EEG is typically bandpass filtered [6], [13]. Since the exact value of $\tilde{f}_0$ depends on the state of the brain and is a priori unknown, the EEG is typically decomposed into multiple overlapping subbands, e.g. each having bandwidth Ω. Signals in these subbands can be analyzed separately and then the results can be combined.

We assume that the main sources for the EEG are cortical areas located below the electrodes. Although each electrode reflects primarily a signal of a source located exactly below the electrode, crosstalk from other sources exists that is caused by propagation of their signals through biological tissues of the head in the form of electrical fields.

B. Analytical Model

The physiological considerations described above, motivate us to use the following analytical model.

1) Modeled CSSs

For each source q, the CSS is modeled in a single subband of width Ω. This subband corresponds to a bandpass filtered EEG. Each modeled CSS contains a passband Gaussian noise signal, which mimics the background signal of P_{q1}, and may also contain a sinusoidal signal of frequency $\tilde{f}_0$, which mimics the oscillatory signal of P_{q2}. We denote the sampling frequency of the CSSs as $\tilde{f}_s$ and the center frequency of the subband as $\tilde{f}_c$.

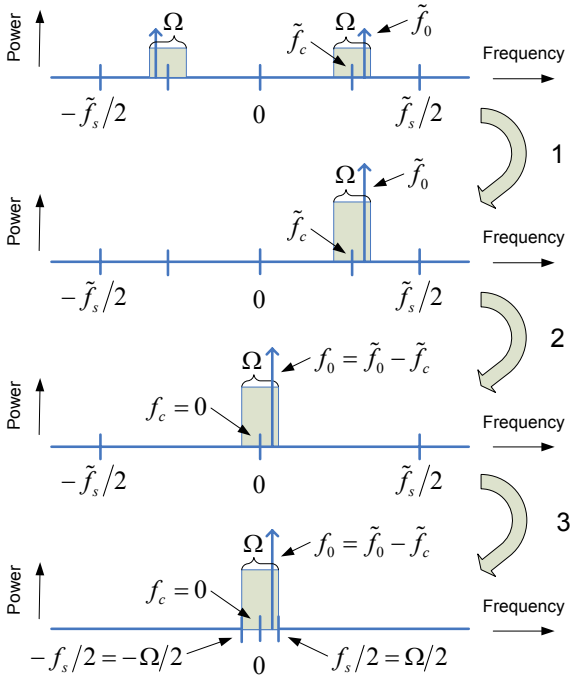

Fig. 1. The transformation of a passband signal into a baseband equivalent signal. Only one period of the spectrum is shown.

In order to make the model mathematically more tractable, we use equivalent baseband signals (without any loss of generality). The transformation of a real-valued passband signal into its baseband equivalent signal is illustrated in Fig. 1 for the fundamental interval of the power spectrum.

Firstly, we remove all negative frequencies and double the amplitudes of the positive ones. The result is called the analytical signal and is typically obtained using the Hilbert transform. Secondly, all frequencies are shifted downwards so that center frequency f_c of the shifted subband becomes zero. This shift can be accomplished by multiplication of the

signal with a complex exponential. Thirdly, we perform downsampling by a factor of $G = \tilde{f}_s / \Omega$, which is equivalent to changing the periodicity of the spectrum to a fundamental interval of width $f_s = \tilde{f}_s / G = \Omega$, where f_s is the re-sampling frequency.

As the result of this transformation, passband Gaussian noise in CSSs becomes complex-valued white Gaussian noise, denoted as n_q, and the sinusoid becomes a complex exponential, denoted as s_q, with different (shifted) frequency $f_0 = \tilde{f}_0 - \tilde{f}_c$. It must be noted that the power and the amount of information in CSSs is fully preserved after the transformation.

The model for the CSS is described by the following formula:

$$x_q[k] \triangleq \begin{cases} n_q[k], \text{ if source } q \text{ is uncoupled} \\ \qquad \text{with any other source;} \\ n_q[k] + s_q[k], \text{ if source } q \text{ is coupled} \\ \qquad \text{with one or more other sources,} \end{cases} \tag{1}$$

where $n_q[k]$ is a white Gaussian noise signal with mean 0 and variance $\sigma^2(n_q)$, and

$$s_q[k] \triangleq A_q e^{j(2\pi k f_0 / f_s + \theta_q)}, \tag{2}$$

where $A_q > 0$ is an amplitude, θ_q is a phase shift, $k = 1...K$ is a discrete time index, and $q = 1...Q$ is a source index. We assume that signals n_q are mutually independent and have equal variance for all sources q. For a source q, we define the signal-to-noise ratio (SNR) in x_q as

$$\mathrm{SNR}_q \triangleq \frac{\langle s_q^2[k] \rangle_k}{\langle n_q^2[k] \rangle_k}, \tag{3}$$

where $\langle \cdot \rangle_k$ means time-average. Without loss of generality, we set $\sigma^2(n_q) = 1$ for all q. In this case, SNR_q can be changed by varying A_q in s_q and simplifies to:

$$\mathrm{SNR}_q = \begin{cases} 0, \text{ if source } q \text{ is uncoupled} \\ \quad \text{with any other source;} \\ A_q^2, \text{ if source } q \text{ is coupled} \\ \quad \text{with one or more other sources.} \end{cases}$$

Here we recall that signal s_q is present in x_q if and only if the source q is coupled with one or more other sources, see (1). We assume A_q to be a priori known in the model. This assumption is needed to analyze the sensitivity of the PLI to noise for different (a priori known) SNRs.

2) Mixing Model

Signals x_q are used in a model for the CSS mixture in the EEG. In Fig. 2, a block diagram is shown for the model. The model includes only two sources $q = 1, 2$ with mutual crosstalk, and is defined by the following formula:

$$\begin{cases} c_1[k] \triangleq x_1[k] + \alpha x_2[k]; \\ c_2[k] \triangleq \alpha x_1[k] + x_2[k], \end{cases} \tag{4}$$

where c_q, $q = 1, 2$ are signals of the EEG channels, and constant α determines the amount of crosstalk.

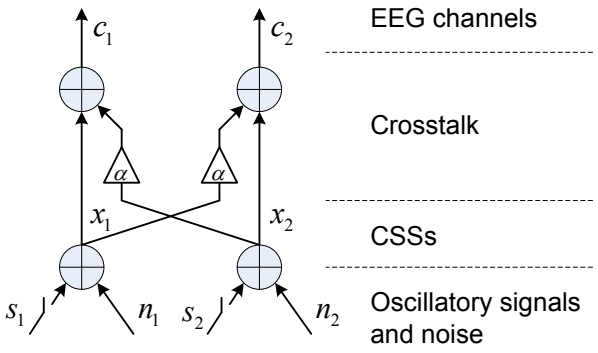

Fig. 2. A block diagram for the simplest model for the CSS mixture. Different parts of the model are explained on the right. See the text for further details.

III. Phase Synchronization and the PLI

The PLI was developed to quantify phase synchronization of oscillatory systems from experimental data. For signals, phase synchronization is typically measured in two steps: a) estimation of instantaneous phases of the signals, and b) statistical quantification of a phase relationship [2], [3].

For the first step, two common methods can be distinguished in the literature. These methods are the convolution of the signals with a complex wavelet, and the Hilbert transform. Both methods provide unambiguous complex-valued representations of the real-valued signals. The previously reported differences between these two methods are minor, and the methods were concluded to be equivalent for neuro-signals [14].

For a complex-valued signal c, the instantaneous phase φ can be obtained analytically as $\varphi \triangleq \mathrm{Im}(\ln(c))$. For two periodic signals c_1 and c_2 with fundamental frequencies f_1 and f_2 that are related as $f_1 \approx f_2$, phase synchronization can be described as a phase locking condition $|\varphi_1 - \varphi_2| < C$, where C is some constant (see [3] for analytical justification and generalization for the case $nf_1 \approx mf_2$, where n and m are some integers). Such synchronization may exist when the noise is negligible. If the noise is strong or if the signals are chaotic, large phase fluctuations and rapid 2π phase jumps (phase slips) may be observed and the condition may not be fulfilled. In this case, phase synchronization should be treated in a statistical sense [2]. It

was shown that the presence of a dominant peak in the distribution of the cyclic relative phase $\Psi \triangleq (\varphi_1 - \varphi_2) \bmod 2\pi$ can be understood as a phase synchronization in a statistical sense [2].

Several methods were proposed to quantify the distribution of Ψ. We use the phase locking index (PLI) described in [2] since it is most widely used, see e.g. [4]. The PLI is defined as:

$$\gamma \triangleq \left| \left\langle e^{j\Psi[k]} \right\rangle_k \right|, \tag{5}$$

where $\langle \cdot \rangle_k$ means time average.

In case of strong synchronization between the signals, γ is close to one. If synchronization is weak, then γ has a small value. It should be noted that γ is sensitive only to the phase difference of the signals.

IV. Results

A. Bandwidth and Effective Number of Samples

The bandwidth Ω is an important parameter for EEG analysis. It determines the sampling frequency $f_s = \Omega$ for equivalent baseband signals. Another relevant parameter is the duration of the EEG epoch, denoted by T. Parameters T and Ω together determine number of data samples K in the equivalent baseband signals: $K = Tf_s = T\Omega$. The larger number of samples in a real-valued passband signal carries no extra information, i.e. some of the samples are redundant. For instance, a single channel of a typical EEG epoch with $T = 10$ s and $\Omega = 2$ Hz and $\tilde{f}_s = 100$ Hz contains $T\tilde{f}_s = 1000$ samples, while an equivalent baseband signal contains only $T\Omega = 20$ samples. This redundancy of the passband signals may cause spurious detection of phase synchronization between signals [12].

B. Working SNR range

Signal processing methods are typically used to facilitate visual EEG analysis. It is useful to assess performances for the methods in a range of SNR values in CSSs that are relevant for visual analysis. In order to assess the bounds for this range we perform a simple experiment with physiological EEGs. We took two EEG signals: a signal recorded from a healthy person (resting with eyes opened) and a signal with clear spike-and-wave complexes recorded from an epileptic patient during a typical absence seizure. We use the first signal as the background activity and the second signal as the activity caused by oscillatory couplings. A human expert (an experienced neurologist) is asked to analyze visually mixtures of these signals with different SNR for the presence of any epileptic pattern (the SNR is defined by (3) for a subband of width $\Omega = 2$ Hz.)

We found that the expert can visually distinguish the epileptic patterns starting from SNR = 8.5 dB and is confident about the type of the patterns starting from SNR = 12.5 dB. Thus, the range

$$R_v = [8.5 \ldots 12.5] \text{ dB} \tag{6}$$

is considered to be of particular interest.

Since epileptic patterns corresponding to SNR < 8.5 dB are not assessable by visual inspection, signal processing methods can be very helpful to detect them. Epileptic patterns corresponding to SNR > 12.5 dB are obvious for a human expert. This, however, does not necessarily imply that they are easy to detect by signal processing methods.

C. Analytical Results

We analyze the probability density function (PDF) of the PLI for the CSS mixture model. Since the exact analysis is mathematically intractable, we use approximations.

1) No crosstalk

In Appendix, the analysis is performed for signals c_q, $q = 1,2$ of the model defined by (4) in the absence of crosstalk (i.e. $\alpha = 0$, $c_q = x_q$) and for a sufficiently large number of samples K. Two approximate PDFs (APDFs) of γ defined in (5), $D_1(\gamma)$ and $\hat{D}_1(\gamma)$, are derived for high SNR in x_q. Furthermore, the exact PDF D_2 is derived for the special case SNR = 0.

The APDF D_1 involves two Gaussians and is shown below:

$$D_1 = \frac{K}{\sigma_R \sqrt{2\pi}} \left(e^{\frac{-(\gamma-\mu)^2 K}{2\sigma_R^2}} + e^{\frac{-(\gamma+\mu)^2 K}{2\sigma_R^2}} \right), \ \gamma \geq 0, \tag{7}$$

where $\mu = e^{-\frac{\sigma^2(\nu)}{2}}$, $\sigma_R^2 = \frac{1}{2}\left(1 - e^{-\sigma^2(\nu)}\right)^2$, and $\sigma^2(\nu) = \frac{\sigma^2(n_1)}{A_1^2} + \frac{\sigma^2(n_2)}{A_2^2}$, see Appendix for details. It should be noted that D_1 is derived for $A_q \gg \sigma^2(n_q)$, $q = 1,2$. However, by means of simulations we will show that D_1 is also a fair approximation for the case $A_q \sim \sigma^2(n_q)$. Furthermore, σ_R^2 and μ are mutually related as: $\sigma_R^2 = \frac{1}{2}\left(1 - \mu^2\right)^2$.

A computationally efficient simplification of D_1 can be obtained that involves only one Gaussian:

$$\hat{D}_1 = \frac{K}{\sigma_R \sqrt{2\pi}} \left(e^{\frac{-(\gamma-\mu)^2 K}{2\sigma_R^2}} \right), \ \gamma \geq 0. \tag{8}$$

This simplification can be made because the second Gaussian in (7) is small compared to the first one for $A_q \gg \sigma^2(n_q)$. The mean and variance for the distribution $\hat{D}_1$ can be computed using standard formulas for the normal distribution.

The PDF D_2 (no approximations are used, SNR = 0) is given by a Rayleigh distribution:

$$D_2 = 2K\gamma e^{-\gamma^2 K},\ \gamma \geq 0. \tag{9}$$

2) Crosstalk present

Let us now investigate how crosstalk between the sources affects APDFs D_1 and $\hat{D}_1$. For $\alpha > 0$, each signal c_q of the model can be rewritten in the following way:

$$\begin{aligned} c_q &= s_q + n_q + \alpha\left(s_p + n_p\right) = \\ &\left(s_q + \alpha s_p\right) + \left(n_q + \alpha n_p\right) = \hat{s}_q + \hat{n}_q, \end{aligned} \tag{10}$$

where $\hat{s}_q \triangleq s_q + \alpha s_p$ is an oscillatory component with the same frequency as s_q but different amplitude and phase, and $\hat{n}_q \triangleq n_q + \alpha n_p$ is a white Guassian noise signal with $\mu = 0$ and $\sigma^2 = \sigma^2\left(n_q\right) + \alpha\sigma^2\left(n_p\right)$, $p = 1,2$, $q = 1,2$, $p \neq q$. It can be shown that in this case $\sigma^2(v)$ (used for μ and σ_R^2 in (7) and (8)) is $\sigma^2(v) = \dfrac{\sigma^2\left(\hat{n}_1\right)}{A_1^2} + \dfrac{\sigma^2\left(\hat{n}_2\right)}{A_2^2}$. Thus, (10) and the (A)PDF together expose the sensitivity of the PLI to the amount of crosstalk between the signals.

The (A)PDFs D_1, $\hat{D}_1$, and D_2 describe the PLI statistically, i.e. the mean and variance as well as other statistics of the PLI can be obtained from them for given SNR and K. In order to assess the accuracy of the functions, we compare the mean and variance obtained analytically from (7)-(9) with the mean and variance obtained numerically using Monte Carlo simulations with 1000 realizations of x_q. For this comparison, we fix $K = Tf_s = 20$, the choice of which corresponds to a typical bandpass filtered signal of ten seconds length, see Section IV-A. For larger values of K, however, the accuracy is improved. The SNR range used for the comparison is $\mathrm{SNR} \in [-20 \ldots 20]$ dB. The lowest SNR = −20 dB corresponds to e.g., the very onset of a focal epileptic seizure. The highest SNR = 20 dB corresponds to e.g., spike-and-wave patterns of a generalized epileptic seizure, when the amplitudes of the patterns can substantially exceed the amplitudes of the spontaneous background signal. Intermediate values of SNR cover most other cases including the range R_v defined by (6). For simplicity we use equal SNR for both sources. We take $\tilde{f}_s = 100$ Hz and $f_s = \Omega = 2$ Hz, which is a commonly used bandwidth for EEG analysis [4], [13]. Furthermore, we assume that $\tilde{f}_0 = 10$ Hz and $\tilde{f}_c = 9.5$ Hz in the passband CSSs, which are some typical values for the EEG. The corresponding frequencies in baseband signals x_q are $f_0 = 0.5$ Hz, $f_c = 0$ Hz, see Section II-B. As mentioned in Section III, the PLI is insensitive to phases θ_q in signals s_q. Therefore, we let phases $\theta_q = 0$.

The results are shown for the mean of the PLI in Fig. 3a and for the variance in Fig. 3b, for the case $\alpha = 0$. The accuracy of the APDFs for other values of α is qualitatively very similar. It can be seen from the figures that the mean and variance derived from D_1 and $\hat{D}_1$ are accurate for SNR above 10 dB and at least fairly good for SNR above −5 dB. We remark that $\hat{D}_1$ is significantly different from D_1 only for SNR < −7 dB. It can also be seen from Fig. 3 that D_2 accurately describes the mean and variance for SNR < −12 dB. Furthermore, we remark that within the SNR range R_v (6), the mean and variance are accurately described by D_1 as well as $\hat{D}_1$. For a typical epoch length of ten seconds, we note that the variance is much smaller compared to the mean.

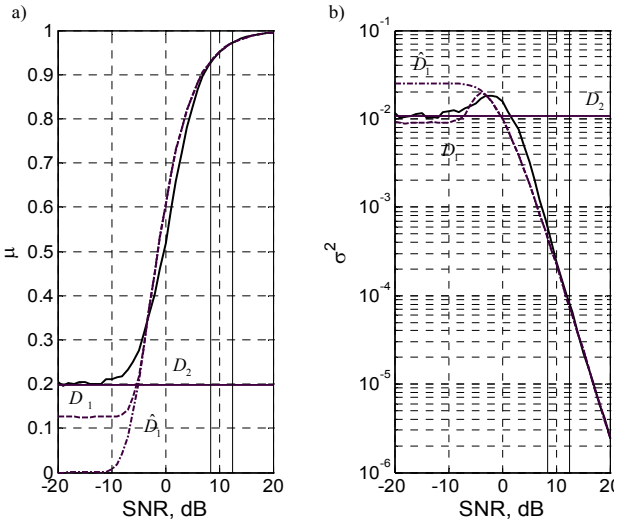

Fig. 3. A comparison of the analytically and empirically computed (a) mean and (b) variance of the PLI. Solid curves corresponds to results of simulations, dashed curves are obtained using D_1, dash-doted curves are based on $\hat{D}_1$, and horizontal solid lines correspond to D_2. The range R_v is shown by vertical lines.

V. Discussion and Conclusion

One of the measures used widely for EEG analysis is the *phase locking index* (PLI). Its usefulness has been confirmed experimentally, at least for some EEGs [6], [4]. In this article, we analyze the PLI as a measure of interdependency of *cortical source signals* (CSSs) recorded in the EEG. We analyze the PLI on a theoretical base. The main objective of the analysis is to derive the probability density function (PDF) of the (measured) PLI that describes sensitivity of the PLI to noise and the number of samples in the CSSs, and to the amount of crosstalk between them. Since the PDF is mathematically intractable, we derive and analyze its approximations (APDFs) for a model of CSSs mixture in the EEG.

In order to build a simple and efficient model, we show that passband signals, which are typically used in EEG analysis, are equivalent to baseband signals with lower sampling frequency and thus fewer samples within the same time internal. Since the baseband signals are analytically more tractable, they are used in the model. Furthermore, the correspondence between passband and baseband signals exposes the relation between the bandwidth and the effective number of samples in the passband signals – an issue which is sometimes overlooked in the (EEG-related) literature.

Approximate probability density functions (APDFs) are derived (7)-(9) that expose the behavior of the PLI for different amounts of noise in the modeled source signals and the epoch length. It is found that APDFs accurately approximate the mean and variance of the PLI for a wide range of SNR in the signals, including the range R_v that is relevant for visual analysis of epileptic EEGs. The APDFs can be used to determine the confidence intervals and significance levels for interdependency detection methods.

We notice that our analysis of the PLI is limited to sinusoidal linearly interdependent signals in the presence of additive Gaussian noise. This substantially simplifies the analysis, and facilitates interpretation of the results. However, the PLI was developed for more complex signals including chaotic ones. We do not use prior information about the signals in the simulations and it is likely that results of the analysis carry over to other (more complex) signals as well.

VI. Appendix

Let us analyze the sensitivity of the PLI to the white additive Gaussian noise in the input signals:

$$c_q[k] = s_q[k] + n_q[k] = A_q e^{j\varphi_q[k]} + B_q[k] e^{j\phi_q[k]}, \qquad (11)$$

where $s_q[k] \triangleq A_q e^{j\varphi_q[k]}$ is an exponential signal with phase $\varphi[k] \triangleq 2\pi k f_0 / f_s + \theta_q$, and $n_q[k] \triangleq B_q[k] e^{j\phi_q[k]}$ is some white additive Gaussian noise with mean 0 and variance $\sigma^2(n_q)$, f_s is the sampling frequency, $A_q, B_q, f_0, f_s \in \mathbb{R}^+$, $\theta_q, \varphi_q, \phi_q \in \mathbb{R}$, $k = 1...K$, $q = 1,2$. For the sake of simplicity, we omit the time index k in the following formulas. We note that this signals correspond to the signals of the CSS mixture model described by (4) with $\alpha = 0$, i.e. without crosstalk.

Our objective is to find the probability density function (PDF) for the PLI computed for c_q. Since the exact PDF appears mathematically intractable for the general case, we use approximations. We proceed as follows. First we derive an approximate PDF (APDF) D_1 and a computationally more efficient APDF $\hat{D}_1$ for the case $A_q \gg \sigma^2(n_q)$. Then we derive a PDF D_2 for the special case $A_q = 0$. The accuracies of D_1, $\hat{D}_1$ and D_2 are evaluated in Section IV-C.

Mathematically, the phase of any non-zero complex number c, denoted as $\measuredangle c$, equivalently can be computed as $\mathrm{Im}(\ln(c))$. Therefore, it follows that:

$$\measuredangle c_q \triangleq \mathrm{Im}(\ln(c_q)) = \mathrm{Im}\left(\ln\left(A_q e^{j\varphi_q} + B_q e^{j\phi_q}\right)\right) =$$
$$\mathrm{Im}\left(\ln\left(e^{j\varphi_q}\right) + \ln(A_q) + \ln\left(1 + \frac{B_q}{A_q} e^{j(\phi_q - \varphi_q)}\right)\right) =$$
$$\varphi_q + \measuredangle(1 + \nu_q),$$

where $\nu_q \triangleq \frac{B_q}{A_q} e^{j(\phi_q - \varphi_q)}$ is some modified noise with distribution $\mathrm{N}(0, \sigma^2(\nu_q))$, where $\sigma^2(\nu_q) \triangleq \frac{\sigma^2(n_q)}{A_q^2}$. We notice that the variance $\sigma^2(\nu_q)$ is the inverted SNR in the signal c_q, and $\sigma^2(\nu_q) \ll 1$ since $A_q^2 \gg \sigma^2(n_q)$.

Since $\sigma^2(\nu_q)$ is small, $\measuredangle(1+\nu_q)$ can be approximated by Taylor expansion as

$$\measuredangle(1+\nu_q) = \mathrm{Im}(\ln(1+\nu_q)) \approx \mathrm{Im}(\nu_q) + O(\nu_q^2).$$

Ignoring high order terms, $\measuredangle(1+\nu_q)$ is distributed as $\mathrm{N}\left(0, \frac{\sigma^2(\nu_q)}{2}\right)$ assuming that the real and imaginary parts of n_q, and therefore of ν_q, have the same variance.

Now we can write the phase difference of two signals $q = 1,2$ as:

$$\measuredangle c_1 - \measuredangle c_2 = \triangle\varphi + \measuredangle(1+\nu_1) - \measuredangle(1+\nu_2) \approx$$
$$\triangle\varphi + \mathrm{Im}(\nu_1) - \mathrm{Im}(\nu_2) = \triangle\varphi + \nu, \qquad (12)$$

where $\nu \triangleq \mathrm{Im}(\nu_1) - \mathrm{Im}(\nu_2)$ is approximately Gaussian $\mathrm{N}(0, \sigma^2(\nu))$, and $\sigma^2(\nu) = \sigma^2(\nu_1) + \sigma^2(\nu_2)$ because $\mathrm{Im}(\nu_1)$ and $\mathrm{Im}(\nu_2)$ are independent.

Now let us consider the PLI for the signals c_1 and c_2 as defined by (5): $\gamma \triangleq \left|\left\langle e^{j(\measuredangle c_1 - \measuredangle c_2)} \right\rangle\right|$, where $\langle\cdot\rangle$ denotes the average over time. According to the approximation (12): $\gamma \approx \left|\left\langle e^{j(\triangle\varphi + \nu)} \right\rangle\right| = \left|\left\langle e^{j\nu} \right\rangle\right|$.

Let us denote $r \triangleq e^{j\nu}$ for convenience. Since $\langle r \rangle$ is computed by averaging of a large data set r, we may apply the Central Limit Theorem. The Central Limit Theorem (CLT) states that for sufficiently large size of data K, $\langle r \rangle$ approaches the normal distribution $\mathrm{N}\left(\mu(r), \frac{\sigma^2(r)}{K}\right)$, or

equivalently $\langle r \rangle \approx \mu(r) + \omega$, where ω is $\mathrm{N}\left(0, \frac{\sigma^2(r)}{K}\right)$. We omit index r for $\mu(r)$ and $\sigma^2(r)$ in formulas below for convenience.

It can be shown that $\mathrm{Re}(\mu) = e^{-\frac{\sigma^2(\nu)}{2}}$ and $\mathrm{Im}(\mu) = 0$. We denote the real and imaginary parts of ω as ω_R and ω_I respectively, and the variances of them respectively as σ_R^2/K and σ_I^2/K. In order to simplify computation of $\gamma \triangleq |\langle r \rangle|$ we will use an approximation $|\langle r \rangle| \approx |\mu + \omega_R|$ which we will justify shortly for $\sigma^2(\nu) \ll 1$. Given that the distribution of ν is $\mathrm{N}(0, \sigma^2(\nu))$, the following expressions can be obtained:

$$\begin{aligned} \sigma_R^2 &\triangleq \mathrm{E}\left(\mathrm{Re}^2(r)\right) - \mathrm{E}\left(\mathrm{Re}(r)\right)^2 = \frac{1}{2}\left(1 - e^{-\sigma^2(\nu)}\right)^2; \\ \sigma_I^2 &\triangleq \mathrm{E}\left(\mathrm{Im}^2(r)\right) - \mathrm{E}\left(\mathrm{Im}(r)\right)^2 = e^{-\sigma^2(\nu)} \sinh\left(\sigma^2(\nu)\right). \end{aligned} \quad (13)$$

It should be noted that due to a non-linear transformation of input noise ν, $\sigma_R^2 \neq \sigma_I^2$, i.e. the noise is unequally distributed among the real and imaginary parts of r.

Now, we can write:

$$\gamma \triangleq |\langle r \rangle| = |\mu + \omega_R + j\omega_I| = \sqrt{(\mu + \omega_R)^2 + \omega_I^2}.$$

Furthermore, $(\mu + \omega_R)^2 \gg \omega_I^2$ since $\sigma^2(\nu) \ll 1$ that justifies the approximation $|\langle r \rangle| \approx \sqrt{(\mu + \omega_R)^2} = |\mu + \omega_R|$.

Therefore, $\langle r \rangle \approx \mu + \omega_R$ and is approximately

$$\langle r \rangle \sim \mathrm{N}\left(\mu, \frac{\sigma_R^2}{K}\right). \quad (14)$$

Recalling that $\gamma \triangleq |\langle r \rangle|$, we can obtain APDF for γ denoted as D_1 as a sum of two Gaussians that are (14) and its reflection with respect to the ordinate axis $\mathrm{N}\left(-\mu, \frac{\sigma_R^2}{K}\right)$. The APDF D_1 is presented by (7).

Furthermore, a simplification of D_1 can be obtained which involves only one Gaussian (14), and is presented by (8).

Let us analyze the case $A_q = 0$ (4), i.e. when the exponential signals s_q are absent. In this case, $\mu = 0$ and $\sigma^2 = 1$. The distribution of $\langle r \rangle$ follows immediately from the CLT: $\langle r \rangle \sim N(0, 1/K)$. Finally, the PDF of $\gamma = |\langle r \rangle|$ is known as Rayleigh distribution with the parameter $b^2 = \frac{1}{2K}$, assuming that the real and imaginary parts are independent. This distribution is presented by (9). For this distribution, $\mu = b\sqrt{\frac{\pi}{2}} = \sqrt{\frac{\pi}{4K}}$ and $\sigma^2 \triangleq \frac{4-\pi}{2} b^2 = \frac{4-\pi}{4K}$.

As shown in Section IV-B, D_1, $\hat{D}_1$ and D_2 can also be utilized in the presence of crosstalk, i.e. for $\alpha > 0$.

REFERENCES

[1] E. Niedermeyer and F. Lopes da Silva, *Electroencephalography:basic principles, clinical applications, and related fields*, 4 ed. Lippincott Williams and Wilkins, 1999.

[2] P. Tass, M. G. Rozenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, and H.-J. Freund, "Detection of n:m phase locking from noisy data: application to magnetoencephalography," *Phys.Rev.Lett.*, vol. 81, no. 15, pp. 3291-3294, Oct. 1998.

[3] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, "Phase synchronization of chaotic oscillators," *Phys.Rev.Lett.*, vol. 76, no. 11, pp. 1804-1807, Mar. 1996.

[4] M. Chavez, M. Van Quyen, V. Navarro, M. Baulac, and J. Martinerie, "Spatio-temporal dynamics prior to neocortical seizures: amplitude versus phase couplings," *IEEE Trans.Biomed.Eng.*, vol. 50, no. 5, pp. 571-583, May 2003.

[5] F. Mormann, R. G. Andrzejak, T. Kreuz, C. Rieke, P. David, C. E. Elger, and K. Lehnertz, "Automated detection of a preseizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients," *Phys.Rev.E.Stat.Nonlin.Soft.Matter Phys.*, vol. 67, no. 2 Pt 1, pp. 021912, Feb. 2003.

[6] R. Quian Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, "Performance of different synchronization measures in real data: a case study on electroencephalographic signals," *Phys.Rev.E.Stat.Nonlin.Soft.Matter Phys.*, vol. 65, no. 4 Pt 1, pp. 041903, Apr. 2002.

[7] M. Steriade, "Corticothalamic resonance, states of vigilance and mentation," *Neuroscience*, vol. 101, no. 2, pp. 243-276, 2000.

[8] F. H. Lopes da Silva, W. Blanes, S. N. Kalitzin, J. Parra, P. Suffczynski, and D. N. Velis, "Dynamical diseases of brain systems: different routes to epileptic seizures," *IEEE Trans.Biomed.Eng.*, vol. 50, no. 5, pp. 540-548, May 2003.

[9] T. Kreuz, F. Mormann, R. G. Andrzejak, A. Kraskov, K. Lehnertz, P. Grassberger, "Measuring synchronization in coupled modeled systems: A comparison of different measures," *Phys.D*, vol. 225, no. 1, pp. 29-42, Jan. 2007.

[10] A. Porta, N. Montano, R. Furlan, C. Cogliati, S. Guzzetti, T. Gnecchi-Ruscone, A. Malliani, "Automatic classification of interference patterns in driven event series: application to single sympathetic neuron discharge forced by mechanical ventilation," *Biol.Cyb.*, vol. 91, no. 4, pp. 258-273, Oct. 2004.

[11] C. J. Stam, "Nonlinear dynamical analysis of EEG and MEG: review of an emerging field," *Clin.Neurophysiol.*, vol. 116, no. 10, pp. 2266-2301, Oct. 2005.

[12] L. Xu, Z. Chen, K. Hu, H. E. Stanley and P. Ch. Ivanov, "Spurious detection of phase synchronization in coupled nonlinear oscillators," *Phys.Rev.E*, vol. 73, 065201, Jun. 2006.

[13] K. Ansari-Asl, J. J. Bellanger, F. Bartolomei, F. Wendling, and L. Senhadji, "Time-frequency characterization of interdependencies in nonstationary signals: application to epileptic EEG," *IEEE Trans.Biomed.Eng.*, vol. 52, no. 7, pp. 1218-1226, Jul. 2005.

[14] M. Van Quyen, J. Foucher, J. Lachaux, E. Rodriguez, A. Lutz, J. Martinerie, and F. J. Varela, "Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony," *J.Neurosci.Methods*, vol. 111, no. 2, pp. 83-98, Oct. 2001.