



HAL
open science

Cerebral Synchrony Assessment Tutorial: A General Review on Cerebral Signals' Synchronization Estimation Concepts and Methods

Esmail Seraj

► **To cite this version:**

Esmail Seraj. Cerebral Synchrony Assessment Tutorial: A General Review on Cerebral Signals' Synchronization Estimation Concepts and Methods. [Research Report] Shiraz University; Georgia Institute of Technology, USA. 2018. hal-01376284v4

HAL Id: hal-01376284

<https://hal.science/hal-01376284v4>

Submitted on 26 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Cerebral Synchrony Assessment Tutorial: A General Review on Cerebral Signals' Synchronization Estimation Concepts and Methods¹

by

Esmail Seraj^{†,*} ²

[†]*Signal Processing Center, Department of Computer Science and Engineering and Information Technology, School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran³*

^{*}*School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia, United States⁴*

July 2018⁵

arXiv:1612.04295v2 [q-bio.NC] 6 Jul 2018

¹Version #3.2

²Email: esmaeil.seraj09@gmail.com

³Previous affiliation

⁴Current affiliation

⁵Last update

Abstract

The human brain is ultimately responsible for all thoughts and movements that the body produces. This allows humans to successfully interact with their environment. If the brain is not functioning properly many abilities of human can be damaged. The goal of cerebral signal analysis is to learn about brain function.

The idea that distinct areas of the brain are responsible for specific tasks, the functional segregation, is a key aspect of brain function. Functional integration is an important feature of brain function, it is the concordance of multiple segregated brain areas to produce a unified response. There is an amplified feedback mechanism in the brain called reentry which requires specific timing relations. This specific timing requires neurons within an assembly to synchronize their firing rates. This has led to increased interest and use of phase variables, particularly their synchronization, to measure connectivity in cerebral signals. Herein, we propose a comprehensive review on concepts and methods previously presented for assessing cerebral synchrony, with focus on phase synchronization, as a tool for brain connectivity evaluation.

Keywords: Synchronization, Connectivity, Cerebral Connectivity, Phase synchrony, Phase Estimation, PSD, Correlation, Coherence, Magnitude Squared Coherence, MSC, Higher Order Spectra, HOS, Bispectrum, Multiple MSC, Thomson Multitapers, Imaginary MSC, Phase Lag Index, PLI, WPLI, Parametric Spectral Estimation, Phase Locking Value, PLV, PLS, Mean Phase Coherence, MPC, Wavelet Coherence, WC, WPLV, Empirical Mode Decomposition, EMD, Phase Analysis

Contents

1	Introduction	1
1.1	Cerebral Signals: Early Analysis	2
1.2	Cerebral Signal: Modern Analysis-Cerebral Connectivity	8
1.3	Synchronization	9
1.3.1	Linear Synchronization Methods	9
1.3.2	Nonlinear Synchronization Methods	10
1.3.3	Information-based Techniques	12
1.4	Cerebral Signal: Modern Analysis-Higher order Spectra	12
1.4.1	Moments and Cumulants of order n	13
1.4.2	Bispectrum	13
1.5	Major Obstacles	15
1.5.1	Important Issues Regarding Instantaneous Phase Estimation: Phase Slipping Problem	15
1.5.2	Volume Conduction	15
1.5.3	Presence of a Common Reference	16
1.5.4	Noise Sources and Spurious Synchrony	16
1.5.5	Phase Enslaving	16
1.6	Applications	17
1.7	Summary	17
2	Phase Analysis Methods	18
2.1	Coherence	18
2.1.1	Conventional MSC	18
2.1.2	MSC for Deterministic Evoked Signals	19
2.1.3	Multiple MSC	20
2.1.4	MSC Based on Thomson Multitapers	21
2.1.5	Imaginary Part of Coherence	21
2.1.6	Phase Lag Index	23
2.1.7	Weighted Phase Lag Index	24
2.1.8	MSC Based on Parametric Spectral Estimation Methods	25
2.2	Phase Synchronization	26
2.2.1	Phase Locking Value	26
2.2.2	Mean Phase Coherence	29
2.2.3	Wavelet Coherence	30
2.2.4	Weighted Phase Locking Value	30
2.2.5	Reduced Interference Rihaczek Distribution Phase Synchrony	31
2.2.6	Phase Synchrony Based on Empirical Mode Decomposition	32
A	Which Coordinates?	33

Chapter 1

Introduction

Cerebral signals are among the easiest, most useful and efficient tools for studying brain and its different cognitive states. According to their concept (i.e. embedded information) and also the recording procedure, these signals are called in many different names such as EEG (electroencephalogram), EMG (magnetoencephalogram), ERP (event related potential) or etc.. Analyzing these signals needs some powerful and discriminative features extracted from them. Amplitude is one of these powerful features. Amplitude of Biological signals has been shown that is very informative and is used as main analysis factor in former studies.

In some applications, because of the environment being too much noisy, or presence of other biological contaminants, the amplitude information of captured signals become contaminated and inadequate. As an example of such applications we can refer to ERP signals studies in BCI purposes that the spontaneous EEG and other biological signals such as EOG play the contaminant role and make the amplitude of ERP signals inadequate in information. As an other instance, investigating fetal brain signals needs recordings from maternal abdominal. Also in this application there are some biological contaminants such as fetal and maternal cardiac activity signals (fECG, mECG). These signals are stronger, i.e. 10–100 fold, in amplitude than fetal brain signals. Thus, the amplitude solely can not yield adequate information. In such cases we need to look for new features to use and capture information from new aspects that are less contaminated and noisy.

Phase of this biological signals is an informative useful alternative feature in these cases that has been of interest in past few years. In many recent researches it has been shown that phase has valuable information and can be used as a complement for amplitude information. Thus, using phase of brain signals in cases that amplitude informations are contaminated or inadequate, or the phase itself has adjunct information over amplitude (i.e. brain cognitive response investigations), would be very helpful.

Phase of cerebral signals has been extracted and used through various procedures in past studies such as *Fourier* and *Wavelet* transforms in frequency domain, *Analytic* form of signals and also using *multi-channel* signals and the angles between signal subspaces. Each of these methods are applied on data sets related to a specific application such as brain computer interface purposes, cognitive researches, event related potentials studies and etc. Since for different applications the result of all these various methods will vary, the important confronting problem here is that there is not a comprehensive and reliable comparison between these methods. So we decided to first specify these different methods and extract a well-defined conceptual framework out of them to prevent from being perplexed in using phase analysis methods. After that we will implement some of these methods on specific data sets in specific applications to compare the results and specify the properties of each method in details. We hope that the results of this study will be useful in applications such as ERP signals analysis (i.e. BCI systems studies) and also brain cognitive responses investigations.

To this goal, in Chapter 1, the brain signals' phase analysis methods from two viewpoints are presented: (1) the early analysis and (2) the modern analysis. Different related quantities are defined

and the statistics are presented and some benefits and limitations confronting in both the early and the modern viewpoints are discussed. Finally at the end of this section the related applications are introduced. Phase analysis is of an extended interest in this applications because of their particular characteristics.

In Chapter 2, phase analysis methods are presented and investigated afterward in different fields such as measuring connectivity in distinct brain areas, phase modulations and demodulations and high order statistics. The different approaches in each specific case such as *Coherence*, *Phase Synchronization* and *Desynchronization* and etc. are presented in this Section. The statistics and the motivation of each method are proposed and discussed for each part.

1.1 Cerebral Signals: Early Analysis

The human brain is ultimately responsible for all thought and movement that the body produces. This allows humans to successfully interact with their environment. If the brain is not functioning properly many abilities of human can be damaged. The goal of cerebral signal analysis is to learn about brain function.

Cerebral signals are susceptible to be contaminated by various noise sources such as muscular activities (i.e. blinking, jaw clenching or frowning), artifact generated by the electrical equipments which are near recording areas and also other biological signals (i.e. EMG or EOG). Artifact detection and removal is a necessary step for many types of cerebral signal analysis [1].

Early cerebral signal recording systems were very simple in comparison to the their modern species. The analogue EEG signals recorded at that time were analyzed visually and the qualitative features of them were been investigated usually. Features of the EEG signal were classified as either paroxysmal (transient, bursting) or on-going (background, spontaneous) activity. A simple feature of EEG recordings was measured by counting the number of oscillations in one second.

Because of recordings being too much noisy the quantification and interpretation of the amplitude of a desired frequency component were too difficult, since the peak-to-peak amplitude of the signals changed over time. One approximation used was to create an envelope for the signal by joining together all the peaks/troughs and taking an average of the resulting lines [2]. Another used measure was the duration of time that a recording would spend oscillating in a particular frequency band [3].

Many complex waveform that exist in EEG recording such as the alpha rhythm which was the first EEG waveform reported by Hans Berger [4] (also called as the Berger waveform) and the K-complex which is a transient waveform that occurs during sleep were discovered [5](many more examples can be found in [6]). Also topographic maps of the scalp potentials were used to visually identify spatial patterns in EEG signals, however creating these maps were difficult from analogue recordings.

Often visual analysis was not sufficient to discriminate between components and furthermore, more complex events were very cumbersome with analogue recordings. Analogue-to-digital converters (ADC) are used to create digital EEG recordings. Digital EEG recordings opened new possibilities for analyzing the spectral content of brain signals. Creating topographic maps was made more practical with digital recordings and systems. Using the digital EEGs, computers used interpolation to create detailed contour plots of the scalp potential. However, there are also some ambiguities remaining when interpolation is used, due to projecting a 3D volume onto a 2D surface [7].

Based on the idea of *Fourier analysis*, using the Fourier transform leads to measuring the *power spectral density*(PSD) and so the *mean squared coherence* (MSC) (which employs the PSD) and also extracting significant frequency specific components such as instantaneous phase and amplitude.

1.1.0.1 Power Spectral Density

Let $x(t)$ be a stochastic stationary process, The Fourier transform of $x(t)$ is shown by $X(f)$ and is determined as below:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (1.1)$$

Then the power spectral density $PSD_x(f)$ of $x(t)$ is calculated as follow:

$$PSD_x(f) = \sum_{i=-\infty}^{\infty} |X_i(f)|^2 \quad (1.2)$$

The PSD can also be calculated between two distinct signals and it is called as cross-spectrum function. Cross-spectrum is a complex function that its amplitude and phase show the correlation between two signals. For instance, the cross-spectrum of two sinusoidal signals of the same frequency has a sharp peak in that frequency. The phase of the cross-spectrum at that frequency equals the phase difference between the records. If two signals share a sinusoidal activity, but each also contains other, unshared activity, their cross-spectrum has a peak at the shared frequency only.

1.1.0.2 Correlation Function

Another measure availed through digital signal recordings is the correlation function. Auto-correlation function $R_x(\tau)$ of $x(t)$ is measured as below:

$$R_x(\tau) = x(\tau)x(-\tau) = \int_{-\infty}^{\infty} x(t)x(t + \tau)dt \quad (1.3)$$

The Fourier transform of the Auto-correlation function does indeed equal to the PSD of our stochastic stationary signal $x(t)$.

$$PSD_x(f) = \int_{-\infty}^{+\infty} R_x(\tau)e^{j2\pi f\tau} d\tau \quad (1.4)$$

The proof is quite straightforward, by mean of Fourier transform of the correlation function $R_x(\tau)$ we have:

$$\begin{aligned} f\{R_x(\tau)\} &= \int_{-\infty}^{\infty} R(\tau)e^{-j2\pi f\tau} d\tau \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t)x(t + \tau)e^{-j2\pi f\tau} dt d\tau \\ &= \int_{-\infty}^{\infty} x(t) \int_{-\infty}^{\infty} x(t + \tau)e^{-j2\pi f\tau} d\tau dt \\ &= X(f) \int_{-\infty}^{\infty} x(t)e^{j2\pi ft} dt \\ &= X(f)X^*(f) = |X(f)|^2 \end{aligned}$$

1.1.0.3 Phase & Amplitude Extraction

Two other quantities availed through digitalized early cerebral signal analysis are the instantaneous phase and amplitude. There are two conceptually distinct set of techniques for the purpose of capturing the amplitude and phase of a signal: (1) linear techniques and (2) nonlinear techniques. In one hand, the linear techniques assume a constant amplitude and phase within the estimation

window [8]. As a common example, Fourier transform represents the stationary process $x(t)$ in the frequency space as a phasor $X(f)$:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt = A(f) \cos \theta(f) + jA(f) \sin \theta(f) = A(f)e^{j\theta(f)}$$

where $A(f)$ and $\theta(f)$ are the amplitude and phase of signal $x(t)$ within the estimation window respectively.

In the other hand, nonlinear techniques try to measure the time-dependent instantaneous amplitude and phase which are illustrating the moment-to-moment change in signal. Instantaneous amplitude and phase are usually obtained using either the *Hilbert* or *Wavelet* transforms which has been illustrated that both produce similar results [9]. There are other recently proposed methods for phase extraction purpose which are represented here.

1.1.0.3.1 Hilbert Transform The Hilbert transform of a real valued process $x(t)$ is defined as:

$$x_h(t) = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (1.5)$$

where $x_h(t) = H\{x(t)\}$ represents the Hilbert transform of process $x(t)$ and the integral is taken using the Cauchy principal value due to the potential singularity at $t = \tau$. The Hilbert transform is the convolution of the signal $x(t)$ with the function $\frac{1}{\pi t}$ which means the PSD of $x_h(t)$ is equal to the product of the PSD of $x(t)$ and the PSD of $\frac{1}{\pi t}$. Besides, the Fourier transform of the function $\frac{1}{\pi t}$ is equal to:

$$F\left\{\frac{1}{\pi t}\right\} = -i \operatorname{sgn}(f)$$

which states that the Hilbert transform is simply a $\frac{\pi}{2}$ shift in the phase of the original signal. Thus using the analytic extension of a signal as presented below does not change the PSD of that signal [10]:

$$Z(t) = x(t) + jx_h(t) \quad (1.6)$$

Now, the instantaneous amplitude and phase of $x(t)$ can be defined as the magnitude and argument of the analytic extension as follow:

$$Z(t) = A(t)e^{j\phi(t)} \quad (1.7)$$

where $A(t)$ is the instantaneous amplitude and $\phi(t)$ is the instantaneous phase which can be calculated as below:

$$\phi(t) = \arctan\left(\frac{x_h(t)}{x(t)}\right) \quad (1.8)$$

To have a physically meaningful interpretation of the instantaneous phase and amplitude, they have to be extracted from a narrow-band signal or in a specific frequency of interest, so a filtering pre-step before applying Hilbert transform is needed [11].

1.1.0.3.2 Wavelet Transform As mentioned before, another useful procedure for measuring the instantaneous amplitude and phase of a signal is through time-frequency transformations such as Wavelet transform. By mean of the standard Morlet-Wavelet, the Wavelet coefficients of signal $x(t)$ can be determined as follow [12]:

$$W_x(t, f) = \int_{-\infty}^{\infty} x(u)\psi_{t,f}^*(u) du \quad (1.9)$$

where the $\psi_{t,f}(u)$ is a modulated Gaussian function (with center of t and variance of σ) by a sinusoidal wave with frequency of f :

$$\psi_{t,f}(u) = \sqrt{f} e^{j2\pi f(u-t)} e^{-\frac{(u-t)^2}{2\sigma^2}}$$

Then two important parameters obtained from Wavelet transform are *spectrogram* and *Instantaneous Phase* [12]. We have:

$$Spectrogram = |W_x(t, f)|^2 \quad (1.10)$$

and

$$\exp(j\phi_x(t, f)) = \frac{W_x(t, f)}{|W_x(t, f)|} \quad (1.11)$$

which in latter the $\phi_x(t, f)$ represents the instantaneous phase.

1.1.0.3.3 Phase Extraction Based on Complex Energy Density Function “Rihaczek derived the signal energy distribution in time and frequency by application of the complex signal notation” [71].

Considering two complex signals $x(t)$ and $y(t)$ which are at similar frequency, the total complex energy can be defined as:

$$CE(t, f) = \int_{-\infty}^{\infty} x(t)y^*(t)dt \quad (1.12)$$

Rihaczek utilized this idea to compute the interaction energy of a signal within some frequency band. This leads to the *Complex Energy Density Function* (CEDF), stated as below [72]:

$$CEDF(t, f) = \frac{1}{\sqrt{2\pi}}x(t)X^*(f)\exp(-j2\pi ft) \quad (1.13)$$

CEDF measures the complex energy of a signal around time t and frequency f . As stated in [71], CEDF provides a better understanding of phase-modulated signals. Now, the time-varying phase for signal $x(t)$ in time-frequency plane is defined as below:

$$\phi_x(t, f) = \arg\left(\frac{CEDF(t, f)}{|CEDF(t, f)|}\right) \quad (1.14)$$

1.1.0.3.4 Phase Extraction Based on Empirical Mode Decomposition The *Empirical Mode Decomposition* EMD technique was first proposed by Huang in 1998 [76], and is based on the assumption that any signal consists of different simple intrinsic independent modes of oscillation. Each of these modes will have the same number of extrema and zero-crossings and there is only one extremum between successive zero-crossings [74]. Here we present the statistics given in [76] for EMD algorithm with respect to the signal $x(t)$:

- Find all the local *extrema*. Then connect all the local maxima together and also all the local minima by a cubic spline as the *upper and lower envelope* respectively, ($e_{max}(t)$ and $e_{min}(t)$).
- Extract the *detail* as the difference between the signal and the mean of the upper and lower envelope values : $d(t) = x(t) - \frac{e_{max}(t)+e_{min}(t)}{2}$.
- If $d(t)$ is not an IMF, treat it as original signal and repeat first two steps until $d(t)$ becomes an IMF, then $d(t) = I_1(t)$ is the first IMF.
- After getting the first IMF, remove it from the original signal and obtain the residual r_1 : $x(t) - I_1(t) = r_1(t)$.
- Treat the $r_1(t)$ as the original signal and repeat the above steps to obtain the second IMF, $I_2(t)$.
- Repeat the process n times. The decomposition process can be stopped when no more IMF components can be extracted from the last residual.

Thus, the EMD algorithm decomposes the signal $x(t)$ into n empirical modes as:

$$x(t) = \sum_{i=1}^n I_i(t) + r_n(t) \quad (1.15)$$

The IMFs $I_1(t), \dots, I_n(t)$, include different frequency bands ranging from high to low [74]. Now, after decomposing signal $x(t)$ into its relative IMFs, for the purpose of phase extraction, Hilbert transform can be applied to the IMFs and produce the instantaneous phase sequences.

The main difference between this method and other phase extraction method is that, here, phase sequences are extracted from IMFs and there is no need for band-pass filtering the signal anymore. Also EMD based methods have some other merits which are discussed in next chapter while introducing EMD based phase analysis method for assessing synchronization in brain.

1.1.0.4 Coherence

The idea that distinct areas of the brain are responsible for specific tasks, the functional segregation, is a key aspect of brain function. Functional integration is another important feature of brain function, it is the concordance of multiple segregated brain areas to produce a unified response [13], [14].

To measure functional integration in brain, different procedures were introduced to quantify functional connectivity, a statistical dependence between different cerebral signals captured from different brain regions. The first famous measure of functional connectivity was coherence.

Coherence determines the correlation between the signals at specific frequencies [15], [16]. Coherence in a simple definition is a measure of synchronization and correlation between two random process, stochastic, wide-sense stationary (*WSS*) signals that is computable through the coherence function. Coherence is widely used to study dependency and relationship between different brain regions particularly during a specific cognitive task or while experiencing a specific stimuli [17].

In general Coherence function is a complex valued number with both an amplitude and phase that are used to measure phase synchronization in signals. The main advantage of coherence over other correlation calculating methods (i.e. in [18]) is that in fact coherence gives the correlation between two signals as a function of frequency that provides the possibility of studying spatial correlation in different frequency bands [17].

According to the different characteristics of signals (i.e. stochastic or deterministic) and also different procedures presented for determining PSDs in literature, several approaches for MSC calculation exist. Here we represent the most widely used approaches and in next Chapter after reintroducing the conventional MSC, The rest of the approaches will be discussed.

1.1.0.4.1 MSC Using FFT In [17] a method for measuring coherence function is presented. The method is based on weighted windowing of the Fourier transform of signals. Let $x(t)$ and $y(t)$ be two random process, zero mean, wide-sense stationary and ergodic signals with length l . A summary of the method in [17] is given below:

- Divide the signals into N equal parts with T samples (different part can be overlapped or disjoint)
- Multiply the samples of each ensemble in a weighting function (i.e. cosine)
- Take *FFT* of each weighted ensemble
- Measure power spectral densities of signals (*PSD*)
- Finally, measure the coherence function (or as said “*Mean Squared Coherence (MSC)*”)

If $X_i(f)$ and $Y_i(f)$ be the Fourier transform of i -th ensemble, the PSDs are calculated as follow:

$$PSD_{xx}(f) = \frac{1}{NT} \sum_{i=1}^N |X_i(f)|^2 \quad (1.16)$$

$$PSD_{yy}(f) = \frac{1}{NT} \sum_{i=1}^N |Y_i(f)|^2 \quad (1.17)$$

$$PSD_{xy}(f) = \frac{1}{NT} \sum_{i=1}^N X_i(f)Y_i^*(f) \quad (1.18)$$

With the ergodicity of signals being neglected we have:

$$PSD_{xx}(f) = E\{|X_i(f)|^2\} \quad (1.19)$$

$$PSD_{yy}(f) = E\{|Y_i(f)|^2\} \quad (1.20)$$

$$PSD_{xy}(f) = E\{X_i(f)Y_i^*(f)\} \quad (1.21)$$

where $E\{\cdot\}$ illustrates the mathematical expectation. Using these PSDs we compute the mean squared coherence MSC as:

$$|MSC(f)|^2 = \frac{|PSD_{xy}(f)|^2}{PSD_{xx}(f)PSD_{yy}(f)} = \frac{\frac{1}{N} \sum_{i=1}^N |(PSD_{xy}(f))_i|^2}{PSD_{xx}(f)PSD_{yy}(f)} \quad (1.22)$$

Since $PSD_{xy}(f)$ is complex, the *Cartesian* or *Polar* coordinates can be used for averaging. We discussed This issue in Appendix A.

1.1.0.4.2 MSC Using Correlation Functions Another method for obtaining complex coherence is presented in [20] which computes the PSDs using auto and cross correlations. In this method after dividing the signals into N separate parts, first the auto/cross correlations R_{xx} , R_{yy} and R_{xy} are calculated and then the PSDs are measured by taking Fourier transform of these correlation functions:

$$R_{xx}(\tau) = E\{x(t)x(t+\tau)\} \Rightarrow PSD_{xx} = \int_{-\infty}^{+\infty} R_{xx}(\tau)e^{j2\pi f\tau} d\tau \quad (1.23)$$

$$R_{yy}(\tau) = E\{y(t)y(t+\tau)\} \Rightarrow PSD_{yy} = \int_{-\infty}^{+\infty} R_{yy}(\tau)e^{j2\pi f\tau} d\tau \quad (1.24)$$

$$R_{xy}(\tau) = E\{x(t)y(t+\tau)\} \Rightarrow PSD_{xy} = \int_{-\infty}^{+\infty} R_{xy}(\tau)e^{j2\pi f\tau} d\tau \quad (1.25)$$

The equations (1.23), (1.24) and (1.25) can also be illustrated using inner product by the complex sinusoidal mother function as follow:

$$PSD_{xx} = \langle R_{xx}, e^{j2\pi ft} \rangle \quad (1.26)$$

$$PSD_{yy} = \langle R_{yy}, e^{j2\pi ft} \rangle \quad (1.27)$$

$$PSD_{xy} = \langle R_{xy}, e^{j2\pi ft} \rangle \quad (1.28)$$

This method also takes all the assumptions considered in former method. Here, as well, if the ergodicity assumption is neglected thus the expectation term in correlation calculation could be changed by time averaging. Finally the MSC is obtained as below:

$$MSC(f) = \frac{PSD_{xy}(f)}{\sqrt{PSD_{xx}(f)PSD_{yy}(f)}} = \frac{F\{R_{xy}(f)\}}{\sqrt{F\{R_{xx}(f)\}F\{R_{yy}(f)\}}} \quad (1.29)$$

where $F\{\cdot\}$ denotes the Fourier transform.

MSC varies between 0 to 1 ($0 < MSC < 1$). The more two signals are related and synchronized, the greater the MSC value will become. In fact a MSC value close to 1 means the activity between corresponding signals follows a linear transformation in that frequency of interest and a MSC near zero shows two non-related signals in that frequency.

As discussed before, cross-spectrum is a complex function that its amplitude and phase show the correlation between two signals. For instance, the cross-spectrum of two sinusoidal signals of the same frequency has a sharp peak in that frequency. The phase of the cross-spectrum at that frequency equals the phase difference between the records. If two signals share a sinusoidal activity, but each also contains other, unshared activity, their cross-spectrum has a peak at the shared frequency only. The MSC has the same interpretation to the cross-spectrum and concluded from the discussion above, it is said that the phase of MSC is meaningful and equals the phase difference between two corresponding signals, if the amplitude in that frequency is considerable. The large value of coherence amplitude in a phase of 0 degree indicates the large positive correlation and in a phase of 180 degree it's the vice versa and the the rest of the phase values are interpreted similarly due to the amplitude in that frequency [18].

The discriminant factor between methods presented in [17] and [20] is the procedure of measuring power spectral densities of signals. As reported in [20], the PSD can be estimated using non-parametric methods such as Blackman–Tukey method, weighted overlapped segment averaging (*WOSA*), also known as Welch's method, Lag-resaping method or parametric ones, as discussed later in next Chapter.

1.2 Cerebral Signal: Modern Analysis-Cerebral Connectivity

Brain connectivity depicts patterns of links in the brain. Most of the brain functions, for instance the ones involved in learning, memory, behavior adaptation to stimuli, emotions, as well as pathological processes in some brain disorders like epilepsy, autism or schizophrenia, are based on interactions between neuronal assemblies distributed within and across distinct cerebral regions [21]. For instance, it has been shown that segregated areas may activate in response to a particular cognitive task. Neural assemblies are the basic concept of current scientific model for how cerebral functional integration is achieved. Integration of cerebral areas can be measured by assessing brain connectivity [21], [22], [23]. Brain connectivity can be separated into three different concepts: (1) *structural* (or *neuroanatomical*) connectivity, which denotes anatomical links, (2) *functional* connectivity, which rely on statistical dependencies between signals from different brain areas and (3) *effective* connectivity, which introduces causal interactions among these signals [21], [23], [24].

There are reasons discussed in [25] which make *structural connectivity* difficult to define intrinsically. *functional connectivity* is defined as the temporal correlation among the activity of different neural assemblies whereas the direct or indirect effect that one neural system exerts over another is defined as *effective connectivity* [26], [27].

Functional and effective connectivity techniques are intensely dependent on measuring the correspondence of neural signals over time. Thus cerebral signal recording techniques such as EEG and MEG with good temporal resolution, are optimal for calculating such connectivity. Most frequently used measures for connectivity are: correlation, coherence, mutual information, transfer entropy, generalized synchronization [28], continuity measure [29], synchronization likelihood [30] and phase synchronization [31].

There is an amplified feedback mechanism called reentry which requires specific timing relations. This specific timing requires neurons within an assembly to synchronize their firing rates. This has led to increased interest and use of phase variables, particularly their synchronization, to measure

connectivity in cerebral signals [16]. Moreover, our focus in this work is based on signal phase and its analysis methods. Thus, we will introduce the phase synchrony as a subset of modern methods for cerebral signal analysis and discuss various approaches presented to achieve this quantity in next section.

The next section is dedicated to introducing synchronization and its different aspects. Each case is introduced and its characteristics are represented briefly. Finally the phase synchronization as one of the major investigated fields in this thesis and its various approaches are presented.

1.3 Synchronization

Connectivity measures are widely used to determine the level of synchronization between distinct brain regions using cerebral recordings. As discussed before, since *structural connectivity* is difficult to define inherently, only the two remaining connectivity measures (*functional connectivity* and *effective connectivity*) are used for determining synchronization in brain.

As discussed in [23], methods to measure *effective connectivity* can be subdivided into two main categories: (1) model-based and (2) data-driven techniques. The former case assumes theoretical models that describe how brain areas interact and influence each other, whereas the latter case does not assume any specific underlying model or prior knowledge concerning spatial or temporal relationships. Granger-causality (GC), directed transfer function (DTF), directed coherence (DC) and its extension partial directed coherence (PDC) are the most common data-driven connectivity techniques. It is notable that DTF, DC and PDC are developed out of the GC method. The PDC technique will be presented in next chapter as an extension of coherence methods.

The *functional connectivity* methods are subdivided into three main categories: (1) linear, (2) nonlinear and (3) information-based techniques [23]. Each of these categories and the corresponding most common approaches are introduced in next subsections.

1.3.1 Linear Synchronization Methods

1.3.1.1 Cross Correlation

Cross-correlation and MSC are the most commonly used linear synchronization methods. As we discussed before (see Sections 1.2.1.2 and 1.2.1.4), assuming two discrete time processes x_n and y_n where $n = 1, \dots, N$, then the cross-correlation function R_{xy} is defined as:

$$R_{xy}(\tau) = \frac{1}{N - \tau} \sum_{n=1}^{N-\tau} \left(\frac{x_n - \mu_x}{\sigma_x} \right) \left(\frac{y_{n+\tau} - \mu_y}{\sigma_y} \right) \quad (1.30)$$

where μ and σ denote mean and variance, respectively, while τ is the time lag.

1.3.1.2 MSC

Another linear technique for synchronization measurements is the MSC or simply coherence. MSC is defined precisely in Section (1.1.0.4). To remind what the basic idea is, the main formula to calculate MSC is given below:

$$|MSC(f)|^2 = \frac{\frac{1}{N} \sum_{i=1}^N |(PSD_{xy}(f))_i|^2}{PSD_{xx}(f)PSD_{yy}(f)} \quad (1.31)$$

which is averaged over trials.

Although MSC is a successful measure of functional integration, it has several limitations [16]:

- Stationarity: Because of using FFT, existence of a PSD depends on the stationarity of a process. As reported in [32], there is evidence that EEG recordings longer than a particular period of time are non-stationary.
- Linearity: Coherence is based on linear correlation in the process. Thus, if information is being transferred in a nonlinear manner, it may not be detected.
- Direction: When there is synchronization between two signals, coherence does not clarify which signal is driving the flow of information.

1.3.2 Nonlinear Synchronization Methods

Many crucial neural processes have nonlinear characteristics. Due to the lack of linear synchronization methods to detect this nonlinear connectivity, basic motivation for developing nonlinear methods were not put on outperforming linear methods but rather providing complementary information under certain assumptions [23].

Within the field of neuroscience, synchronization is commonly illustrated in the concept of phase synchrony. There are various approaches to obtain this quantity based on different characterization procedures. The main ideas are described briefly in next subsection and more comprehensively in next chapter. Another nonlinear connectivity measure termed as *Generalized synchronization* is also introduced very briefly.

1.3.2.1 Phase Synchronization

Phase synchronization (PS) is defined as temporal adjustment of rhythms of a pair of signals in a frequency band of interest, whereas the amplitudes can remain uncorrelated [33]. PS also can be determined through successive frequency bands to provide a complete measure of synchrony between two signals. Phase locking-value (PLV) is the most commonly used measure of PS. In order to detect periods of phase synchrony (i.e. PLV) in cerebral signals, there are four main steps summarized as follow [34], [35], [36], [37], [38]:

- The first step is to calculate a sequence consisting instantaneous frequency specific phase values within a temporal window.
- The second step simply is to determine the instantaneous phase-differences (in frequency f and time t) from the corresponding phase sequences captured in first step.
- Quantifying the local stability of this phase-differences across trials of cerebral signals.
- Determining the degree of statistical significance of each quantity obtained in step three.

The first step can be performed through one of the two common signal phase extraction methods: (1) Hilbert Transform or (2) Wavelet transform (see Section 1.2.1.3). The second step actually depends on the third step. The quantification in step three can be performed through various statistical dependence parameters such as mean phase difference, circular variance, standard deviation, Shannon entropy, or mutual information [33]. Depending on which of these procedures to be selected, the process in second step might be required or not. If the phase differences have to be found, there are different ways to this goal depending on the method used for phase extraction. This might differ between simply subtracting the corresponding time-dependent phases captured through *Hilbert Transform* or using description below due to phase extraction through *Wavelet Transform*:

$$\exp(j(\phi_y(t, f) - \phi_x(t, f))) = \frac{W_x(t, f)W_y^*(t, f)}{|W_x(t, f)||W_y(t, f)|} \quad (1.32)$$

where the term $\phi_y(t, f) - \phi_x(t, f)$ represents the phase-difference between processes y and x . The main conventional approaches to assess the synchrony between neural signals are distinct in their technique to quantify the captured instantaneous phase-differences. Finally, for the last step commonly surrogate data are utilized. There are different techniques presented for generating surrogate data in literature such as shifting or scrambling the original series [35], [37]. The degree of statistical significance of the PLVs is then determined by comparing them to values obtained from surrogate data.

Generally speaking, for two neural signals $x(t)$ and $y(t)$ and their corresponding phases $\phi_x(t)$ and $\phi_y(t)$, phase synchrony in its most general form is defined as:

$$n\phi_x(t) - m\phi_y(t) = \text{const.} \quad (1.33)$$

where n and m are integers indicating the ratios of possible frequency locking [31]. In what follows we assume $n = m = 1$ for simplicity. Because of problems such as *Volume Conduction* (introduced later in current chapter) or *background spontaneous brain signals* the true synchrony is buried in a considerable background noise [38]. In this case, as in [31], this condition of phase locking can be replaced by the weaker condition as stated below:

$$|n\phi_x(t) - m\phi_y(t)| < \text{const.} \quad (1.34)$$

As discussed in [33], such approaches to phase synchrony impose two main limitations as presented below:

- As mentioned, only synchrony between pairs of signals can be directly studied. This extremely increases the computational cost. For instance, when there is n signals, $d = \frac{n^2-n}{2}$ pairs of signal exist.
- As illustrated in [33], synchronizations with frequency non-stationarity can be observed in brain signals (frequency of synchronization alters continuously through time, although the phase-difference remains stable). Thus, defining phase synchronization in this manner will prevent such phase synchronizations from being detected.
- It has been demonstrated in [40] that synchrony-measurement methods based on this definition may be insensitive to very short periods of phase synchronization (see also [41]).

The relation between coherence and phase synchronization is investigated in [60] and according to the reported results there is a very close relationship between these concepts. In many studies it has been reported that a complete phase synchronization is manifested by highly coherent phases and correlated amplitudes. In fact, both the phase synchronization and the coherence analysis are looking for periods of phase synchrony in various frequency bands but there is some differences between their sensitivity to detecting them. With focusing on Signals phase, the coherence analysis has some objections (see Section 1.3.1.2), however, in [60] it is reported that phase synchronization analysis is not superior than coherence considerably.

1.3.2.2 Generalized synchronization

After the successful application of PS in EEG analysis, the concept of Generalized synchronization (GS) was introduced [42]. GS approaches were proposed to investigate the dependencies between nonlinear signals without any knowledge about the governing equations [22]. GS represents how much neighborhoods of one chaotic attractor can be mapped onto the other [23]. Although this mapping is prone to stationarity deficits, it is considered to be a robust way of evaluating the GS [42]. Generating such attractors needs delay vectors to be constructed from time series using a procedure known as time-delay embedding [43].

1.3.3 Information-based Techniques

In probability theory and information theory, Mutual Information (MI) measures the mutual dependencies between two random-variable processes by quantifying the amount of information gained about one signal from measuring the other. If two signals have a statistical dependence, observing one signal gains information about the other and reduces the entropy because of the knowledge of other signal. The benefit of using information-based techniques is that these techniques are sensitive to both linear and nonlinear statistical dependencies between signals [16], [23].

MI analysis can be used to assess functional connectivity between cerebral signals. For this purpose a measure of entropy is needed [44]. Defining the entropy for a signal $x(t)$ as:

$$H(x) = - \int_x P(x) \log(P(x)) \quad (1.35)$$

we have the mutual information between signals $x(t)$ and $y(t)$ determined as:

$$I(x, y) = \int_y \int_x P(x, y) \log\left(\frac{P(x, y)}{P(x)P(y)}\right) dx dy \quad (1.36)$$

where $P(x, y)$ is the joint probability density function of $x(t)$ and $y(t)$, and $P(x)$ and $P(y)$ are the marginal probability density functions of $x(t)$ and $y(t)$ respectively [44]. MI has been used for evaluating functional connectivity in EEG recordings (see [45] and [46]). Also MI and entropy based features have been used previously for CAP detection in sleep EEG recordings [92].

1.4 Cerebral Signal: Modern Analysis-Higher order Spectra

The previous introduced statistical tools utilize first and second order statistics to extract information from random signals. However, the main problem confronts in the presence of nonlinearity in systems where the first and second order statistics are unable to adequately analyze many signals [78]. Presenting *High-order Spectra (Statistics)* (HOS) is mainly motivated to overcome this problem and plays an important role in digital signal processing. HOS methods are very useful in problems where non-Gaussian, non-minimum phase, phase coupling or nonlinear behavior and robustness to additive noise are important [78]. They are also used in detection and classification in communication and pattern recognition applications.

“In general, there are three motivations behind the use of HOS in signal processing which can be considered as its property over lower order spectra: (1) suppressing Gaussian noise of unknown mean and variance; (2) detecting and characterize nonlinearities in the data and (3) reconstructing and preserving the phase as well as the magnitude response of signals or systems” [79]. Also the applications of HOS on biomedical signals are based on these properties. Particularly in this contribution we focus on the latter property of HOS. In fact, measuring Power spectrum causes phase relations and signals phase informations to be suppressed, while HOS, based on the ability of cumulant spectra to preserve the Fourier-phase of signals, contains such informations [80], [81] and [82].

Higher order spectra are functions of two or more component frequencies while the power spectrum and lower order spectra are functions of a single frequency [78]. The *bispectrum* and *bicoherence* and their related extensions (i.e. auto/cross bispectrum and bicoherence) are the third order spectra and also the most used HOS quantities as they are easiest computationally. The other less common HOS quantities than the 3-rd orders are the (auto/cross) *trispectrum* and *tricoherence*. First, we need to review approaches for calculating moments and cumulants of order n as high order spectra quantities' computation is based on them.

1.4.1 Moments and Cumulants of order n

For a stationary discrete time random process $X(k)$, the moments of order n are given by:

$$m_n(\tau_1, \tau_2, \dots, \tau_{n-1}) = E\{X(k)X(k + \tau_1)\dots X(k + \tau_{n-1})\} \quad (1.37)$$

where $E\{\cdot\}$ denotes expectation.

The $n - th$ order cumulants are functions of the moments of order up to n , for example the first order cumulants can be obtained as below:

$$c_1 = m_1 = E\{X(k)\} \quad (1.38)$$

which is the *mean*. Consequently, the second order cumulants are given by:

$$c_2(\tau_1) = m_2(\tau_1) - (m_1)^2 \quad (1.39)$$

which is the *covariance*. Also for the third order cumulants we have:

$$c_3(\tau_1, \tau_2) = m_3(\tau_1, \tau_2) - (m_1)[m_2(\tau_1) + m_2(\tau_2) + m_2(\tau_2 - \tau_1)] + 2(m_1)^3 \quad (1.40)$$

where $m_3(\tau_1, \tau_2)$ in the third order moment sequence. Following this manner, the $n - th$ order cumulants can be obtained. The general relations between cumulants and moments are given in [81].

1.4.2 Bispectrum

1.4.2.1 Auto-bispectrum

The bispectrum of a signal $x(t)$ can be defined in two ways. First is *the triple product of the DFT's* [81]:

$$B_{xxx}(f_1, f_2) = \lim_{T \rightarrow \infty} \frac{1}{T} E\{X_T(f_1)X_T(f_2)X_T^*(f_1 + f_2)\} \quad (1.41)$$

The bispectrum is complex. It means that it contains Fourier magnitude and phase information [83]. With the above definition we have:

$$B_{xxx}(f_1, f_2) = |B_{xxx}(f_1, f_2)| \exp(j\phi(f_1, f_2)) \quad (1.42)$$

where the magnitude of bispectrum obtains as:

$$|B_{xxx}(f_1, f_2)| = |X(f_1)||X(f_2)||X^*(f_1 + f_2)| \quad (1.43)$$

and the phase of bispectrum (or biphas [83]) is:

$$\phi(f_1, f_2) = \phi(f_1) + \phi(f_2) - \phi(f_1 + f_2) \quad (1.44)$$

Second definition of bispectrum is *the double DFT of the 3rd order cumulant*. Two methods for calculating bispectrum based on this definition in real data are presented later in current section.

The auto-bispectrum is a function of two frequency components, f_1 and f_2 which would have a small value if the biphas varies over the different realizations and conversely, would have a large value if this phase does not vary, which is an indication of quadratic coupling between f_1 and f_2 [84].

1.4.2.2 Cross-bispectrum

Following the definition of the auto-bispectrum, the cross-bispectrum is used to determine the quadratic coupling between frequency components in two different signals. Considering two signals with Fourier transforms as $X(f)$ and $Y(f)$, the cross-bispectrum is defined by [84]:

$$B_{xy}(f_1, f_2) = \lim_{T \rightarrow \infty} \frac{1}{T} E\{X_T(f_1)X_T(f_2)Y_T^*(f_1 + f_2)\} \quad (1.45)$$

Through this definition, the coupling level between two frequency components in $X(f)$ namely f_1 and f_2 and their algebraic sum in $Y(f)$ is determined which is based on the phase relation between these components in the different realizations. A large value of the cross-bispectrum indicates high *Quadratic Phase Coupling* (QPC) between $X(f)$ and $Y(f)$. In fact QPC normally is obtained through the normalized extension of bispectrum known as bicoherence. Bicoherence as a tool for phase analysis is presented within next Chapter [84].

1.4.2.3 Methods for Calculating Bispectrum

There are two main methods for calculating bispectrum through its definition based on 3rd order cumulants: (1) indirect method and (2) direct method.

1.4.2.3.1 Indirect Method Let $x(k), k = 1, \dots, L$ be the available discrete signal. First, we have to segment the data into N parts each with M samples. Now $x_i(k), k = 1, \dots, M$ is the i -th segment. Then, After subtracting the mean of each segment the 3rd order moment of each segment is computed as below:

$$m_3^{x_i}(\tau_1, \tau_2) = \frac{1}{M} \sum_{l=l_1}^{l_2} X_i(l)X_i(l + \tau_1)X_i(l + \tau_2) \quad (1.46)$$

where $l_1 = \max(0, -\tau_1, -\tau_2)$ and $l_2 = \min(M - 1, M - 2)$.

Since we made each segment zero-mean, its third-order moments and cumulants are equal. Thus, now the average cumulants have to be calculated:

$$c_3^x(\tau_1, \tau_2) = \frac{1}{N} \sum_{i=1}^N m_3^{x_i}(\tau_1, \tau_2) \quad (1.47)$$

Finally, the third order spectrum (bispectrum) can be estimated as [79]:

$$B_{xxx}(f_1, f_2) = \sum_{\tau_1=-L}^L \sum_{\tau_2=-L}^L c_3^x(\tau_1, \tau_2) \exp(-j(f_1\tau_1 + f_2\tau_2))\omega(\tau_1, \tau_2) \quad (1.48)$$

where $L < M - 1$ and $\omega(\tau_1, \tau_2)$ is a two-dimensional window, introduced to smooth out edge effects [79]. A complete description of appropriate windows that can be used and their properties can be found in [81].

1.4.2.3.2 Direct Method Regarding to assumed signal $x(k)$ in previous method, again in this method we have to segment the data and zero-mean each segment similar to previous method. After these steps, the Fourier transform of each segment based on M points have to be computed:

$$F_{x_i}(k) = \sum_{l=0}^{M-1} x_i(l) \exp(-j\frac{2\pi}{M}lk) \quad (1.49)$$

where $k = 0, 1, \dots, M - 1$. Now the third order spectrum (bispectrum) of each segment is obtained as:

$$C_3^{x_i}(k_1, k_2) = \frac{1}{M} F_{x_i}(k_1) F_{x_i}(k_2) F_{x_i}^*(k_1 + k_2) \quad (1.50)$$

which is similar to the first definition of bispectrum. The $C_3^{x_i}(k_1, k_2)$ need to be computed only in the triangular region $0 \leq k_2 \leq k_1, K - 1 + k_2 < M/2$. Also a smoothing window for reducing the variance can be performed around each frequency. With these we have:

$$\hat{C}_3^{x_i}(k_1, k_2) = \frac{1}{W^2} \sum_{l_1=-W/2}^{(W/2)-1} \sum_{l_2=-W/2}^{(W/2)-1} C_3^{x_i}(k_1 + l_1, k_2 + l_2) \quad (1.51)$$

where the smoothing window is of size $(W \times W)$. Finally, the third order spectrum (bispectrum) is given as [79]:

$$B_{xxx}(f_1, f_2) = \frac{1}{N} \sum_{i=1}^N \hat{C}_3^{x_i}(f_1, f_2) \quad (1.52)$$

Both the direct and the indirect methods produce asymptotically unbiased and consistent bispectrum estimates [79].

1.5 Major Obstacles

There are some main obstacles confronting through modern cerebral signal analysis methods, particularly when measuring connectivity in brain. Each of them are introduced below and discussed through next chapter within each represented method.

1.5.1 Important Issues Regarding Instantaneous Phase Estimation: Phase Slipping Problem

Phase Slipping is a notorious problem common in phase estimations from cerebral signals which makes the phase sequence to contain fake (unrelated to brain activity) jumps during low foreground (narrow-band filtered signal) SNRs. This problem is rigorously studied in [86, 85]. In [86], a Monte Carlo based statistical framework is proposed for EEG phase and frequency estimation to overcome the problems associated to previous deterministic approaches (introduced in section 1.1.0.3). Also, in [85], despite the basic illustration of this problem, a straightforward approach for EEG phase estimation based on findings in [86] is proposed and is evaluated in a real BCI problem. To help other researches use the findings of these studies, a comprehensive cerebral signal phase analysis toolbox is generated by the authors of [86] and [85]. Please refer to [87] for a detailed description of the provided toolbox

1.5.2 Volume Conduction

One of the most important obstacles confronting in investigating correlation and synchronization between different brain regions using EEG/MEG recordings is a phenomenon known as Volume Conduction (VC). The activity of a particular brain area will be observed in some adjacent electrodes because of conduction in brain volume which produces spurious synchrony. This substantial problem provides false results and shows incorrect related brain regions that can not be neglected. The problem of VC is especially large for scalp EEG and MEG data, because of their low spatial resolution. The difference between estimated connectivity directly from neural sources and the corresponding estimates from scalp recordings is an effect of VC [16].

Several recent methods are presented to overcome this substantial problem and as some of them reported, they obtained successful results. Many of these methods are represented in Section 2.

1.5.3 Presence of a Common Reference

EEG is a bipolar signal recorded from the scalp while MEG is a unipolar record. In a unipolar record, the recording channels are reference free whereas in a bipolar record, channels are constructed by subtracting two unipolar signals captured from two distinct scalp electrodes (i.e. a reference and an arbitrary electrode).

As reported in [47], the confronting problem here is that bipolar recordings impose a distortion in observed synchrony values, such that in general, they destroy the intended physical interpretation of phase synchrony. It is notable to emphasize that the problem of a common reference is an obstacle when using EEG (and not MEG) recording due to their intrinsic bipolarity. The limitations due to this problem in coherence measurements [48], [49], and phase synchrony [47], are widely investigated previously. It is reported that any contamination in the reference channel will significantly affect the coherence measurements [48], [49]. As mentioned before, one of the superb benefits of using PS over conventional coherence is that PS separates the amplitude information from the phase. Nevertheless, problem of common reference imposes limitations to PS too. It has been shown that if two signals are synchronized, their differences from a third signal are not necessarily synchronized. Based on this, the choice of the reference electrode can considerably affect the synchrony values obtained, up to the point that they almost span the entire interval $[0, 1]$ (see [47]).

There are a number of different ways to choose the reference electrode such as *typical EEG reference montages*, which all voltage differences refer to one specific channel, *bipolar montage*, which reference electrode for every channel is different and normally spatially close to it, *average reference* and the *Laplacian reference* which is a mathematical approximation to a reference-free signal. All these techniques are susceptible to common reference problem but it is preferred to use the *Laplacian reference* method [47].

Also as reported in [47], the significance-level in PS measurements will not help and again high synchrony is detected when the reference electrode amplitude is relatively high. As a conclusion, utilizing MEG data in PS measurements and more generally in brain connectivity studies will produce much more reliable results.

1.5.4 Noise Sources and Spurious Synchrony

A noise source or an artifact in cerebral signal is referred to any unwanted contribution to the field potential. This might vary due to different applications. For example, when studying ERPs the spontaneous background EEG/MEG is considered as an artifact while in many other applications the background EEG/MEG is the subject of study. Also there are some common noise sources in EEG recordings that require a preprocessing step where the artifacts are detected and removed. Artifacts generated by the electrical equipment which are near the recording area, contaminations caused by muscle activity (i.e. blinking, jaw clenching or frowning) and also other contaminant biological signals (i.e. EOG) are such artifacts. Neglecting these kind of noise sources which are uncorrelated to the signals under study in phase synchronization and brain connectivity measurements might impose incorrect PS values and cause *Spurious Synchrony*.

1.5.5 Phase Enslaving

“Narrow-band filtering tends to transfer phase behavior consistent with the target frequency (or frequency band) from the time region where this holds to adjacent regions where corresponding oscillations are lacking. Conversely, if the signal exhibits distinct oscillations of some particular frequency f , these may strike through to the phase course even if f is incompatible with the target frequency. Such tends to happen particularly in time-frequency regions where power is low. Since power in EEG signals decays rapidly with increasing frequency, phase enslaving may thus especially affect higher frequency (i.e. gamma) oscillations” [70].

1.6 Applications

Phase analysis has a broad usage in many fields. It has been used widely in communication investigations (i.e. signal phase modulation/demodulation) and also in biomedical researches, particularly in field of neuroscience (i.e. brain functional integration and cerebral signal synchrony measurements).

The most frequent applications in brain connectivity measurements are typically the *Brain-Computer Interface (BCI)*, *Event Related Potential (ERP)* and the *Cognitive* studies. Recent applications include investigation of cognitive processes such as visual perception, mental rotation [50], or pathological states such as epileptic seizures [39], schizophrenia [51], [52], attention deficit (hyperactivity) disorder [53] or migraine [54].

1.7 Summary

In this chapter an overview of frequent cerebral signals, their history and their early and modern analysis methods, particularly in cerebral connectivity measurements were presented.

Most of the early quantities gained by digitalizing the EEG records were introduced. Quantities such as *PSD*, *Correlation Function*, signals' *Instantaneous Phase* and *Amplitude* and also the concept of *Coherence*. In modern analysis the main problem discussed was estimating functional integration between neural regions from spatio-temporal patterns in cerebral signals. To this purpose, the concept of *Synchronization* was introduced. It had been discussed that synchronization can be classified into three main categories: (1) linear, (2) nonlinear and (3) information-based methods. Conventional approaches in these categories were also represented. Finally, the major obstacles confronting in this new fast growing field were discussed and many possible and commonly used applications to this concept were introduced.

The next Chapter is dedicated to delicately representing many introduced *Phase Analysis* methods in various distinct fields which are applicable to brain studies.

Chapter 2

Phase Analysis Methods

Phase information is generally considered to be “purer” and less contaminated or more informative than the amplitude of cerebral recordings, which is more influenced by the impedance of the skull or various artifacts such as eyes or face muscles movements [55]. Biological signal phase is an informative useful alternative feature in aforementioned cases that has been of interest in many different research fields through past few years. Thus, using phase of brain signals in cases that amplitude informations are contaminated or inadequate, or in cases that the phase itself has valuable adjunct information over amplitude (i.e. brain cognitive response investigations), would be very helpful.

As mentioned before, measuring brain connectivity is of a particular interest in neuroscience. Because of the reentry mechanism, phase variables and particularly their synchronization for the purpose of measuring connectivity in cerebral signals became more interesting [16].

Various methods within different fields and frameworks are presented in literature for the purpose of cerebral signal phase analysis. The main frameworks and basic concepts were introduced in previous Chapter. In this Chapter, we try to represent and elaborate these various methods within their corresponding field and framework.

2.1 Coherence

2.1.1 Conventional MSC

As introduced earlier in Section (1.1.0.4.1), the conventional method presented in [17] for measuring coherence is based on weighted windowing of the Fourier transform of signals. Let $x(t)$ and $y(t)$ be two random process, zero mean, wide-sense stationary and ergodic signals with length l . A summary of the method in [17] is given below:

- Divide the signals into N equal parts with T samples (different part can be overlapped or disjoint)
- Multiply the samples of each ensemble in a weighting function (i.e. cosine)
- Take *FFT* of each weighted ensemble
- Measure power spectral densities of signals (*PSD*)
- Finally, measure the coherence function (or as said “*Mean Squared Coherence (MSC)*”)

If $X_i(f)$ and $Y_i(f)$ be the Fourier transform of i -th ensemble, the PSDs are calculated as follow:

$$PSD_{xx}(f) = \frac{1}{NT} \sum_{i=1}^N |X_i(f)|^2 \quad (2.1)$$

$$PSD_{yy}(f) = \frac{1}{NT} \sum_{i=1}^N |Y_i(f)|^2 \quad (2.2)$$

$$PSD_{xy}(f) = \frac{1}{NT} \sum_{i=1}^N X_i(f)Y_i^*(f) \quad (2.3)$$

where the signals are sampled at f_s ($f_s > 2BW$) and the BW is the band width. N is the number of ensembles with 50% overlap, T is the number of samples in ensembles and a p -point FFT is taken ($p = Tf_s$).

With the ergodicity of signals being neglected we have:

$$PSD_{xx}(f) = E\{|X_i(f)|^2\} \quad (2.4)$$

$$PSD_{yy}(f) = E\{|Y_i(f)|^2\} \quad (2.5)$$

$$PSD_{xy}(f) = E\{X_i(f)Y_i^*(f)\} \quad (2.6)$$

where $E\{\cdot\}$ illustrates the mathematical expectation. Using these PSDs we compute the mean squared coherence MSC as:

$$|MSC(f)|^2 = \frac{|PSD_{xy}(f)|^2}{PSD_{xx}(f)PSD_{yy}(f)} = \frac{|\sum_{i=1}^N X_i(f)Y_i^*(f)|^2}{\sum_{i=1}^N |X_i(f)|^2 \sum_{i=1}^N |Y_i(f)|^2} \quad (2.7)$$

and the complex coherence function is obtained as below:

$$MSC(f) = \frac{PSD_{xy}(f)}{\sqrt{PSD_{xx}(f)PSD_{yy}(f)}} \quad (2.8)$$

It is notable that T have to be chosen large enough to reduce the bias and standard deviation of the measurements. Also the Fourier transform of the weight function have to be narrow. Thus a good resolution for PSD measurements is provided when N and T both are chosen large (the best resolution is obtained in 50% overlap between ensembles).

The important question confronting here is that “*what is the optimum way for averaging the MSC over ensembles?*”. As reported in [19], the procedure described above for computing averaged MSC might prevent some coherencies in case of non-stationary phase from being detected. Another method presented for averaging MSC over trials is described below:

$$|MSC(f)|^2 = \frac{\frac{1}{N} \sum_{i=1}^N |(PSD_{xy}(f))_i|^2}{PSD_{xx}(f)PSD_{yy}(f)} \quad (2.9)$$

2.1.2 MSC for Deterministic Evoked Signals

In some applications, for example detecting brain evoked potentials during a rhythmic periodic stimuli the MSC could be used to measure the correlation and synchronization between evoked potentials and background EEG signal. For Instance, in [56] they assumed $x(t)$ to be the evoked potential in response to a rhythmic periodic stimuli and $y(t)$ as the background EEG. If the windowing process is done in the way that all the ensembles contain equal number of the periods of signal $x(t)$ thus a modification as below is possible. Due to main equation presented for MSC in Section (1.2.1.4) we have:

$$|MSC(f)|^2 = \frac{|\sum_{i=1}^N Y_i(f)X_i^*(f)|^2}{\sum_{i=1}^N |X_i(f)|^2 \sum_{i=1}^N |Y_i(f)|^2}$$

since $x(t)$ is deterministic, it can be put out from the summation as below:

$$|MSC(f)|^2|_{x(t)periodic} = \frac{|\sum_{i=1}^N Y_i(f)|^2}{N \sum_{i=1}^N |Y_i(f)|^2} \quad (2.10)$$

where $0 < MSC(f) < 1$, and it reaches to its maximum value (i.e. 1), when there is only the evoked potentials observing in the windows.

2.1.3 Multiple MSC

The extension of coherence concept and increasing in number of existing signals, particularly in case of multi-channel EEG recordings presents another aspect of coherence known as *multiple coherence*. Multiple coherence shows the relation and synchronization between a signal and a group of other signals. This concept is used to increase the evoked potential detection rate for multi-channel EEG recordings in [57]. Similar to previous method, signal $x(t)$ assumed to be the brain response of a rhythmic periodic stimuli (i.e. evoked potential) and so deterministic and $y_i(t)s(i = 1, 2, \dots, L)$ as the different EEG signals recorded from different electrodes. The coherence function is computed as below:

$$|MSC(f)|^2|_{x:y_1,\dots,y_L} = \frac{PSD_{yx}^H(f)PSD_{yy}^{-1}(f)PSD_{yx}(f)}{PSD_{xx}(f)} \quad (2.11)$$

where H indicates *Hermitian* of a matrix and the cross-spectrum vector and auto-spectrum matrix are as follow:

$$PSD_{yx}(f) = \begin{pmatrix} PSD_{y_1x}(f) \\ \vdots \\ PSD_{y_Lx}(f) \end{pmatrix} \quad (2.12)$$

$$PSD_{yy}(f) = \begin{pmatrix} PSD_{y_1y_1}(f) & \dots & PSD_{y_1y_L}(f) \\ \vdots & \ddots & \vdots \\ PSD_{y_Ly_1}(f) & \dots & PSD_{y_Ly_L}(f) \end{pmatrix} \quad (2.13)$$

With $x(t)$ being deterministic, same as previous method, its Fourier transform has an identical value in all ensembles and a simplification could be made:

$$PSD_{yx}(f)|_{x(t)periodic} = \begin{pmatrix} \sum_{i=1}^N Y_{1i}^*(f)X(f) \\ \vdots \\ \sum_{i=1}^N Y_{Li}^*(f)X(f) \end{pmatrix} \quad (2.14)$$

$$PSD_{yx}(f)|_{x(t)periodic} = X(f) \begin{pmatrix} \sum_{i=1}^N Y_{1i}^*(f) \\ \vdots \\ \sum_{i=1}^N Y_{Li}^*(f) \end{pmatrix}$$

$$PSD_{yx}(f)|_{x(t)periodic} = X(f)V(f) \quad (2.15)$$

Putting this result into equation (25) results:

$$|MSC(f)|^2|_{x(t)periodic} = \frac{V^H(f)PSD_{yy}^{-1}(f)V(f)}{N} \quad (2.16)$$

This procedure for measuring MSC is helpful calculating phase synchronization in multi-channel EEG recordings and particularly in application of detecting evoked potentials among a group of other channels of background records. As an example of recent studies using MSC for investigating cognitive sources of specific EEG based features within brain, we can mention [89].

2.1.4 MSC Based on Thomson Multitapers

A novel method based on *Thomson Multitapers* [58], for determining coherence function is presented in [59]. In this method, the left and right singular vectors of the cross-covariance matrix corresponding to a known cross-spectrum model are evaluated as multitapers for estimation of cross-spectrum and coherence. A random process, zero mean, stationary signal $x(t)$, is filtered through a causal linear filter $h(n)$ and a stationary sequence $y(t)$ is created and then the MSC function is measured using a multitaper spectrum estimator as:

$$|MSC(f)^{mt}|^2 = \frac{|PSD_{xy}^{mt}(f)|^2}{PSD_{xx}^{mt}(f)PSD_{yy}^{mt}(f)} \quad (2.17)$$

where $-0.5 < f < 0.5$ and:

$$PSD_{xy}^{mt}(f) = \frac{1}{N} \sum_{i=1}^N u_i^H F^H(f) X Y^H F(f) v_i \quad (2.18)$$

$$PSD_{xx}^{mt}(f) = \frac{1}{N} \sum_{i=1}^N u_i^H F^H(f) X X^H F(f) u_i \quad (2.19)$$

$$PSD_{yy}^{mt}(f) = \frac{1}{N} \sum_{i=1}^N v_i^H F^H(f) Y Y^H F(f) v_i \quad (2.20)$$

where n samples of signals, $x = [x(0), \dots, x(n-1)]^T$ and $y = [y(0), \dots, y(n-1)]^T$ are used, and H is showing conjugate transpose and the left and right windows are $u_i = [u_i(0), \dots, u_i(n-1)]^T$ and $v_i = [v_i(0), \dots, v_i(n-1)]^T$ respectively and $F = \text{diag}[1, e^{-j2\pi f}, \dots, e^{-j2\pi(n-1)f}]$ is the Fourier transform matrix.

According to [59], using this procedure for measuring MSC has some advantages over former methods (i.e. based on weighted overlapped segment averaging) in many applications.

2.1.5 Imaginary Part of Coherence

A new method based on imaginary part of the complex coherence function is presented in [61] and it has been shown that the results of this method can not be caused by *Volume Conduction* (VC). This method assumes that there is no time lag between the potential recorded on the scalp and the source of this potential. This assumption is proved through an investigation in [62]. During an accurate study, they do not observed any time lag or any shift in phase of the recorded potential compared with source of the corresponding signal for frequencies under $100Hz$.

A very striking point discussed in [61] is that they argued about using coherency for non-stationary signals and reported that the coherence analysis is feasible in non-stationary processes as strong as stationary ones.

The basic idea used in method presented by [61] is using the real and imaginary part of the coherency instead of its amplitude and phase. Since the real and imaginary parts of coherency are just another representation of complex coherency function, no new quantities are being calculated and it is just a new perspective.

According to [61], the (complex) coherency of the non-interacting areas of brain are necessarily real. Therefore the imaginary part of the coherency can be a good choice for investigation on interacting areas of brain. To see this, assume that the M possible active sources at any time are given as follow:

$$S = [S_1, S_2, \dots, S_M]^T$$

and the L scalp electrodes as:

$$X = [X_1, X_2, \dots, S_L]^T$$

now we have:

$$\begin{pmatrix} x_1(t) \\ \cdot \\ \cdot \\ \cdot \\ x_L(t) \end{pmatrix} = \begin{pmatrix} a_{11} & \cdot & \cdot & a_{1M} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{L1} & \cdot & \cdot & a_{LM} \end{pmatrix} \begin{pmatrix} s_1(t) \\ \cdot \\ \cdot \\ \cdot \\ s_L(t) \end{pmatrix} \quad (2.21)$$

where a_{lm} denote the degree to which source s_m project to electrode x_l . With the coefficient a_{lm} assume to be constant, we have in frequency domain:

$$\begin{pmatrix} X_1(f) \\ \cdot \\ \cdot \\ \cdot \\ X_L(f) \end{pmatrix} = \begin{pmatrix} a_{11} & \cdot & \cdot & a_{1M} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{L1} & \cdot & \cdot & a_{LM} \end{pmatrix} \begin{pmatrix} S_1(f) \\ \cdot \\ \cdot \\ \cdot \\ S_L(f) \end{pmatrix} \quad (2.22)$$

Now assume that the signals recorded from two adjacent electrodes, $x(t)$ and $y(t)$ are superposed of K independent sources, then according to above equations we have:

$$X(f) = \sum_{i=1}^K a_i S_i(f) \quad (2.23)$$

$$Y(f) = \sum_{j=1}^K b_j S_j(f) \quad (2.24)$$

Now based on the main assumption of this method, while computing the cross-spectrum of these signals we will have:

$$PSD_{xy}(f) = E\{X(f)Y^*(f)\} = \sum_i \sum_j a_i b_j E\{S_i(f)S_j^*(f)\} \quad (2.25)$$

$$\Rightarrow PSD_{xy}(f) = \sum_i a_i b_i E\{S_i(f)S_i^*(f)\} = \sum_i a_i b_i |S_i(f)|^2 \quad (2.26)$$

which is real. Thus the normalization to auto-spectrum terms will also be real and so will be the coherence function. Based on this result and the assumption, if the VC causes no time lag (which has been shown in [62]), the imaginary part of the coherency has no effect of VC artifact in it. This also has been shown in [61] by adding non-interacting signal (i.e. noise) and experiencing a decrease in imaginary part of coherency.

In fact, the imaginary part of coherency shows the phase synchronization between time-lagged signals. This claim is coming from the fact that time-lag in time domain causes a phase shift in frequency domain and vice versa. Therefore, it could be said that the **imaginary part of coherency** specifically shows the synchronization between **time-lagged** signals.

According to [61], however the amplitude and phase of the complex coherence function provide similar information to its real and imaginary part, but looking to the imaginary part instead of phase have some advantages as below:

- Non-interacting sources do not lead to small but rather to random phases. We cannot interpret a phase without having an estimate of its significance at the same time.

- One usually calculates coherency with respect to a baseline (a rest condition). Since in the individual coherencies the real parts are typically much larger than the imaginary parts, the phase flips by π depending on whether the real part of coherency is larger in the rest or active condition. The interesting structure is easily obscured by this rather meaningless effect.
- Phase is usually regarded as an additional information about time delay between two processes. However, volume conduction strongly affects the real part but does not create an imaginary part. Processes can appear to be synchronized with almost vanishing time delay while it is only the volume conducted copies of the signals which do not have a time delay.

Finally, stochastic given by [61] to measure the imaginary part of coherency is presented in follow. It has been shown that for two signals $x(t)$ and $y(t)$, the imaginary part of power spectral density is determined as below:

$$\text{Imag}\{X(f)Y^*(f)\} = \sum_{i=1}^K \sum_{j=1}^K A_i A_j (a_i b_j - a_j b_i) \sin(\Delta\theta_{ij}) \quad (2.27)$$

where i and j indicate the independent sources and finally the imaginary part of coherency (ImagC) is obtained as:

$$\text{Imag}C_{xy}(f) = \frac{E\{\text{Imag}\{X(f)Y^*(f)\}\}}{\sqrt{E\{A_x\}^2 E\{A_y\}^2}} = \frac{E\{\text{Imag}\{PSD_{xy}(f)\}\}}{\sqrt{E\{A_x\}^2 E\{A_y\}^2}} \quad (2.28)$$

which varies between 0 and 1 (similar to $MSC(f)$).

2.1.6 Phase Lag Index

The ImagC normalizes the components of cross-spectrum to the amplitudes of signals. Thus it is not true if one says it is only related to the imaginary components of cross-spectrum. Adding uncorrelated noise sources causes the amplitude of cross-spectrum to increase while the numerator of ImagC fraction stays unaffected. This problem is investigated in [63] and by simulations it has been shown that the normalization of ImagC to the amplitudes of signals increases its sensitivity to additional uncorrelated noise sources and changes in phase of coherency. Due to this, another method for measuring phase synchronization is represented in [63] to confront these problems. As a potential improvement on the ImagC, [63] proposed the phase lag index (PLI). The PLI estimates, for a particular frequency, to what extent the phase leads and lags between signals from two sensors are non-equiprobable, irrespective of the magnitude of the phase leads and lags. In simulations, the PLI performed better than the ImagC in detecting true changes in phase synchronization, and was less sensitive to the addition of volume conducted noise sources [64]. The main purpose of PLI was to solve two major problems (1) *volume conduction* and (2) *common reference electrode*.

“The central idea is to discard phase differences that center around $0 \bmod \pi$. One way to realize this is to define an asymmetry index for the distribution of phase differences, when the distribution is centered around a phase difference of zero. If no phase coupling exists between two time series, then this distribution is expected to be flat. Any deviation from this flat distribution indicates phase synchronization. Asymmetry of the phase difference distribution means that the likelihood that the phase difference $\Delta\phi$ will be in the interval $-\pi < \Delta\phi < 0$ is different from the likelihood that it will be in the interval $0 < \Delta\phi < \pi$. This asymmetry implies the presence of a consistent, nonzero phase difference (lag) between the two time series. The existence of such a phase difference or time lag, however, cannot be explained by the influences of volume conduction from a single strong source or an active reference, since these influences are effectively instantaneous. The distribution is expected to be symmetric when it is flat (no coupling), or if the median phase difference is equal to or centers around a value of $0 \bmod \pi$ ” [63].

An index of the asymmetry of the phase difference distribution can be obtained from a time series of phase differences $\Delta\phi_{t_k}$, $k = 1, \dots, N$ in the following way:

$$PLI = |E\{sign[\Delta\phi_{t_k}]\}| \quad (2.29)$$

Phase lag index measures phase differences as they occur on the unit circle by first thresholding each $\Delta\phi$ using the signum function and then averaging over successive data points [63]. The PLI ranges between 0 and 1. A PLI of zero indicates either no coupling or coupling with a phase difference centered around $0 \bmod \pi$. A PLI of 1 indicates perfect phase locking at a value of $\Delta\phi$ different from $0 \bmod \pi$. The stronger this nonzero phase locking is, the larger PLI will be. Note that PLI does no longer indicate, which of the two signals is leading in phase. Whenever needed, however, this information can be easily recovered, for instance, by omitting the absolute value in above equation [63].

2.1.7 Weighted Phase Lag Index

The PLI is said to have some defects. Recently, in a comprehensive study in [64], it has been shown that PLI has some problems with VC, noise sources and detecting changes in phase synchronization due to its discontinuity. It is said that in PLI a small perturbation turns phase lags into phase leads and vice versa. Based on this, another quantity for measuring phase synchronization known as *Weighted Phase Lag Index* (WPLI) is presented in [64]. It is said that WPLI is an improvement over both the conventional PLI and the ImagC. According to [64], the advantage of WPLI over PLI is that it weights the observed phase lags and leads by the magnitude of the imaginary components of cross-spectrum that causes less sensitivity to additional uncorrelated noise sources and more power of detecting phase synchronization statistically.

As mentioned before (see Section 1.4), in case of phase synchronization measurements four problems are well known:

- The presence of a common reference signal
- Volume Conduction
- The presence of noise sources
- Sample-size bias

VC, and in case of EEG (but not MEG) data, the use of a common reference, can spuriously inflate phase synchronization indices. The problem of VS is especially large for scalp EEG and MEG data, because of their low spatial resolution.

Due to these problems, to increase the detection rate of true changes in phase synchronization and also to decrease the effects of additional uncorrelated noise sources, coherency phase changes and the common reference problem, the represented statistics for WPLI are as follow:

$$WPLI(f) = \frac{|E\{Imag\{X(f)Y^*(f)\}\}|}{E\{|Imag\{X(f)Y^*(f)\}\}|} = \frac{|E\{Imag\{PSD_{xy}(f)\}\}|}{E\{|Imag\{PSD_{xy}(f)\}\}|} \quad (2.30)$$

$$WPLI(f) = \frac{|E\{|Imag\{PSD_{xy}(f)\}\}sign\{Imag\{PSD_{xy}(f)\}\}|}{E\{|Imag\{PSD_{xy}(f)\}\}|} \quad (2.31)$$

As it is clear, unlike the ImagC, the WPLI uses the components of imaginary part of the cross-spectrum to normalize and also unlike the conventional PLI, $sign\{Imag\{PSD_{xy}(f)\}\}$ is weighted by $|Imag\{PSD_{xy}(f)\}|$. The WPLI also varies between 0 to 1 and find its maximum value in:

$$sign\{Imag\{PSD_{xy}(f)\}\} = \pm 1$$

when:

$$|E\{|Imag\{PSD_{xy}(f)\}|sign\{Imag\{PSD_{xy}(f)\}\}| = E\{|Imag\{PSD_{xy}(f)\}|\}$$

The performance of WPLI is evaluated in confront with the mentioned phase synchronization measurement difficulties in a wide simulation experiment in [64] and it has been reported that the WPLI performed much better in comparison with both PLI and ImagC. As an instance of recent studies using PLI for BCI application we can mention [90].

2.1.8 MSC Based on Parametric Spectral Estimation Methods

Each of these represented method for analyzing of signal phase exhibit some sensitivity to both changes in background power and the instantaneous effects caused by volume conduction. As an improvement to all these methods, the traditional spectral estimation based on Fourier Transform of the raw EEG data, can be replaced by a parametric estimation derived using the *Yule-Walker* AR approach. After this as before, we have to look for periods of phase synchronization [65].

Let p be the model order, $v_i(t)$ be the noise term. For two assumptive signals $x(t)$ and $y(t)$ the AR model could be derived as:

$$x(t) = \sum_{q=1}^p b_{11}(q)x(t-q) + \sum_{q=1}^p b_{12}(q)y(t-q) + v_1(t) \quad (2.32)$$

$$y(t) = \sum_{q=1}^p b_{21}(q)x(t-q) + \sum_{q=1}^p b_{22}(q)y(t-q) + v_2(t) \quad (2.33)$$

which in frequency domain turns into:

$$X(f) = \sum_{q=1}^p B_{11}(q)X(f)e^{-j2\pi fq} + \sum_{q=1}^p B_{12}(q)Y(f)e^{-j2\pi fq} + V_1(f) \quad (2.34)$$

$$Y(f) = \sum_{q=1}^p B_{21}(q)X(f)e^{-j2\pi fq} + \sum_{q=1}^p B_{22}(q)Y(f)e^{-j2\pi fq} + V_2(f) \quad (2.35)$$

To use this model in cross-spectrum measurement and then coherence function estimation using EEG/MEG signals we can extend this to a multi-channel case (i.e. multi-channel EEG). Let $x(t) = [x_1(t), \dots, x_L(t)]^T$ be the scalp electrodes. Derivation of **Multy-channel AR Model** gives ($B(0) = -I$, $F.T.\{v(t)\} = N(f)$) [65]:

$$x(t) = \sum_{q=1}^p b(q)x(t-q) + v(t) \quad (2.36)$$

taking Fourier transform of this equation gives:

$$X(f) = \sum_{q=1}^p B(q)X(f)e^{-j2\pi fq} + N(f) \quad (2.37)$$

$$- \sum_{q=1}^p B(q)X(f)e^{-j2\pi fq} + X(f) = N(f)$$

$$(I - \sum_{q=1}^p B(q)e^{-j2\pi fq})X(f) = N(f)$$

$$\begin{aligned} \left(-\sum_{q=0}^p B(q)e^{-j2\pi fq}\right)X(f) &= N(f) \\ \hat{B}(f)X(f) &= N(f) \end{aligned} \quad (2.38)$$

where $\hat{B}(f) = -\sum_{q=0}^p B(q)e^{-j2\pi fq}$. If $H(f) = \hat{B}^{-1}(f)$ multiplying in equation above gives the cross-spectrum as:

$$\begin{aligned} X(f) &= H(f)N(f) \\ X(f)X^*(f) &= H(f)N(f)N^*(f)H^*(f) \end{aligned} \quad (2.39)$$

which is an estimation of cross-spectrum needed for establishing the coherence function. Normalization of the temporal averaged version of this cross-spectrum (i.e. expectation) to the amplitude of the estimated AR coefficients gives the coherence function. To estimate the coefficients of AR model, as presented in [66], the least squares modified Yule–Walker equation (*LSMYWE*) estimator [67], can be used:

$$\hat{b} = -(\hat{R}^H \hat{R})^{-1} \hat{R}^H \hat{r} \quad (2.40)$$

$$\hat{R} = \begin{pmatrix} \hat{r}_{xx,p} & \hat{r}_{xx,p-1} & \cdot & \cdot & \hat{r}_{xx,1} \\ \hat{r}_{xx,p+1} & \hat{r}_{xx,p} & \cdot & \cdot & \hat{r}_{xx,2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \hat{r}_{xx,M-1} & \hat{r}_{xx,M-2} & \cdot & \cdot & \hat{r}_{xx,M-p} \end{pmatrix} \quad (2.41)$$

$$\hat{r} = \begin{pmatrix} \hat{r}_{xx,p+1} & \hat{r}_{xx,p+2} & \cdot & \cdot & \hat{r}_{xx,M} \end{pmatrix} \quad (2.42)$$

where $\hat{r}_{xx,n}$ shows the i -th element of auto-correlation estimation for signal x_n , p is the AR model order and $M = 2p$.

As reported in [65], the parametric and non-parametric spectral estimation methods perform similarly in low noise power while in high noise power the parametric methods act much better in detecting periods of phase synchronization. Using this parametric model in first step of the presented phase synchronization detection methods will improve the results of them and make them less impressible to noise, background power changing and volume conduction effects. The main problem confronting here is dealing with number of channels and the optimum model order.

It is notable that using non-parametric methods have some important advantages too. For instance, parametric methods assume that signals are generated by a linear auto-regressive process but the non-parametric methods do not consider this assumption. Moreover, there are less parameters to be adjusted in non-parametric methods (i.e. the window length), but in parametric methods, after selecting the number of channels and adjusting the model order, the AR model coefficients have to be estimated and after that other parameters such as window length and etc. are set.

2.2 Phase Synchronization

As discussed before, in synchronization concept, coherence deals with several objections. The linearity, stationarity and also the lack of ability to separate the effects of amplitude and phase informations are the main abjections confronting. Thus, another concept called *phase locking* was introduced to overcome these obstacles.

2.2.1 Phase Locking Value

There are two major approaches to calculate this quantity. Both of these methods are based on the framework presented in Section (1.3.2.1). It means in first step they utilize one of the phase extraction methods and then after simply calculating the phase-differences, they quantify the local

stability in this phase-differences and finally, in significance-level the significance of captured values are determined. The major difference between methods presented for *Phase Locking Value* (PLV) is in the manner they use to deal with the first and third steps. Due to different techniques utilized in these steps, we represent various statistics for PLV as follow.

2.2.1.1 Phase-Locking Statistics

This method, called *Phase-Locking Statistics* (PLS), measures the significance of the phase-covariance between two signals. PLS separates the phase and amplitude components and can be directly interpreted in the framework of neural integration [34].

To have a meaningful interpretation from instantaneous phase, it has to be extracted through a narrow filter which isolates the frequency of interest. In this method, First of all the signal is band-pass filtered in its frequency of interest and a short arbitrary interval around it (i.e. $f \pm 2$). After that, its convolution with a complex **Gabor-Wavelet** centered at frequency f is calculated and phase of this convolution is extracted for all time-bins t , trials $n = 1, \dots, N$, and for each of the pair of electrodes as discussed before in Section (1.2.1.3.2). As mentioned before, when using Wavelet Transform description below gives back the phase-differences which is calculated across trials, as in the second step [34], [35]:

$$\exp(j(\phi_y(t, f) - \phi_x(t, f))) = \frac{W_x(t, f)W_y^*(t, f)}{|W_x(t, f)||W_y(t, f)|} \quad (2.43)$$

where the term $\phi_y(t, f) - \phi_x(t, f)$ represents the phase-difference between processes $y(t)$ and $x(t)$.

As the third step, the stability of phase-differences across trials is quantified by a phase locking value:

$$PLV(t, f) = \left| \frac{1}{N} \sum_{i=1}^N \exp(j(\phi_{y,i}(t, f) - \phi_{x,i}(t, f))) \right| \quad (2.44)$$

where N is the total number of trials and i indicates the index of trial. PLV varies between 0 and 1 which indicate the completely non-synchronized and synchronized signals respectively. PLV is said to has phase enslaving problem when the power is low.

As the final step in significance-level, the degree of statistical significance of each PLV value is determined by comparing it to the values obtained by the *surrogate data* which is a statistical test with the null hypothesis that states: “the value of original data is not significantly different from the specific class of signals” [60]. Efficient surrogate data are generated in a way that they simulate all signals’ characteristics, but destroy only the one property suspected to be the source of outstanding values of measure computed [60]. For this purpose, as in [35], the surrogate data is generated by shifting the trials of one of the main signals (i.e. $y(t)$) in such a way that the phase-differences are no longer computed in simultaneously recorded trials. It means we have:

$$PLV_{surrogate}(t, f) = \frac{1}{K} \sum_{j=1}^K \left| \frac{1}{N} \sum_{i=1}^N \exp(j(\phi_{y,j(i)}(t, f) - \phi_{x,i}(t, f))) \right| \quad (2.45)$$

The proportion of surrogate values greater than the original PLV for a tim t is called PLS. Normally a criterion of 5% ($PLS < 5\%$) is used to characterize significant synchrony [35]. As direct applications of PLV in recent studies we can mention [85, 86, 88, 89] and [90]. Further implementation details and MATLAB codes related to these studies can be found in [87] and [91].

2.2.1.2 Single-trial Phase Locking Statistics

As an extension of PLS, *Single-trial Phase Locking Statistics* (S-PLS) is presented in [35] in order to estimate PLS in single-trials. S-PLS shares the whole idea with PLS except at a slight cost of

temporal resolution, in this method the phase-difference variability is calculated through successive time-steps (not across trials). Based on this, the S-PLV for each individual trial is obtained as below:

$$S - PLV(t, f) = \left| \frac{1}{\delta} \int_{t-\frac{\delta}{2}}^{t+\frac{\delta}{2}} \exp(j(\phi_y(\tau, f) - \phi_x(\tau, f))) d\tau \right| \quad (2.46)$$

As PLV, the S-PLV varies between 0 and 1 with 1 indicating the strongest phase locking [35].

The significance-level in this method is performed similar to that of PLS except that here surrogate values are generated by estimating the maximum PLV between pairs of independent white noise signals. Finally, S-PLS is obtained through a similar criteria to that of PLS.

2.2.1.3 Phase Synchrony Based on Shannon Entropy

The instantaneous phase sequence which is used to estimate the PLV can also be captured through the Hilbert transform (statistics are given in Section 1.2.1.3.1). In a method presented in [36], as the first step they extract the instantaneous phase of signals by Hilbert transform. A merit of the this approach is that the phase can be easily obtained for an arbitrary broad-band signal. Nevertheless, instantaneous amplitude and phase have a clear physical interpretation only if the signals are narrow-band [38]. Therefore, an extra filtering pre-step is required to separate the frequency band of interest. After the pre-step filtering and extracting the phase sequence from the frequency band of interest, Tass et al. [36], proposed that *Shannon Entropy* can be used to characterize the statistical strength of phase synchrony. For this purpose, “the deviation of the actual distribution of the phase-difference between recording signals from a uniform one must be quantified” [38]. The synchronization index proposed by Tass et al. in [36] can be illustrated as follow:

$$\gamma = \frac{H_{max} - H}{H_{max}} \quad (2.47)$$

where H is the entropy and defined by (see Section 1.3.3):

$$H = - \sum_{i=1}^N P_i \ln P_i \quad (2.48)$$

which N is the total number of trials, and $H_{max} = \ln(N)$ is the maximal entropy, and P_i the relative frequency of finding the phase-differences within the $i - th$ trial [38].

As reported in [36], the optimal number of trials can be obtained as:

$$N = \exp(0.626 + 0.4 \ln(n - 1)) \quad (2.49)$$

where n is the number of samples.

Finally, the degree of statistical significance of the phase-locking values was determined through generating surrogate values obtained from surrogate data. Surrogate data is constructed by *scrambling* the original series. This randomization destroys any temporal structure, if present in the original series [36], [38]. The γ is a normalized quantity ($0 < \gamma < 1$) which measures the strength of phase synchronization and the $\gamma = 1$ indicates the perfect synchrony.

2.2.1.4 Phase Synchrony Based on Mutual Information

Another phase-locking statistic was introduced in [37]. This method is also similar to previous methods with the whole idea except in procedure presented for quantification the phase synchrony strength. The phase synchrony index in this method is as below:

$$\rho = \frac{I}{I_{max}} \quad (2.50)$$

where $I_{max=\ln(N)}$ and N is total number of trials. I is the mutual information between phase sequences and can be determined as follow:

$$I(\phi_x, \phi_y) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} P_{x,y}(\phi_x, \phi_y) \log\left(\frac{P_{x,y}(\phi_x, \phi_y)}{P_x(\phi_x)P_y(\phi_y)}\right) d\phi_x d\phi_y \quad (2.51)$$

which $P_x(\phi_x)$ and $P_x(\phi_y)$ are the probability distribution of the phases ϕ_x and ϕ_y , and $P_{x,y}(\phi_x, \phi_y)$ is their joint distribution. The *Mutual Information* (MI) in this method measures the dependence between phase sequences of signals $x(t)$ and $y(t)$.

It has been reported that MI produces surplus amount of information caused by an incorrect assumption that two processes are independent [68]. Also, as discussed in Section (1.3.3), MI can be rewritten as entropy as follow:

$$I(\phi_x, \phi_y) = H(\phi_x) + H(\phi_y) - H(\phi_x, \phi_y) \quad (2.52)$$

Finally, the statistical significant of obtained phase synchrony values is tested through significant-level similar to that of previous method. Normalized in this way, just like previous methods, ρ varies between 0 and 1 where latter indicates the complete synchrony [37].

It seems that this method does not provide any further information in comparison to previous method. Although, more accurate results will be presented in next Chapter after implementing these methods using similar data sets and in a unique situation.

2.2.2 Mean Phase Coherence

Another statistical measure for phase synchrony called as *Mean Phase Coherence* (MPC) is presented in [39]. This method utilize circular variance [69] to characterize the dependences between instantaneous phases of two signals. The method presented in [39] shares the two first steps with previous methods where as first step the instantaneous phases are extracted through Hilbert transform and as second step the phase-differences are calculated by simply subtracting the phase sequences. Finally, as a measure for synchronization the MPC of an angular distribution is defined as:

$$MPC = \left| \frac{1}{N} \sum_{i=0}^{N-1} \exp(j\phi_{x,y}(i\Delta t)) \right| \quad (2.53)$$

where $\phi_{x,y}$ indicates the phase differences between phases ϕ_x and ϕ_y and $\frac{1}{\Delta t}$ is the sampling rate of discrete series. Whit utilizing the Euler's formula, the above equation turns into:

$$MPC = \left(\left[\frac{1}{N} \sum_{i=0}^{N-1} \cos(\phi_{x,y}(i\Delta t)) \right]^2 + \left[\frac{1}{N} \sum_{i=0}^{N-1} \sin(\phi_{x,y}(i\Delta t)) \right]^2 \right)^{\frac{1}{2}} \quad (2.54)$$

Clearly from this notation MPC is restricted to the interval $[0, 1]$ [39]. As same as all previous methods presented for phase synchrony, reliability of these obtained synchrony values is evaluated by the surrogate data approach. In [60] they offered to use *Z-score* statistics for the statistical test in significance-level. *Z-score* statistic calculates how different is the value of original data from the mean value of all surrogates in the units of standard deviation [60]:

$$Z = \frac{MPC_{org} - \mu(MPC_{surr})}{\sigma(MPC_{surr})} \quad (2.55)$$

where μ and σ are the mean value and the standard deviation of all surrogates respectively.

2.2.3 Wavelet Coherence

Another method presented in [35] called as *Wavelet Coherence* (WC) uses the wavelet coefficients of signals and the coherence function to provide a **weighted** version of PLV. The basic idea for presenting this method is that instantaneous phase may be meaningless where power is low and former methods are susceptible to a notorious problems called as *phase enslaving* [70]. AS in [35], WC for time delays t and frequency(bands) f can be obtained as below:

$$WC(t, f) = \frac{|\frac{1}{N} \sum_{i=1}^N W_{x_i}(t, f) W_{y_i}^*(t, f)|}{\sqrt{\frac{1}{N} \sum_{i=1}^N |W_{x_i}(t, f)|^2 \frac{1}{N} \sum_{i=1}^N |W_{y_i}(t, f)|^2}} \quad (2.56)$$

Since for both signals $x(t)$ and $y(t)$ we have:

$$\exp(j\phi(t, f)) = \frac{W(t, f)}{|W(t, f)|} \quad (2.57)$$

and also from the original PLV equation:

$$PLV(t, f) = \left| \frac{1}{N} \sum_{i=1}^N \exp(j(\phi_{y,i}(t, f) - \phi_{x,i}(t, f))) \right| \quad (2.58)$$

then we got:

$$W_{x_i}(t, f) W_{y_i}^*(t, f) = \exp(j\Delta_i(t, f)) |W_{x_i}(t, f)| |W_{y_i}(t, f)| \quad (2.59)$$

then assuming:

$$\lambda_i(t, f) = \frac{|W_{x_i}(t, f)| |W_{y_i}(t, f)|}{\sqrt{\frac{1}{N} \sum_{i=1}^N |W_{x_i}(t, f)|^2 \frac{1}{N} \sum_{i=1}^N |W_{y_i}(t, f)|^2}} \quad (2.60)$$

thus WC can be rewritten as a *weighted* variant of PLV as below [70]:

$$WC(t, f) = \left| \frac{1}{N} \sum_{i=1}^N \exp(j\Delta_i(t, f)) \lambda_i(t, f) \right| \quad (2.61)$$

Due to multiplication by $\lambda_i(t, f)$, wavelet amplitudes are low and this clearly reduces the potential consequences of *phase enslaving*. Nevertheless, WC has some defects such as it does not separate the effects of amplitude and phase in the interrelation between signals [70].

2.2.4 Weighted Phase Locking Value

The basic idea in this method presented in [70] is completely similar to WC method. The only different is between procedures for obtaining the weights. Description below known as *Weighted Phase Locking Value* (WPLV) is presented to measure the statistical strength of phase synchrony between signal pairs:

$$WPLV(t, f) = \left| \frac{1}{N} \sum_{i=1}^N \exp(j\Delta_i(t, f)) \omega_{x_i}(t, f) \omega_{y_i}(t, f) \right| \quad (2.62)$$

In contrast to WC, here weights are not proportional to amplitudes. As offered in [70], the weights differ from 1 —which gains the conventional PLV— only when the amplitudes are considered “not large enough” and that depends on the respective spectral characteristics. For calculating the weights assume that for signal $x(t)$ we have (the procedure and statistics for $y(t)$ is repeated):

$$P_x(t, f) = \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} |W_x(t, f)|^2 df \quad (2.63)$$

$$P_x(f) = \frac{1}{T} \int_T P_x(t, f) dt \quad (2.64)$$

where $P_x(t, f)$ denotes the mean wavelet power in frequency band f at time t and $P_x(f)$ is its time average across the relevant trial T . Now, assessment of weights can be summarized as below [70]:

- Large enough: $\omega_x(t, f) = 1$, if $P_x(t, f) > \frac{3}{2}P_x(f)$
- Too small: $\omega_x(t, f) = 0$, if $P_x(t, f) \leq \frac{3}{10}P_x(f)$
- Between: $\omega_x(t, f) = \frac{\log(\frac{P_x(t, f)}{P_x(f)}) - \log(\frac{3}{10})}{\log(5)}$

For both the previous and current methods, since the wavelet transform has to be computed anyway, it is preferred to use it also for assessing the instantaneous phases. As discussed before in previous Chapter, the wavelet phases for frequency bands can be determined as the phase angle of the band-averaged complex wavelet coefficients as follow:

$$\exp(j\phi_x(t, f)) = \frac{\overline{W}_x(t, f)}{|\overline{W}_x(t, f)|} \quad (2.65)$$

where

$$\overline{W}_x(t, f) = \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} W_x(t, f) \quad (2.66)$$

It is notable that *phase enslaving* is less serious problem in case of cerebral signals and it is normally serious when test signals with a sharply defined spectrum is used. Nevertheless, it is expected that WPLV yield more accurate results than the conventional PLV because of discussed problem.

2.2.5 Reduced Interference Rihaczek Distribution Phase Synchrony

This phase synchronization method is based on *Cohen's Class of Distributions* and the *Complex Energy Density Function* (CEDF) phase extraction method. Cohen's class of distributions are bilinear time-frequency distributions and can be illustrated as below [73]:

$$CEDF(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\theta, \tau) A(\theta, \tau) \exp(-j(\theta t + 2\pi\tau f)) d\tau d\theta \quad (2.67)$$

where $\phi(\theta, \tau)$ is the kernel function and $A(\theta, \tau)$ is the ambiguity function and can be obtained as below:

$$A(\theta, \tau) = \int_{-\infty}^{\infty} x(u + \frac{\tau}{2}) x^*(u - \frac{\tau}{2}) \exp(j\theta u) \quad (2.68)$$

“The major differences between Cohen's class of distributions compared to other time-frequency representations such as the wavelet transform are the nonlinearity of the distribution, energy preservation, and the uniform resolution over entire time-frequency plane” [71]. This distributions describe energy of a signal over time and frequency simultaneously but they can not be used for describing phase information of a signal. Thus, we need complex time-frequency distributions to obtain both the energy and the phase information. For this purpose, as discussed earlier in Section (1.2.1.3.3), Rihaczek derived the signal energy distribution in time and frequency by application of the complex signal notation. This CEDF, known as Rihaczek distribution, encounters a major disadvantage due to producing crossterms in case of multi-component signals. These crossterms are at the same time and frequency as the original signals and will lead to biased energy and phase estimates [71]. Assume $z(t) = x(t) + y(t)$, then the Rihaczek distribution is as follows:

$$CEDF(t, f) = \frac{1}{\sqrt{2\pi}} (x(t)X^*(f) \exp(-j2\pi ft) + y(t)Y^*(f) \exp(-j2\pi ft))$$

$$+x(t)Y^*(f) \exp(-j2\pi ft) + y(t)X^*(f) \exp(-j2\pi ft) \quad (2.69)$$

where the last two terms are crossterms. Therefore, a filtering is needed to remove these crossterms out of calculations.

For this purpose, this method proposes a method based on Cohen's class of distributions and uses its kernel function as the required filter. This is called *Reduced Interference Rihaczek Distribution* (RID) and utilizes the *Choi-Williams* kernel to filter out the crossterms. As in [71], the *Reduced Interference Complex Energy Density Function* or simply (RI-CEDF) can be expressed as:

$$RI - CEDF(t, f) = \int \int \exp(-\frac{(\theta\tau)^2}{\sigma}) \exp(j\frac{\theta\tau}{2}) A(\theta, \tau) \exp(-j(\theta t + 2\pi\tau f)) d\tau d\theta \quad (2.70)$$

where $\exp(j\frac{\theta\tau}{2})$ is the Rihaczek distribution kernel function and the integrals are from $-\infty$ to ∞ .

After defining this RI-CEDF, as introduced previously, the phase in time-frequency plane can easily be defined as:

$$\phi_x(t, f) = \arg\left(\frac{RI - CEDF(t, f)}{|RI - CEDF(t, f)|}\right) \quad (2.71)$$

Now, as the second step in phase synchronization evaluation process, we need to estimate the phase-difference between two signals $x(t)$ and $y(t)$. The phase-difference between time-varying phases can be computed as follow:

$$\phi_{x,y}(t, f) = \arg\left(\frac{RI - CEDF_x(t, f)RI - CEDF_y^*(t, f)}{|RI - CEDF_x(t, f)||RI - CEDF_y(t, f)|}\right) \quad (2.72)$$

Finally in this method, as the final step to assess phase synchrony, the conventional PLV measure is used to quantify the phase synchrony strength between pairs of electrodes.

2.2.6 Phase Synchrony Based on Empirical Mode Decomposition

Another set of methods for evaluating phase synchrony utilize *Empirical Mode Decomposition* (EMD) and separate data into *Intrinsic Mode Functions* (IMFs) and then, extract the instantaneous phases of the IMFs and finally use a phase synchrony quantification procedure, for example in [74], [75] and [77].

Extracting signals phase sequences through EMD algorithm was discussed in previous Chapter. First, signals are decomposed into their relative IMFs and then the Hilbert transform is applied to these IMFs to extract phases. As the next step, for evaluating phase synchrony, the phase-differences have to be calculated. Assuming signals $x(t)$ and $y(t)$ and their corresponding phases $\phi_x(t)$ and $\phi_y(t)$ which are extracted using EMD, the phase-difference is simply calculated as:

$$\delta\phi_{x,y}(t) = \phi_y(t) - \phi_x(t) \quad (2.73)$$

Finally, as in [74] and [75], the phase synchrony strength is evaluated by this phase-difference through measuring the conventional PLV.

Phase synchrony measurements using EMD have some advantages over other introduced previous methods [74]. First, it is said that the discriminative band of the phase synchrony phenomenon differs not only with subjects, but also with time. Nevertheless, EMD algorithm can decompose the non-stationary signal into its relative IMFs. Therefore, no subject-specific or frequency-specific factors need to be considered to get the discriminative band [74]. Second, band-pass filter is a required necessary step in previous methods which causes time-delay in output signal with respect to the input signal while EMD based methods prevent this time-delay.

As reported in [77], there are two main obstacles confronting while using this method for phase synchrony purposes: (1) it is computationally inefficient (2) misleading results can be obtained as components of different frequencies can temporarily appear synchronized. To overcome these problems, [77] proposes to use the *Complex Extension of EMD*. For more information refer to [77].

Appendix A

Which Coordinates?

since $PSD_{xy}(f)$ is complex, the *Cartesian* or *polar* coordinates can be used for averaging. In [19] this problem is investigated. Using *Cartesian* coordinate we got:

$$(PSD_{xy}(f))_i = a_i(f) + jb_i(f) \quad (\text{A.1})$$

$$\Rightarrow PSD_{xy}^c(f) = \frac{1}{N} \sum_{i=1}^N (a_i(f) + jb_i(f)) \quad (\text{A.2})$$

where $PSD_{xy}^c(f)$ indicates the cross-spectrum averaged in *Cartesian* coordinates and its amplitude and phase are obtaining as follow:

$$|PSD_{xy}^c(f)| = \sqrt{\left(\frac{1}{N} \sum_{i=1}^N a_i(f)\right)^2 + \left(\frac{1}{N} \sum_{i=1}^N b_i(f)\right)^2} \quad (\text{A.3})$$

$$\angle PSD_{xy}^c(f) = \arctan\left(\frac{\sum_{i=1}^N b_i(f)}{\sum_{i=1}^N a_i(f)}\right) \quad (\text{A.4})$$

Similarly, in *Polar* coordinates we have:

$$PSD_{xy}^p(f) = \rho(\cos(\phi) + jsin(\phi)) \quad (\text{A.5})$$

where $PSD_{xy}^p(f)$ indicates the cross-spectrum averaged in *Polar* coordinates and ρ its amplitude and ϕ its phase are obtaining as follow:

$$\rho = \frac{1}{N} \sum_{i=1}^N |(PSD_{xy}^p(f))_i| = \frac{1}{N} \sum_{i=1}^N \sqrt{a_i^2(f) + b_i^2(f)} \quad (\text{A.6})$$

$$\phi = \frac{1}{N} \angle(PSD_{xy}^p(f))_i = \frac{1}{N} \sum_{i=1}^N \arctan\left(\frac{b_i(f)}{a_i(f)}\right) \quad (\text{A.7})$$

It is clear that equations (A.3) and (A.6) and also equations (A.4) and (A.7) are equal when all $a_i(f)$ s are equal and all $b_i(f)$ s are equal too. In all any other conditions the amplitude and phase averaged in *Polar* coordinates are larger than the amplitude and phase averaged in *Cartesian* coordinates. Due to EEG signal non-stationarity this condition will not occur and thus as reported in [19], the optimum averaged coherence function performance can be obtained through presentation and averaging the power spectral density functions in *Polar* coordinates.

Bibliography

- [1] F. Lopes da Silva. Biological aspects of eeg and magnetoencephalogram generation. In E. Niedermeyer and F. Lopes da Silve, editors, *Electroencephalography : Basic Principles, Clinical Applications and Related Fields*, chapter 5. Urban and Schwarzenberg, 1987.
- [2] R.D. Walter and C.L. Yeager. Visual imagery and electroencephalographic changes. *Electroencephalogr Clin Neurophysiol*, 8:193–199, 1956.
- [3] A. Glass and S.F. Gurney. A percent time alpha computer for the electroencephalograph. *Wld Med Electron Instrum*, 2:206–210, 1964.
- [4] H. Berger. Uber das elektrenkephalogramm des menschen. *Archiv fur Psychiatrie und Nervenkrankheiten*, 87:527–570, 1929.
- [5] H. Davis, P.A. Davis, A.L. Loomis, E.N. Harvey, and G. Hobart. Electrical reactions of the human brain to auditory stimulation during sleep. *J Neurophysiol*, 2:500–514, 1939.
- [6] G.E. Chatrian, L. Bergamini, M. Dondey, D.W. Klass, M. Lennox-Buchtal, and I. Petersen. Report of the committee on terminology. *Electroencephalogr Clin Neurophysiol*, 37, 1974.
- [7] C.D. Binnie, P.A. Ward, and J. Heywood. EEG contour mapping I: Theory and practice. *Proc. Electrophysiol. Technol. Ass.*, 18, 1971.
- [8] C. Bingham, M.D. Godfrey, and J.W. Tukey. Modern techniques of power spectrum estimation. *IEEE Trans on Audio and Electroacoustics*, AU-15:56–66, 1967.
- [9] M. Le Van Quyen, J. Foucher, J-P. Lachaux, E. Rodriguez, A. Lutz, J. Martinerie, and F.J. Varela. Comparison of hilbert transform and wavelet methods for the analysis of neural synchrony. *J Neurosci Methods*, 111:83–98, 2001.
- [10] D. Gabor. Theory of communication. *J Inst Electr Eng*, 93:429–457, 1946.
- [11] M. Chavez, M. Besserve, C. Adam, and J. Martinerie. Towards a proper estimation of phase synchronization from time series. *J Neurosci Methods*, 154:149–160, 2006.
- [12] S. Mallat. "A wavelet tour of signal processing". San Diego: Academic Press (1999).
- [13] D.J. Felleman and D.C. Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*, 1:1–47, 1991.
- [14] L.H. Finkel and G.M. Edelman. Integration of distributed cortical systems by reentry: A computer simulation of interactive functionally segregated visual areas. *J Neurosci*, 9:3188–3208, 1989.
- [15] D.O. Walter. Spectral analysis for electroencephalograms: Mathematical determination of neurophysiological relationships from records of limited duration. *Exp Neurol*, 8:155–181, 1963.

- [16] W. Marshall, "Statistical Analysis of EEG Phase Shift Events", PhD Thesis in Statistics, University of Waterloo, Ontario, Canada, 2014.
- [17] G. Clifford Carter, CH. Knapp, and AH. Nuttall. "Estimation of the Magnitude-Squared Coherence Function Via Overlapped Fast Fourier Transform Processing". IEEE Transaction on Audio and Electroacoustics, VOL. AU-21, NO. 4, (1973).
- [18] DO. Walter. "Spectral Analysis for Electroencephalograms: Mathematical Determination of Neuro physiological Relationships from Records of Limited Duration". Experimental Neurology 8, (1963), 155–181.
- [19] G.D. Brushe, J.R. Waller, "On the Computation of an Averaged Coherence Function", Digital Signal Processing 11, 110–119 (2001).
- [20] G. Clifford Carter. "Coherence and Time Delay Estimation". Proceedings of the IEEE, VOL. 75, NO. 2, (1987)
- [21] M. Hassan, A. Mheich, F. Wendling, O. Dufor, C. Berrou, "Graph-based analysis of brain connectivity during spelling task", 2nd International Conference on Advances in Biomedical Engineering, Sep 2013, Tripoli, Lebanon. pp.191-194, 2013.
- [22] F. Wendling, K. Ansari Asl, F. Bartolomei, L.Senhadji, "From EEG signals to brain connectivity: A model-based evaluation of interdependence measures", Journal of Neuroscience Methods, 183, 9–18, (2009).
- [23] V. Sakkalis, "Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG", Computers in Biology and Medicine 41, 1110–1117, (2011).
- [24] O. Sporns, "Networks of the Brain", MIT Press, 2010.
- [25] A. van Ooyen, Competition in the development of nerve connections: a review of models, Network 12 (1) (2001) R1–47.
- [26] A.A. Fingelkurts, A.A. Fingelkurts, S. Kahkonen, "Functional connectivity in the brain is it an elusive concept?" Neurosci. Biobehav. Rev. 28 (8), (2005), 827–836.
- [27] B. Horwitz, The elusive concept of brain connectivity, NeuroImage 19 (2 Pt 1) (2003) 466–470.
- [28] Pereda, E.; Quiroga, R. Q.; Bhattacharya, J. (2005). "Causal Influence: Nonlinear multivariate analysis of neurophysical signals". Prog Neurobiol 77 (1–2): 1–37.
- [29] Netoff, I.; Carroll, T. L.; Pecora, L. M.; Schiff, S. J. (2006). "Detecting coupling in the presence of noise and nonlinearity". In Schelter, J.; Winterhalder, W.; Timmer. Handbook of Time Series Analysis. Wiley-B.W.
- [30] Stam, C. J.; Van Dijk, B. W. (2002). "Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets". Physica D 163 (3–4): 236–251. Bibcode:2002PhyD..163..236S.
- [31] M.G. Rosenblum, A.S. Pikovsky, J. Kurths, "Phase Synchronization of Chaotic Oscillators", Physical Review Letters, V. 76, N. 11, 1996.
- [32] B.A. Cohen and A. Sances. Stationarity of the human electroencephalogram. Med Biol Eng Comput, 15:513–518, 1977.

- [33] D. Rudrauf,, A. Douiri, C. Kovach, J.P. Lachaux, D. Cosmelli, M. Chavez, C. Adam, B. Renault, J. Martinerie, M. Le Van Quyen, “Frequency flows and the time-frequency dynamics of multivariate phase synchronization in brain signals”, *NeuroImage*, 31, 209 – 227, (2006).
- [34] J.P. Lachaux, E. Rodriguez, J. Martinerie, F.J. Varela. “Measuring phase synchrony in brain signals”. *Hum. Brain Mapp.* 8 (4), 194–208, 1999.
- [35] J.P. Lachaux, E. Rodriguez, M. Le Van Quyen, A. Lutz, J. Martinerie, F.J. Varela. “Studying single-trials of phase-synchronous activity in the brain”. *Int J Bifurc Chaos*, 2000b.
- [36] P. Tass, M.G. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, H.J. Freund. “Detection of n:m phase locking from noisy data: application to magnetoencephalography”. *Phys. Rev. Lett.* 81 (15), 3291–3294, 1998.
- [37] M. Palus. “Detecting phase synchronization in noisy systems”. *Phys Lett A*, 255:341–51, 1997.
- [38] M. Le Van Quyen, J. Foucher, J.P. Lachaux, E. Rodriguez, A. Lutz, J. Martinerie, F.J. Varela, “Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony”, *Journal of Neuroscience Methods*, 111,83–98, (2001).
- [39] F. Mormann, K. Lehnertz, P. David, C.E. Elger. “Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients”. *Physica D*, 144, 358–369, (2000).
- [40] M. Breakspear. Nonlinear phase desynchronization in human electroencephalographic data. *Hum. Brain Mapp.* 15 (3), 175–198, 2002.
- [41] W.J. Freeman, L.J. Rogers. Fine temporal resolution of analytic phase reveals episodic synchronization by state transitions in gamma EEGs. *J. Neurophysiol*, 87 (2), 937–945, 2002.
- [42] J. Arnhold, et al., A robust method for detecting interdependences: application to intracranially recorded EEG. *Phys. D: Nonlin. Phenom.* 134 (1999).
- [43] F. Takens, D.A. Rand, L.S. Young. Detecting strange attractors in turbulence, *Dynamical systems and Turbulence, Lecture Notes in Mathematics*, 366–381, (1981).
- [44] J.A. Vastano and H.L. Swinney. Information transport in spatiotemporal system. *Phys Rev Letters*, 60:1773–1776, 1988.
- [45] J. Xu, Z. Liu, R. Liu, and Q. Yang. Information transmission in human cerebral cortex. *Physica D*, 106:363–374, 1997.
- [46] N.J.I. Mars, P.M. Thompson, and R.J. Wilkus. Spread of epileptic seizure activity in humans. *Epilepsia*, 26:85–94, 1985.
- [47] R. Guevara, J.L. Pérez Velazquez, V. Nenadovic, R. Wennberg, G. Senjanovic, L.G. Dominguez. “Phase Synchronization Measurements Using Electroencephalographic Recordings: What Can We Really Say About Neuronal Synchrony?”. *Neuroinformatics*, (2005).
- [48] G. Fein, J. Raz, F.F. Brown, E.L. Merrin. Common reference coherence data are confounded by power and phase effects. *Electroencephalogr. Clin. Neurophysiol.* 69, 581–584, (1988).
- [49] H. Zaveri, R.B. Duckrow, S.S. Spencer. The effect of a scalp reference signal on coherence measurements of intracranial electroencephalograms. *Clin. Neurophysiol.* 111, 1293–1299, (2000).

- [50] J. Bhattacharya, H. Petsche, U. Feldmann, B. Rescher. EEG gamma-band phase synchronization between posterior and frontal cortex during mental rotation in humans. *Neurosci Lett* 311 (1), 29–32, 2001.
- [51] P. Bob, M. Paluš, M. Šusta, K. Glaslová. EEG phase synchronization in patients with paranoid schizophrenia. *Neurosci Lett* 447 (1), 73 – 77, 2008.
- [52] K.M. Spencer, P.G. Nestor, M.A. Niznikiewicz, D.F. Salisbury, M.E. Shenton, R.W. McCarley. Abnormal neural synchrony in schizophrenia. *J Neurosci* 23 (19), 7407 – 7411, 2003.
- [53] G. Tcheslavski, A. Beex. Phase synchrony and coherence analyses of EEG as tools to discriminate between children with and without attention deficit disorders. *Biomed Signal Process Contr* 1 (2), 151 – 161, 2006.
- [54] L. Angelini, M.D. Tommaso, M. Guido, K. Hu, P.C. Ivanov, D. Marinazzo. Steady-state visual evoked potentials and phase synchronization in migraine patients. *Phys Rev Lett* 93 (3), 038103, 2004.
- [55] A. Galka. *Topics in Nonlinear Time Series Analysis*. World Scientific. 2000.
- [56] A. Mauricio F.L. Miranda de Sá, A. Fernando C. Infantosi, D.M. Simpson. “A Statistical Technique for Measuring Synchronism Between Cortical Regions in the EEG During Rhythmic Stimulation”. *IEEE Transactions on Biomedical Engineering*, VOL. 48, NO. 10, October 2001.
- [57] A. Mauricio F.L. Miranda de Sá, L.B. Felix, A. Fernando C. Infantosi. “A Matrix Based Algorithm for Estimating Multiple Coherence of a Periodic Signal and Its Application to the Multichannel EEG During Sensory Stimulation”. *IEEE Transactions on Biomedical Engineering*, VOL. 51, NO. 7, July 2004.
- [58] D.J. Thomson. “Spectrum estimation and harmonic analysis,” *Proc. of the IEEE*, vol. 70, no. 9, pp. 1055–1096, Sept 1982.
- [59] M. Hansson-Sandsten. “Cross-Spectrum And Coherence Function Estimation Using Time-Delayed Thomson Multitapers”. *IEEE, ICASSP*. 2011.
- [60] K. Mezeiova, M.Palus. “Comparison of coherence and phase synchronization of the human sleep electroencephalogram”. *Clinical Neurophysiology*, 123 (9) (2012) 1821-1830
- [61] G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach, M. Hallett. “Identifying true brain interaction from EEG data using the imaginary part of coherency”. *Clinical Neurophysiology*, 115, 2292–2307, (2004).
- [62] J.G. Stinstra, M.J. Peters. “The volume conductor may act as a temporal filter on the ECG and EEG”. *Med Biol Eng Comput*, 36:711–6, 1998.
- [63] C.J. Stam, G. Nolte, A. Daffertshofer. “Phase Lag Index: Assessment of Functional Connectivity From Multi Channel EEG and MEG With Diminished Bias From Common Sources”. *Human Brain Mapping* 28:1178–1193, (2007).
- [64] M.Vinck, R. Oostenveld, M. van Wingerden, F. Battaglia, C. Pennartz. “An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias”. *NeuroImage*, 55, 1548–1565, (2011).
- [65] S.M. Gordon, P.J. Franaszczuk, W.D. Hairston, M. Vindiola, K. McDowell. “Comparing parametric and nonparametric methods for detecting phase synchronization in EEG”. *Journal of Neuroscience Methods*. 212, 247–258, (2013).

- [66] G.V. Tcheslavski, A.A. Beex. "Phase synchrony and coherence analyses of EEG as tools to discriminate between children with and without attention deficit disorder". *Biomedical Signal Processing and Control*, 1, 151–161, (2006).
- [67] S.M. Kay. "Modern Spectral Estimation. Theory and Application". PrenticeHall Signal Processing Series, 1988.
- [68] C.E. Shannon, W. Weaver. "The Mathematical Theory of Information". Urbana, IL, University of Illinois Press, 1949.
- [69] K.V. Mardia, *Probability and Mathematical Statistics: Statistics of Directional Data*, Academic Press, London, 1972.
- [70] W. Ehm, J. Kornmeier. "Detecting Frequency-specific Phase Synchrony in Brain Electrical Oscillations". Institute for Frontier Areas of Psychology and Mental Health, Freiburg, Germany (2011)
- [71] S.Aviyente, E. M. Bernat, W.S. Evans, S.R. Sponheim. "A Phase Synchrony Measure for Quantifying Dynamic Functional Integration in the Brain". *Human Brain Mapping*, 32:80–93, (2011).
- [72] A. Rihaczek. "Signal energy distribution in time and frequency". *IEEE Trans Inform Theory* 14:369–374, 1968.
- [73] L. Cohen. "Time-Frequency Analysis: Theory and Applications". New Jersey, Prentice Hall, 1995.
- [74] Z. Zhou, B. Wan, D. Ming, H. Qi. "A novel technique for phase synchrony measurement from the complex motor imaginary potential of combined body and limb action". *J. Neural Eng.* 7, (2010).
- [75] C.M. Sweeny-Reed, S.J. Nasuto. "A novel approach to the detection of synchronisation in EEG based on empirical mode decomposition". *Journal of Computational Neuroscience*, vol. 23, no. 1, pp. 79–111, 2007.
- [76] N.E. Huang. "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis". *Proc. R. Soc.* 454 903–95, 1998.
- [77] D. Looney, C. Park, P. Kidmose, M. Ungstrup, D.P. Mandic. "Measuring Phase Synchrony Using Complex Extensions of EMD". *IEEE/SP 15th Workshop on Statistical Signal Processing*, 2009.
- [78] M. Sanaullah, "A Review of Higher Order Statistics and Spectra in Communication Systems". *Global Journal of Science Frontier Research, Physics and Space Science*, 13, 4, Version 1.0, 2013.
- [79] A.P. Petropulu. "Higher-Order Spectral Analysis". Drexel University, by CRC Press LLC, 1999.
- [80] C.L. Nikias, M.R. Raghuveer, "Bispectrum Estimation – A Digital Signal Processing Framework," *Proc. Of the IEEE*, vol. 75, pp. 869–891, 1987.
- [81] C.L. Nikias, A. Petropulu, "Higher-order spectra analysis: A nonlinear signal processing framework". Englewood Cliffs, NJ: Prentice Hall, 1993.
- [82] C.L. Nikias, J.M. Mendel, "Signal processing with higher-order spectra", *IEEE Signal Processing Magazine*, vol. 10, pp. 10–37, 1993.

- [83] J.W.A. Fackrell, S. Mc Laughlin. “DETECTING PHASE COUPLING IN SPEECH SIGNALS”. IEEE Colloquium Digest on Speech and Image Processing (London, England), pp. 4/1-4/8, 2 Ma y 1995 .
- [84] C.C. Chabalko. “Identification of Transient Nonlinear Aeroelastic Phenomena”. PhD Thesis in Engineering Mechanics, faculty of Virginia Polytechnic Institute and State University, 2007
- [85] Esmaeil Seraj and Reza Sameni, “Robust Electroencephalogram Phase Estimation with Applications in Brain-computer Interface Systems,” *Physiological Measurement*, vol. 38, no. 3, pp. 501–523, Feb. 2017. [Published Online] DOI: 10.1088/1361-6579/aa5bba
- [86] Reza Sameni and Esmaeil Seraj, “A Robust Statistical Framework for Instantaneous Electroencephalogram Phase and Frequency Analysis,” *Physiological Measurement*, vol. 38, no. 12, pp. 2141–2163, Nov. 2017. [Published Online] DOI: 10.1088/1361-6579/aa93a1
- [87] Esmaeil Seraj, “Cerebral Signal Instantaneous Parameters Estimation MATLAB Toolbox - User Guide Version 2.3,” arXiv Preprint, Dec. 2017 [Online]. Available: <https://arxiv.org/abs/1610.02249>
- [88] Foroozan Karimzadeh, Reza Boostani, Esmaeil Seraj, Reza Sameni, “A distributed classification procedure for automatic sleep stage scoring based on instantaneous electroencephalogram phase and envelope features,” *IEEE Transactions on Neural Systems & Rehabilitation Engineering*, vol. 26, no. 2, pp. 362-370, Feb. 2018. DOI: 10.1109/TNSRE.2017.2775058
- [89] Esmaeil Seraj, “An Investigation on the Utility and Reliability of Electroencephalogram Phase Signal upon Interpreting Cognitive Responses in the Brain: A Critical Discussion,” *Journal of Advanced Medical Sciences and Applied Technologies*, vol. 2, no. 4, pp. 299-312, Jan. 2017. [Published Online] DOI: <http://dx.doi.org/10.18869/2Fnrip.jamsat.2.4.299>
- [90] Esmaeil Seraj and Foroozan Karimzadeh, “Improved Detection Rate in Motor Imagery Based BCI Systems Using Combination of Robust Analytic Phase and Envelope Features,” 25th Iranian Conference on Electrical Engineering (ICEE), KNT University of Technology, Tehran, Iran, May 2017. [Published Online] DOI: 10.1109/IranianCEE.2017.7985458
- [91] Reza Sameni, The Open-Source Electrophysiological Toolbox (OSET), [Online] version 3.1 (2014). URL <http://www.oset.ir>
- [92] Foroozan Karimzadeh, Esmaeil Seraj, Reza Boostani, and Mohammad Torabi-Nami. ”Presenting efficient features for automatic CAP detection in sleep EEG signals.” *Telecommunications and Signal Processing (TSP)*, 2015 38th International Conference on. IEEE, 2015.