



**TECHNICAL UNIVERSITY OF CRETE  
DEPARTMENT OF ELECTRONIC AND COMPUTER  
ENGINEERING**

# **EEG Principal Component Organization via Time-Frequency Representation**

**DIPLOMA THESIS  
of  
Iordanidou Vasiliki**

**Supervisor: Professor Michalis Zervakis  
Chania, 2009**

To my beloved family.

# Contents

<b>CONTENTS .....</b>	<b>3</b>
<b>ABSTRACT .....</b>	<b>6</b>
<b>PREFACE.....</b>	<b>7</b>
<b>CHAPTER 1: INTRODUCTION.....</b>	<b>8</b>
<b>CHAPTER 2: ELECTROPHYSIOLOGY, AD PATHOLOGY AND EXPERIMENTAL TESTS .....</b>	<b>11</b>
<b>2.1 Brief Brain Anatomy.....</b>	<b>11</b>
<b>2.2 Evolution of EEG Research.....</b>	<b>14</b>
<b>2.3 EEG Procedure.....</b>	<b>15</b>
<b>2.4 EEG Artifacts .....</b>	<b>18</b>
2.4.1 Physiologic Artifacts.....	18
2.4.2 Extraphysiologic Artifacts .....	20
<b>2.5 EEG Signal Patterns Classification.....</b>	<b>21</b>
<b>2.6 Alzheimer's Disease .....</b>	<b>26</b>
2.6.1 AD Hallmarks .....	27
2.6.2 AD Stages and Symptoms .....	28
2.6.3 AD Causes and Risk Groups .....	30
2.6.4 AD Diagnosis and Treatment .....	31
<b>2.7 Event Related Potentials .....</b>	<b>32</b>
2.7.1 Evoked Potential Types.....	32
2.7.2 Important ERP Components .....	34
<b>CHAPTER 3: MATHEMATICAL ANALYSIS: BACKGROUND METHODS.....</b>	<b>36</b>
<b>3.1 Blind source separation.....</b>	<b>36</b>
3.1.1 Second-order Methods.....	37
3.1.1.1 Principal Component Analysis.....	37
3.1.1.2 Factor Analysis .....	38
3.1.2 Higher-order Methods.....	40
3.1.2.1 Projection Pursuit .....	40
3.1.2.2 Redundancy Reduction .....	40
<b>3.2 Independent Component Analysis .....</b>	<b>41</b>
3.2.1 ICA Relation to Classical Methods .....	42
3.2.2 ICA Applications in BSS .....	43
3.2.3 ICA Cost Functions and Algorithms.....	46

3.2.3.1 Likelihood and network entropy .....	46
3.2.4 ICA Optimization Algorithms .....	47
<b>3.3 Time-Frequency Representation .....</b>	<b>48</b>
<b>3.4 Wavelets .....</b>	<b>49</b>
3.4.1 Continuous Wavelet Transform .....	49
3.4.2 Continuous Wavelets vs Fourier Transform .....	51
<b>3.5 Partial Directed Coherence.....</b>	<b>52</b>
<b>3.6 Common Processing of ERP .....</b>	<b>53</b>
3.6.1 ERD-ERS .....	54
 <b>CHAPTER 4: PROPOSED METHODOLOGY.....</b>	 <b>58</b>
<b>4.1 Data acquisition and test description .....</b>	<b>58</b>
<b>4.2 Independent Component Extraction .....</b>	<b>59</b>
<b>4.3 Independent Component Selection .....</b>	<b>59</b>
4.3.1 Phase Intertrial Coherence (PIC).....	60
4.3.2 Phase-shift Intertrial Coherence (PsIC) .....	62
4.3.3 Utilization of Intertrial Coherence Metrics .....	63
<b>4.4 Functional Synchronization of Components .....</b>	<b>65</b>
4.4.1 Partial Directed Coherence (PDC).....	66
<b>4.5 Summary of Algorithmic Steps .....</b>	<b>67</b>
 <b>CHAPTER 5: RESULTS .....</b>	 <b>69</b>
<b>5.1 Application on Real Data .....</b>	<b>69</b>
5.1.1 First Control Subject Analysis .....	69
5.1.2 Second Control Subject Analysis .....	72
5.1.3 First AD Subject Analysis .....	74
5.1.4 Second AD Subject Analysis.....	76
<b>5.2 Statistical Analysis of Components.....</b>	<b>78</b>
5.2.1 Components .....	78
5.2.2 Component relations analysis .....	80
5.2.2.1 One way directed coherence .....	80
5.2.2.2 Both-ways directed (partial-directed) coherence .....	82
5.2.2.3 Either-way (one-way or both-ways) directed coherence .....	83
<b>5.3 Discussion of Results .....</b>	<b>88</b>
5.3.1 Control Subjects.....	89
5.3.2 AD Patients .....	91
 <b>CHAPTER 6: CONCLUSION AND FUTURE WORK.....</b>	 <b>93</b>
 <b>APPENDIX .....</b>	 <b>94</b>
 <b>REFERENCES .....</b>	 <b>102</b>

<b>ACKNOWLEDGMENTS.....</b>	<b>112</b>
-----------------------------	------------

# ABSTRACT

Over the past few years there has been an increased interest in studying the underlying neural mechanism of cognitive brain activity related to memory. In this direction, we study the brain activity based on its independent components instead of the EEG signal itself aiming towards identifying and analyzing induced responses being attributed to oscillatory bursts from local or distant neural assemblies, with variable latency and frequency, in an auditory working memory paradigm. The significance of the components is determined via intertrial coherence measures. The contribution and functional coupling of independent components to evoked and/or induced oscillatory activities is investigated through the concept of the recently introduced partial directed coherence method, which can also reveal the direction of the statistically significant relationships. The results on real data from an oddball experiment are in accordance with previous psychophysiology studies suggesting increased phase locked activity most prominently in the delta/ theta band, while alpha is also apparent in measures of non phase-locked activity. Dynamic synchronization is inferred between the alpha and delta bands, whereas some influence of the theta band is also detected. This study indicates that functional connectivity during cognitive processes may be successfully assessed using spectral power measures applied on independent components, which reflect distinct spatial patterns of activity.

# PREFACE

At this point we will give a preface of the thesis structure:

- **Chapter 1: Introduction:** In the introduction we explain this thesis scope, the need for this work and its novelties. We also give a brief summary of this thesis.
- **Chapter 2: Electrophysiology, AD Pathology and Experimental Tests:** In the second chapter of this thesis, there are presented the main definitions of important aspects of the analysis such as the brain, the EEG, and many more. Also the pathology of AD is briefly described so as to understand the differences an AD patient has from a person who does not have this pathology, as we also have AD cases in our analysis. Finally some experimental tests are introduced such as the popular  $P_{300}$ .
- **Chapter 3: Mathematical analysis: Background Methods:** In this chapter we present the mathematical background of this work. Mathematical tools and methodologies such as ICA and wavelets are demonstrated here.
- **Chapter 4: Proposed Methodology:** In the fourth chapter the methodology we follow in our analysis is step by step demonstrated. Examples and which justify the steps and their necessity are also presented.
- **Chapter 5: Results:** In chapter five of this thesis, are demonstrated the results on real experimental data of our method. In particular, four subjects' results are analytically presented and there follows a statistical analysis on all of the ten subjects of our analysis. Finally, there follows a discussion on these results and on the different cases of AD patients and control subjects.
- **Chapter 6: Conclusion and Further Work:** In the conclusion we close this thesis and we propose alternatives for further improvement of this thesis.

# CHAPTER 1: Introduction

During electroencephalographic (EEG) activity time-locked to an event, neuronal assemblies at different topographic locations (either local or distant) self-organize into transient networks, which synchronize in time and frequency to produce bursts of oscillations contributing to the observable EEG characteristics [123]. Such activity caused by an external or internal event can be categorized to either a phase-locked evoked response or a phase resetting of ongoing EEG activity, also referred to as induced response [30], [124], [98]. Thus, the transient oscillatory event-related activity, has been found to reflect the superimposed activity of several evoked and induced response sources, each having a distinct topographic organization [39], [99], [100]. More specifically, the so called event related potentials (ERPs) have been considered as originating from stable phase-locking due to transient synchronization of underlying neural substrates. In an event-related single trial it is expressed by an evoked response characterized by precise phase-locking to the stimulus onset. Consequently, ERPs can be detected by averaging single trial responses, as to increase the SNR of the low amplitude response in each single trial [36]. Furthermore, other studies have demonstrated the existence of event related oscillations (EROs) as responses from different bands that are non phase-locked to the stimulus. These induced responses have been attributed to oscillatory bursts from local or distant neural assemblies, with variable latency and frequency from trial to trial [30], [36], or to a phase resetting of pre-stimulus EEG activity [98], [99], [100]. This non phase-locked induced activity is quantified differently than that of ERPs, with techniques primarily based on power averaging over specific frequency bands of interest or by the power of the AM demodulated signal after subtraction of the phase-locked activity [29], [35], [93]. This is motivated by the time-locked nature of the envelope of induced activity. In particular, an event-triggered decrease (increase) in alpha power is termed as event related de-synchronization (synchronization), or alpha ERD/ERS, respectively [30]. The events giving rise to such frequency specific band changes of amplitude (and power) in the spontaneous EEG can be either externally (any kind of stimulation) or internally paced (voluntary movement or visual structure binding) [93].

Because of their neurophysiologic origins, phase-locked (evoked) and non phase-locked (induced) responses are different [30] and have different functional roles [39], even though they may reflect similar cognitive events and may correlate in their various parameters. Several senso-motor or cognitive tasks produce a variety of event related activities, both phase-locked and non phase-locked to the event [39], [124], [98]. For instance, periodic visual stimulation with object recognition experiments have demonstrated the induction of alpha rhythm and alpha non-phase locked activity (event related de-synchronization) maximized at central and lateral posterior regions [30], [98], [93], early visually evoked ERP component at gamma band localized in central regions [36], [39], a visual short-memory ERP component at occipito-temporal regions (possibly linking vision with working memory) , object-binding gamma oscillations that are not phase-locked to the event and express wide spread topographies [39], theta band oscillations contributing to P300 ERP for oddball experiments [39], [25], [29], just to mention a few of the reported components.



Furthermore, due to their different neurophysiological origin, phase-locked evoked activity (ERP) and ongoing EEG rhythms of induced oscillations (such as ERD/ERS) may be considered as coupled processes progressing in time with different spatial localization of activated neuronal assemblies and partially overlapping frequency content [30]. In essence, pre-stimulus EEG activity may remain on post-stimulus response and can affect the characteristics (amplitude, latency, frequency) of the evoked and induced responses.

It becomes obvious that event related brain dynamics entail a variety of activations and oscillations, from phase resetting of ongoing EEG activity in the alpha and theta bands [99] to phase-locked evoked and non phase-locked induced oscillations especially in delta, theta and gamma bands [30], [36]. Their origins relate to multiple task conditions and many stimulus types engaged during the event presentation and execution of its consequent actions [100], which define distinct brain functions, some operating independently and some being coupled. Thus, the separation and analysis of independent activities of different nature and origin is of primary importance in considering alterations in EEG recordings due to brain pathologies or in developing algorithms for brain-computer interfaces. The separation of EEG components has been approached through several signal decomposition methods, including time-frequency [39], time-space [98], or even three way (time-frequency-space) techniques [123]. Efficient decomposition frameworks include wavelet analysis and independent component analysis (ICA) [125] followed by a variety of methods to characterize the nature of derived components in terms of their time/frequency activity and topographic origin [123], [100].

As already mentioned, the analysis of averaged ERPs has been preferred for the consideration of phase-locked activities, since the averaging increases the signal to noise ratio. Furthermore, the analysis of band-pass power content averaged over trials has been extensively used for the study of event induced but not phase-locked activities. Independent Component Analysis (ICA) [98] has been successfully applied on continuous or event related EEG to decompose it into a sum of spatially fixed and temporally independent components that can lead in different spatial distribution patterns, which in turn may be directly attributed to underlying cortical activity. Single-trial ERP analysis is gaining grounds over other methods for the analysis of evoked phenomena on an individualized basis, despite the problems of low SNR and instability of algorithms due to the small number of training samples [39], [98]. Single-trial ERP component analysis has been used primarily for relating a specific task (sensory, motor or cognitive) with the topography and/or frequency content of its components [126]. However, little research effort has been directed towards annotating the contribution of components to phase or non phase-locked phenomena, which can trigger the characterization of sub-activities involved in the performance of a task (e.g. attention, visual cortex organization, binding effects, working memory, etc.) [39], [98] and further facilitate brain-computer interfaces. Even less effort has been devoted to the functional synchronization of these components, which may provide a means of brain organization and synchronization of neural assemblies during the performance of a task [127]. Studies with more detailed MEG signals have revealed local synchronization patterns and cortico-cortical interactions involved in several cognitive operations [38], with composite subtasks being triggered within different brain regions by unitary brain sources and subsequently synchronize to complete the task. Thus, the dynamics of interaction among independent components (rather than among channels, as it is often considered) may be used for indexing

neural synchrony of such local or distant brain sources [128]. An additional reason in support of evaluating synchronization on independent components relates to artifacts produced on electrodes by the volume conduction effect [127]. The EEG recordings on the scalp electrodes form mixings or superpositions of the true source signals, which can lead to fictitious (spurious) synchronization. The components of ICA instead, derived from linear un-mixing transforms, remove such spurious synchronization but preserve the real synchronization of the sources, which act as coupled oscillators [129]. Synchronization among stimulus-locked components can be studied using pair-wise linear (cross-coherence or phase-coherence) [128], [130] or nonlinear dynamics and models [131].

The above considerations necessitate new directions of analysis, besides the study of components in terms of their spatial, temporal and oscillatory nature. In this paper we study the involvement of several brain sources in performing a working memory cognitive task. Our working assumption is that the performance of the task triggers certain evoked and induced responses, as expressed by synchronization between different neural assemblies. Since we are interested in identifying distinct signal components and analyzing their coupling, we focus on decomposing the EEG signal into ICA components and then analyzing their time-frequency content along with their spatial distribution. In addition, besides the study of their content, we consider their phase-locking characteristics across the experimental trials. After the characterization of significant components, we consider their dynamic coupling (synchronization) by deriving and exploiting information on the “driver and response” relationship between observations [132]. The first aspect of component characterization has been partially addressed with measures that can reveal phase locking effects [39], [98]. Besides these measures, we introduce a metric for considering stimulus-locked but not phase-locked activity. The second aspect related to synchronization is addressed through both linear and nonlinear synchronization measures to reveal coupling characteristics.

For the purpose of this analysis we use an auditory oddball experiment. This experimental set-up is expected to produce both phase-locked oscillations, especially in the theta and delta bands related to P300 activity (including P3a and P3b components [85]), and non phase-locked (induced) oscillatory activity, particularly related to alpha-range event related desynchronization (ERD). The P300 ERP wave is elicited after novel or task relevant stimuli requiring cognitive effort from the participant. It is composed of multiple temporally-overlapping components related to orientation of attention to the stimuli (P3a) and recognition with memory processes (P3b) [133]. It has shown remarkable promise as an indicator of human cognition in normals, as well as clinical populations [30]. Furthermore, non phase-locked ERD is elicited by oddball experiments in various frequency bands [20], [29], with most prominent activity in the alpha band. Thus, we expect to recover a variety of components, consistently contributing to the recorded single-trial signals. Altogether alpha, theta and delta bands represent the set of processes that activate in sequence during oddball processing.

# CHAPTER 2: Electrophysiology, AD Pathology and Experimental Tests

Brain is considered to be one of the most complicated parts of the body. It would not be an exaggeration to say that it is the seat of intelligence, interpreter of the senses, initiator of body movement, and controller of behaviour. Lying in its bony shell and washed by protective fluid, the brain is the source of all the qualities that define our humanity.

Scientists are for centuries trying to explore the secrets of the brain, which though until recently seemed to be incomprehensible. Now, however, modern technology's achievements have been very helpful because the researchers have the opportunity to "observe" the brain and its functionalities through different kinds of brain scans. Some of the most important technologies used for brain scan are Electroencephalogram (EEG), Magnetic Resonance Imaging (MRI), Functional Magnetic Resonance Imaging (fMRI), Magnetoencephalography (MEG), Computerized Axial Tomography (CAT) and Positron Emission Tomography (PET).

The EEG is the recording of the electrical activity derived from the brain with electrodes placed on the skull. The signals recorded can be used for clinical and research reasons. EEG can also be used for diagnostic purposes in pathological cases (Alzheimer disease (AD), epilepsy, etc.) and brain-computer interfaces (BCI).

## 2.1 Brief Brain Anatomy

1.5 kilograms and several thousand miles of interconnected nerve cells (about 100 billion) control every movement, thought, sensation, and emotion that comprises the human experience. Within the brain and spinal cord there are ten thousand distinct varieties of neurons, trillions of supportive cells, a few more trillion synaptic connections, a hundred known chemical regulating agents, miles of minuscule blood vessels, axons ranging from a few microns to well over a foot and a half in length, and untold mysteries of how -almost flawlessly- all these components work together [1], [2].

The main parts of a brain can be observed in the following figure:

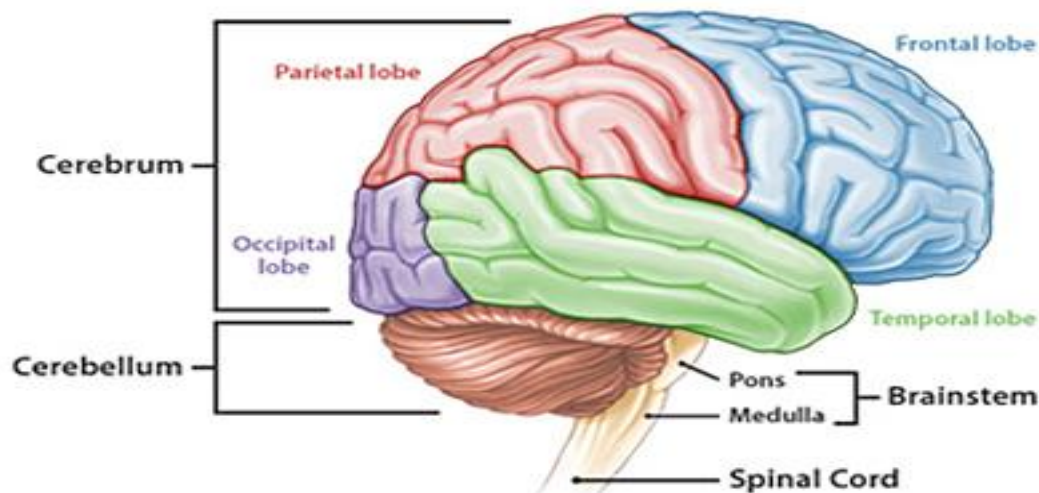


Fig.2.1: A brain anatomy [15].

The parts of the brain which can be clearly distinguished are the followings:

- Frontal lobe
  - Parietal lobe
  - Occipital lobe
  - Temporal lobe
  - Cerebellum
  - Pons
  - Medula
- } Cerebrum
- } Brainstem

- ***Frontal lobe:*** It is located at the front of each cerebral hemisphere and positioned anterior to (in front of) the parietal lobes and above and anterior to the temporal lobes. The frontal lobe reaches full maturity around age 25, marking the cognitive maturity associated with adulthood. The executive functions of the frontal lobe involve the ability to recognize future consequences resulting from current actions, to choose between good and bad actions (or better and best), override and suppress unacceptable social responses, and determine similarities and differences between things or events. Therefore, it is involved in higher mental functions. It also plays an important part in retaining longer term memories which are not task-based [3]. More specifically it is associated with reasoning, motor skills, higher level cognition, and expressive language [5].
- ***Parietal lobe:*** It is positioned above (superior to) the occipital lobe and behind (posterior to) the frontal lobe. The parietal lobe integrates sensory information from different modalities, particularly determining spatial sense and navigation. It also plays important roles in integrating sensory information from various parts of the body such as pressure, touch, and pain, knowledge of numbers and their relations, and in the manipulation of objects. Portions of the parietal lobe are involved with visuospatial processing [4], [5], [6].
- ***Occipital lobe:*** The occipital lobe is located in the rearmost portion of the skull. The occipital lobe is the smallest of four lobes in the human cerebral cortex and is defined as the part of the cerebral cortex that lies underneath the occipital bone. The most important functional aspect of the occipital lobe is that it contains the primary visual cortex. The occipital lobe is divided into several functional visual areas. Each visual area contains a full map of the visual world. Although there are no anatomical markers distinguishing these areas, physiologists have used electrode recordings to divide the cortex into different functional regions [5], [7], [8].
- ***Temporal lobe:*** It is a region of the cerebral cortex that is located beneath the Sylvian fissure (Sylvian fissure divides the frontal lobe and parietal lobe above from the temporal lobe below) on both the left and right hemispheres of the brain. The temporal lobe is involved in

auditory processing and is home to the primary auditory cortex. It is also important for the processing of semantics in both speech and vision. The temporal lobe plays a key role in the formation of long-term memory and control of spatial memory and behavior [9].

- ***Cerebellum***: The cerebellum (Latin for little brain) is a region of the brain that plays an important role in the integration of sensory perception, coordination and motor control. In order to coordinate motor control, there are many neural pathways linking the cerebellum with the cerebral motor cortex (which sends information to the muscles causing them to move) and the spinocerebellar tract (which provides proprioceptive feedback on the position of the body in space). The cerebellum integrates these pathways, like a train conductor, using the constant feedback on body position to fine-tune motor movements [10]. Modern research shows that the cerebellum has a broader role in a number of key cognitive functions, including attention and the processing of language, music, and other sensory temporal stimuli [11].
- ***Pons***: It is above the medulla and anterior to the cerebellum. The pons relays sensory information between the cerebellum and cerebrum, aids in relaying other messages in the brain, controls arousal, and regulates respiration. In some theories, the pons has a role in dreaming [12].
- ***Medula***: It is the last portion of the brain before the spinal cord. It refers to the middle of something, and derives from the Latin word "marrow" (the Latin equivalent of the Greek stem myelo). It contains nerve tracts of the motor and sensory pathways. It also contains autonomic centers for regulating heart rate, vasomotion, and respiratory rhythmicity [13].

## 2.2 Evolution of EEG Research

After the brief introduction, we can understand that it is of great importance to record and decode the complicated brain activity from the different parts of the brain so as to understand which parts of the brain are activated in which activity, what kind of activations we have so as to recognize pathological cases, create brain-computer interfaces (BCI) and try to understand the way the brain reacts in different circumstances. The EEG is one of the most sufficient methods so as to record the brain activity.

A first course on the EEG covers its origins and evolution throughout time. A physician named Richard Caton in 1875 was the first to describe sensory evoked response and observed ‘continuous spontaneous electrical activity’ from the brain surface. In 1912 Pravdich-Neminsky, a Russian physiologist, recorded EEG and evoked potentials on photographs and coined the term ‘electrocerebrogram’. German physiologist and psychiatrist Hans Berger (1873–1941) began his studies of the human EEG in 1920. He gave the device its name and is sometimes credited with inventing the EEG, though others had performed similar experiments. His work was later expanded by Edgar Douglas Adrian. In 1934, Fisher and Lowenback first demonstrated epileptiform spikes. In 1935 Gibbs, Davis and Lennox described interictal spike waves and the 3 cycles/s pattern of clinical absence seizures, which began the field of clinical electroencephalography. Subsequently, in 1936 Gibbs and Jasper reported the interictal spike as the focal signature of epilepsy. The same year, the first EEG laboratory opened at Massachusetts General Hospital. Franklin Offner (1911–1999), professor of biophysics at Northwestern University developed a prototype of the EEG which incorporated a piezoelectric inkwriter called a Crystograph (the whole device was typically known as the Offner Dynograph). In 1947, The American EEG Society was founded and the first International EEG congress was held. In 1953 Aserinsky and Kleitman describe REM sleep. In the 1950s, William Grey Walter developed an adjunct to EEG called EEG topography which allowed for the mapping of electrical activity across the surface of the brain. This enjoyed a brief period of popularity in the 1980s and seemed especially promising for psychiatry. It was never accepted by neurologists and remains primarily a research tool.

It was though in the 1990s when the increasing use of EEG combined with neuroimaging techniques emerged. Real-time digital EEG monitoring for critical care became increasingly important in intensive care units, operating rooms and emergency rooms [14].

## 2.3 EEG Procedure

In order to take the final EEG signal there is a procedure which has to be followed. In conventional scalp EEG, the recording is obtained by placing electrodes on the scalp with a conductive gel or paste, usually after preparing the scalp area by light abrasion to reduce impedance due to dead skin cells.

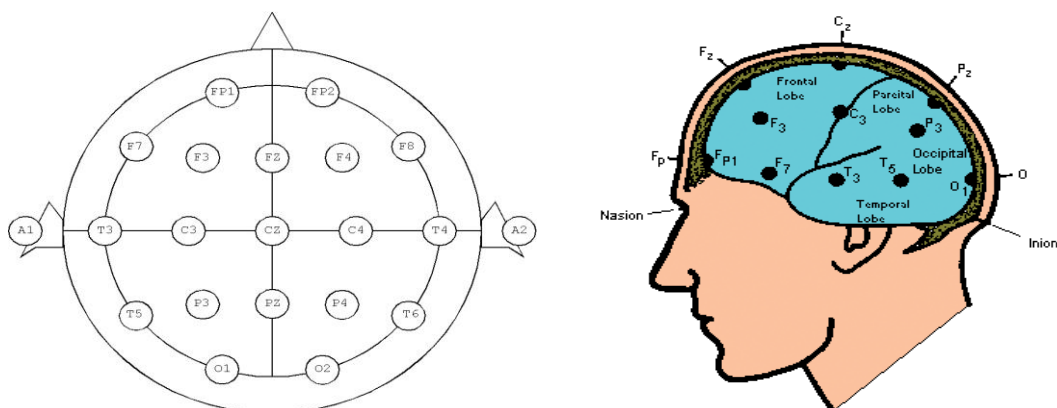


**Fig. 1.2: EEG Electrodes on the skull.**



**Fig. 1.3: EEG electrode types.**

Electrode locations and names are specified by the International 10–20 system for most clinical and research applications (except when high-density arrays are used). This system ensures that the naming of electrodes is consistent across laboratories. The “10” and “20” refer to the fact that the actual distances between adjacent electrodes are either 10% or 20% of the total front-back or right-left distance of the skull. In most clinical applications, 19 recording electrodes (plus ground and system reference) are used. A smaller number of electrodes are typically used when recording EEG from neonates. Additional electrodes can be added to the standard set-up when a clinical or research application demands increased spatial resolution for a particular area of the brain. High-density arrays (typically via cap or net) can contain up to 256 electrodes more-or-less evenly spaced around the scalp [17].



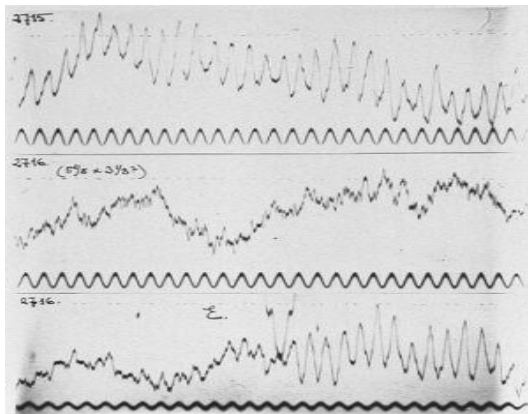
**Fig. 1.4 a,b; 10-20 International electrode location system [16].**

Each electrode is connected to one input of a differential amplifier (one amplifier per pair of electrodes). The naming of the electrodes has the following reasoning (according to the place of the scalp the electrodes are localized):

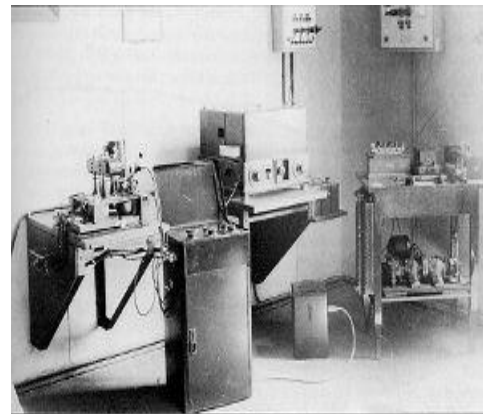
- F - Frontal
- T - Temporal
- C - Central
- P – Parietal
- O – Occipital

The electrodes placed in the midline are followed by the letter “z” (zero). Also, even numbers (2,4,6,8) refer to electrode positions on the right hemisphere, whereas odd numbers (1,3,5,7) refer to those on the left hemisphere [17].

A common system reference electrode is connected to the other input of each differential amplifier. These amplifiers amplify the voltage between the active electrode and the reference (typically 1,000–100,000 times, or 60–100 dB of voltage gain). In analog EEG, the signal is then filtered, and the EEG signal is output as the deflection of pens as paper passes underneath. Most EEG systems these days, however, are digital, and the amplified signal is digitized via an analog-to-digital converter, after being passed through an anti-aliasing filter. Analog-to-digital sampling typically occurs at 256-512 Hz in clinical scalp EEG, though sampling rates of up to 20 kHz are used in some research applications [18].



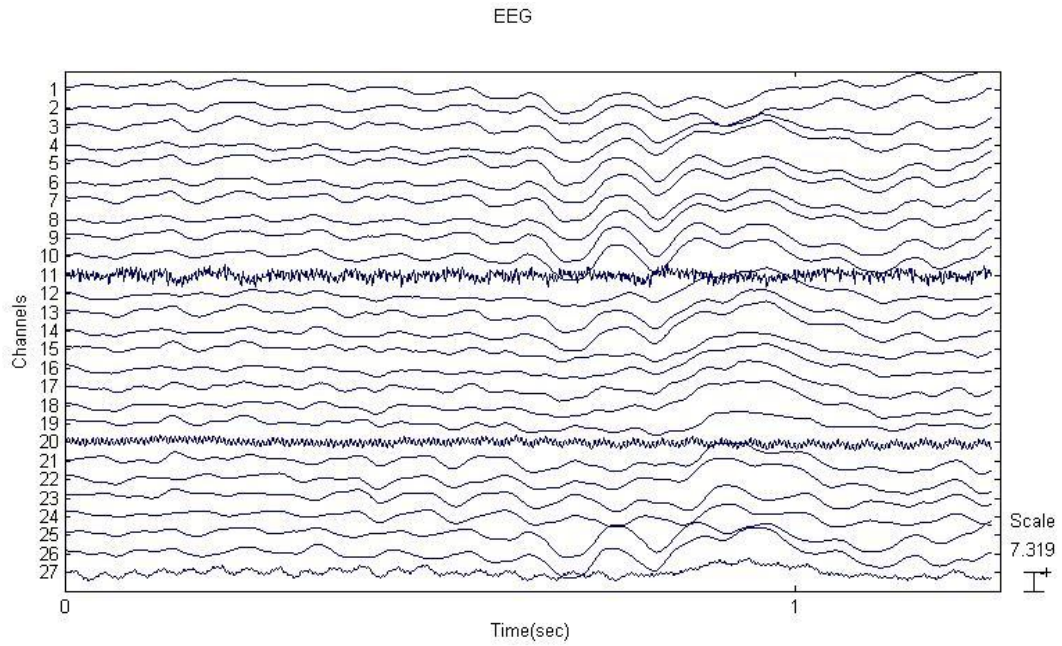
**Fig. 1.5: Berger Examples of EEG.**



**Fig. 1.6: Berger's system for recording EEG, 1926.**

Usually, during the recording, a series of activation procedures are used. These procedures may induce normal or abnormal EEG activity that might not otherwise be seen. These procedures include hyperventilation, photic stimulation, eye closure, mental activity, sleep and sleep deprivation.





**Fig. 1.7: An 1.27sec. EEG recording consisted of 27 channels.**

For each electrode we have a recording which contains information in the time domain for the brain activity, in the specific location the electrode is placed. So as to have a full “picture” of the brain activity, we have to examine the recordings of all the electrodes. It has to be noted that the information taken from the electrodes most of the times has artifacts. Artifacts give information not relative to the events examined in the EEG and are produced by eye closure, heart beating etc. Often there is a confusion in EEG analysis because of the artifacts, though nowadays methods for the recognition and rejection/reconstruction of artifact components have been introduced.

## 2.4 EEG Artifacts

As it has been noted before in most of the EEG signals artifact patterns can be recognized. Artifacts are electrical signals detected along the scalp by an EEG, but that originate from non-cerebral origin. The amplitude of artifacts can be quite large relative to the size of amplitude of the cortical signals of interest. This is one of the reasons why it takes considerable experience to correctly interpret EEGs clinically. Artifacts can be divided into physiologic and extraphysiologic artifacts. While physiologic artifacts are generated from the patient, they arise from sources other than the brain (body), extraphysiologic artifacts arise from outside the body (equipment, environment, etc).

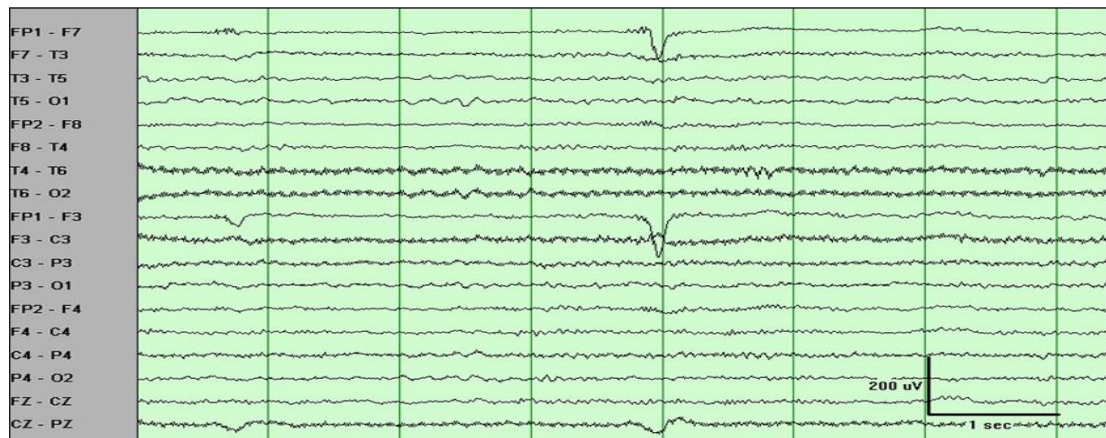
### 2.4.1 Physiologic Artifacts

Physiologic artifacts are those caused by normal functional activities of the subject. Some of the most common physiologic artifacts are the followings:

- ***Eye artifacts:*** On all EEGs eye movements are observed and are useful in identifying sleep stages. The eyeball acts as a dipole with a positive pole oriented anteriorly (cornea) and a negative pole oriented posteriorly (retina). When the globe rotates about its axis, it generates a large-amplitude alternate current field, which is detectable by any electrodes near the eye. The other source of artifacts comes from electromyogram (EMG) potentials from muscles in and around the orbit. Vertical eye movements typically are observed with blinks (Bell phenomenon). A blink causes the positive pole to move closer to frontopolar electrodes, producing symmetric downward deflections. During downward eye movement the positive pole of the globe moves away from frontopolar electrodes, producing an upward deflection best recorded in channels 1 and 5 in the bipolar longitudinal montage [19].
- ***Electrocardiogram (ECG) artifacts:*** The field of the heart potentials over the surface of the scalp often causes some individual variations in the amount and persistence of ECG artifact. Generally, people with short and wide necks have the largest ECG artifacts on their EEGs. The voltage and apparent surface of the artifact vary from derivation to derivation and, consequently, from montage to montage. The artifact is observed best in referential montages using earlobe electrodes A1 and A2. ECG artifact is recognized easily by its rhythmicity and coincidence with the ECG tracing. The situation becomes difficult when cerebral abnormal activity (eg, sharp waves) appears intermixed with EEG artifact, and the former may be overlooked. The EEG technologist should apply electrodes routinely to record the ECG [19].
- ***Electromyogram (EMG) artifacts:*** Myogenic potentials are also common artifacts. Frontalis and temporalis muscles such as muscles of the face, the neck or the scull are common causes. Generally, the potentials generated in the muscles are of shorter duration than those generated in the brain and are identified easily on the basis of duration, morphology, and rate of firing (frequency). Particular patterns of

electromyogram (EMG) artifacts can occur in some movement disorders. Rhythmic 4 to 6Hz sinusoidal artifacts that may mimic cerebral activity can be produced Essential tremor and Parkinson disease [19].

- ***Glossokinetic artifacts:*** In addition to muscle activity, the tongue functions as a dipole. In this case, the tip of the tongue is the most important part because it is more mobile. The artifact produced by the tongue has a broad potential field that drops from frontal to occipital areas, although it is less steep than that produced by eye movement artifacts [19].
- ***Respiration artifacts:*** Two kinds of artifacts can be produced by respiration. One type is in the form of slow and rhythmic activity, synchronous with the body movements of respiration and mechanically affecting the impedance of one electrode. The other type can be slow or sharp waves that occur synchronously with inhalation or exhalation and involve those electrodes on which the patient is lying [19].
- ***Skin artifacts:*** Biological processes and defects on the skin surface may alter impedance and cause artifacts. Sweat is a common cause because sodium chloride and lactic acid from sweating may react with metals of the electrodes and produce huge slow baseline sways. Skull defects also can be the source of asymmetry. In this situation, amplitudes are greater in derivations from electrodes overlying or adjacent to skull defects [19].



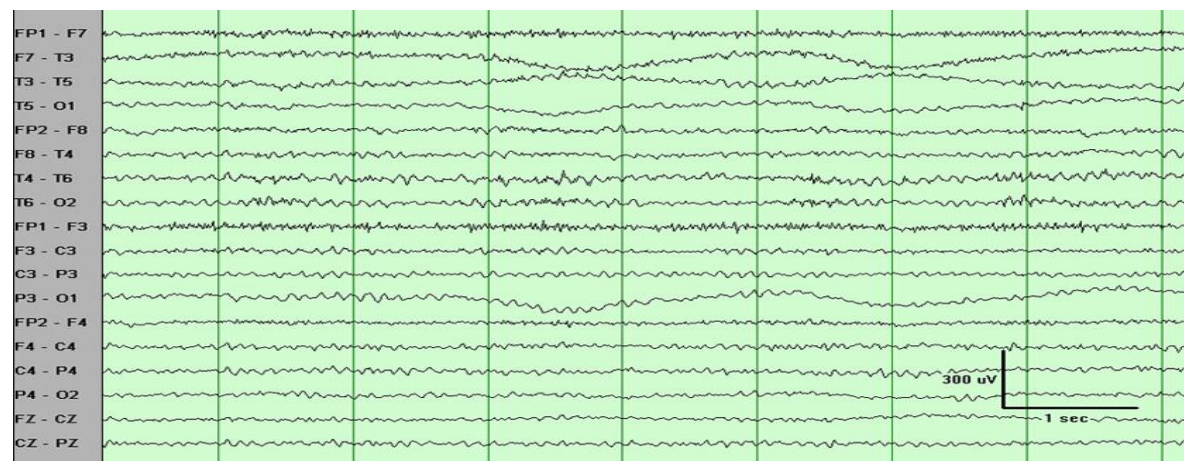
**Fig. 1.7: Eye Artifact. [19].**

In the figure 1.7 we have a left frontal artifact in the fourth second. This is not limited to a single electrode and has the morphology of an eye movement, but it is unilateral. This is an eye movement in a patient who has a glass left eye.

## 2.4.2 Extraphysiologic Artifacts

Extraphysiologic artifacts are those caused by the environment of the examination. Some of the most common extraphysiologic artifacts are the followings:

- **Electrodes artifacts:** The most common electrode artifact is the electrode popping. Morphologically this appears as single or multiple sharp waveforms due to abrupt impedance change. It is identified easily by its characteristic appearance and its usual distribution, which is limited to a single electrode [19].
- **Altering current artifact:** Adequate grounding on the patient has almost eliminated this type of artifact from power lines. The problem arises when the impedance of one of the active electrodes becomes significantly large between the electrodes and the ground of the amplifier. In this situation, the ground becomes an active electrode that, depending on its location, produces the 60-Hz artifact [19].
- **Other artifacts:** Movement of other persons around the patient, electrostatic changes on the drops, or interference from high-frequency radiation from radio, television, hospital paging systems, and other electronic devices can also cause artifacts on the EEG. Those artifacts though are not specific and easily recognised so they are usually omitted [19].



**Fig. 1.8: Sweat artifact. [19].**

In the figure above we have the effect of sweat in some electrodes of the EEG. This is characterized by very low-frequency (here, 0.25- to 0.5-Hz) oscillations. The distribution here (midtemporal electrode T3 and occipital electrode O1) suggests sweat on the left side.

## 2.5 EEG Signal Patterns Classification

The EEG contains frequency components that can be measured and analyzed, and these frequency components have interesting and valuable properties. A great deal of history is involved in the definition, naming, and use of these frequency bands. They are named using Greek letters, a convention that was begun by Hans Berger. He observed all of the rhythms known today (except the 40 Hz “gamma” band), and described many of their basic properties. Since then, our definitions and understandings of the rhythms have been refined.

Nowadays, the patterns seen in the EEG, initiating from changes in the frequency amplitude with respect to time, have been classified into the internationally standard bands of delta, theta, alpha, beta, gamma which are more or less empirical frequency limits. Since the arousal of EEG analysis these patterns have been related with different brain arousal states, functions or pathologies. In different individuals these patterns are found to vary, though the frequency seems to remain the same. This note was the reason why the frequency bands were defined in a strict way. A more detailed description of the frequency bands follows.

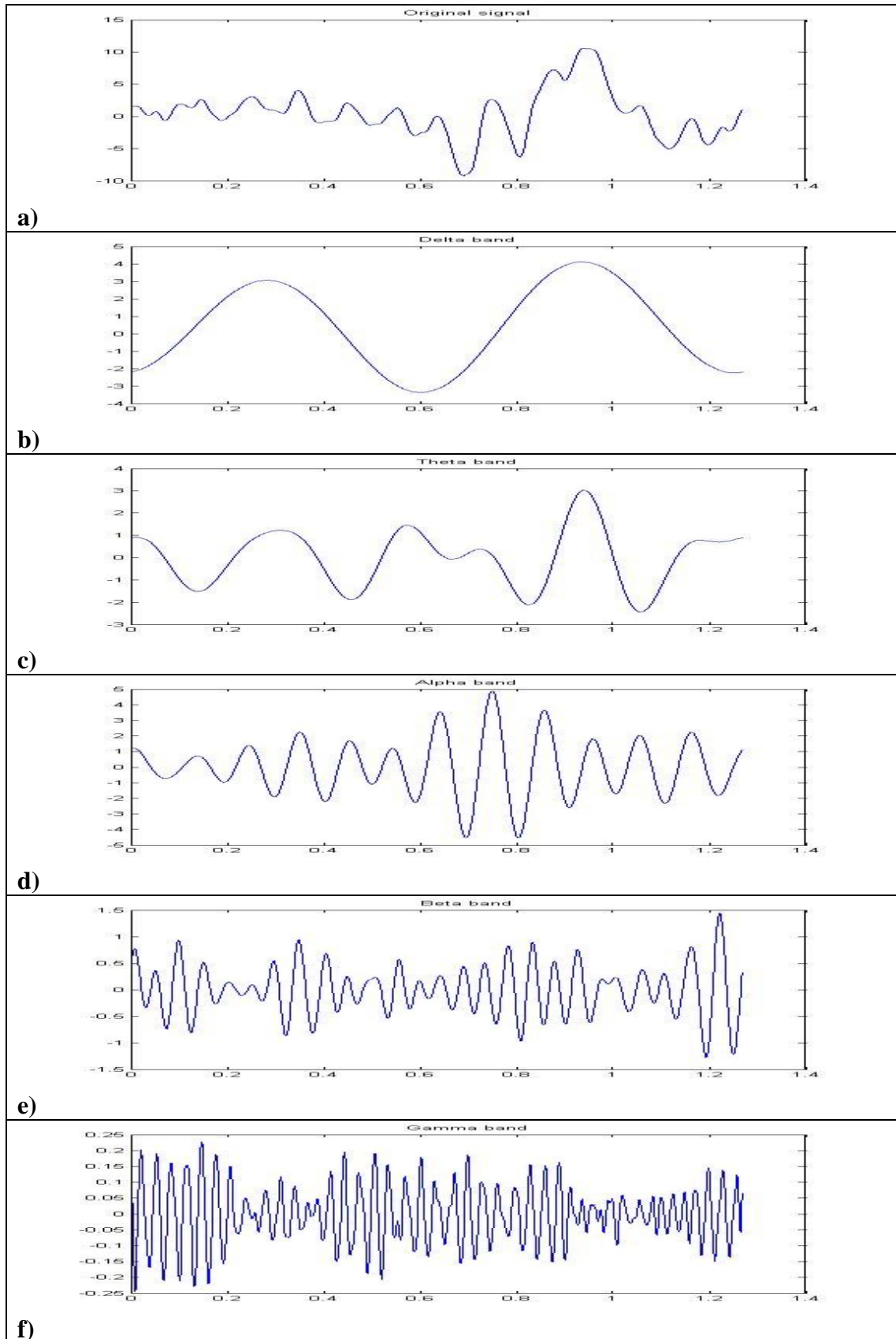
- ***Delta band (0-3Hz):*** It tends to be the highest in amplitude and the slowest waves. It is seen normally in adults in slow wave sleep. It is also seen normally in babies. Past developmental studies showed that delta percentage gradually decreases and, at the same time, faster rhythms increase with age. Instead, in healthy adults, delta band is strongly reduced during awaking states, is typically observed in the deepest stages of sleep and is considered a marker of brain sufferance or pathological condition when it appears in the waking brain [134]. Several studies found that slow wave activity, particularly in the range of delta frequencies, marks pathological brain abnormality resulting from neurological damage, such as cerebral infarct, contusion, local infection, tumor or subdural hematoma [134]. It is usually most prominent frontally in adults and posteriorly in children [18].
- ***Theta band (3-7Hz):*** Theta is seen normally in young children. It may be seen in drowsiness or arousal in older children and adults; it can also be seen in meditation [24]. Theta activity has been associated with reports of relaxed, meditative, and creative states [18]. Theta band seems to associate with working memory functions and correlates with memory load, task demand and episodic (short-term) memory processes [21]. It also relates to successful encoding of new information and attention. In increasing task demands phasic power increase in Theta ERS and ERD is observed [22]. Frontal Theta correlates with mental tasks requiring attention (arithmetic, reasoning). On the contrary excess theta for age represents abnormal activity. It can be seen as a focal disturbance in focal subcortical lesions; it can be seen in generalized distribution in diffuse disorder or metabolic encephalopathy or deep midline disorders or some instances of hydrocephalus [18]. Moreover, language problems relate to working memory and Theta ERD/ERS. Background Theta increases in epilepsy and more specifically children with epilepsy and dyslexia seem to have alterations on Theta response suggesting problems in audio or visual short-term memory encoding [23].

- **Alpha band (7-12Hz):** Hans Berger named the first rhythmic EEG activity he saw, the “alpha wave”. This is activity in the 8-12 Hz range seen in the posterior regions of the head on both sides, being higher in amplitude on the dominant side. It is brought out by closing the eyes and by relaxation. It was noted to attenuate with eye opening or mental exertion. This activity is now referred to as “posterior basic rhythm”, the “posterior dominant rhythm” or the “posterior alpha rhythm” [18]. Alpha band has been associated with attention and cognitive processes requiring access and retrieval of information from long term memory (semantic memory) [31]. Event related Alpha response is auditory stimuli. Encoding of auditory material such as learning produces Alpha event-related synchronization (ERS) [32]. Auditory recognition or comparison (retrieval) as well as cognitive processes such as mental arithmetic, reading, tasks conditions requiring attention and memory processes brings out Alpha event-related desynchronization (ERD) [30]. Alpha ERD (suppression) is also induced by sensory (auditory visual and somatosensory) stimulation. Lower Alpha reflects phasic alertness and has ERD in warning signal. Intermediate Alpha reflects expectancy and desynchronizes seconds before target or non-target appears [29]. Upper Alpha ERD appears only after the target something which reflects performance of task (counting of targets). Thus upper Alpha ERD is associated with processing of sensory-semantic information [26] and lower Alpha ERD reflects attentional processes [27], [28]. Alpha can be abnormal, for example, an EEG that has diffuse alpha occurring in coma and is not responsive to external stimuli is referred to as "alpha coma". An alpha asymmetry always implies brain damage in young subjects or remote infarction in older subjects [18].
- **Beta band (12-30Hz):** It is seen usually on both sides in symmetrical distribution and is most evident frontally. Low amplitude beta with multiple and varying frequencies is often associated with active, busy or anxious thinking and active concentration [18]. More specifically increased Beta is observed in reading and subtraction and Beta oscillatory activity in movement related tasks [33]. Rhythmic beta with a dominant set of frequencies is associated with various pathologies and drug effects, especially benzodiazepines. Activity over about 25 Hz seen in the scalp EEG is rarely cerebral (most often artifactual). It may be absent or reduced in areas of cortical damage. It is the dominant rhythm in patients who are alert or anxious or who have their eyes open [18]. Lower background Beta increases in epilepsy [34].
- **Gamma band (30-100Hz):** Because of the filtering properties of the skull and scalp, Gamma rhythms are usually recorded by electrocorticography or possibly magnetoencephalography. Gamma is seen in occipital sites in visual tasks (posterior sites), in central sites in motor tasks, in posterior sites in reading, in occipital, temporal, parietal sites in word reading and in frontal, central sites in subtraction/mental task. High frequency oscillations (Gamma) are usually low voltage and short lived (100-300 ms) and except an early component they are not

phase-locked to stimulus. These stimulus induced oscillations are probably generated internally via a self-organizing process among stimulus-driven local oscillators that are mutually connected [35]. They are complementary to stimulus-locked evoked oscillations that are directly driven by stimulus transients and which are likely to play a role at the primary information processing stages. Synchronized neural firing in the Gamma band is associated with the binding problem (bottom-up process) [38]. Demonstrated involvement of Gamma band activity in a wider range of top-down driven cognitive processes is suggesting that Gamma synchronization of cortical cell assemblies are involved in cognitive functions. Rhythmic synchronization of neural discharges in Gamma may provide the necessary spatial and temporal links that bind together the processing in different brain areas to build a coherent percept and more generally for the construction of object representations [36]. Induced Gamma activity is observed in response to sensory stimuli and during motor task [37]. Visual attention (as compared to not paying attention to a picture) also increases Gamma band [39].

In the following figures (fig. 1.9a, b, c, d, e, f), we have a real EEG signal (fig. 1.9a), which then has been filtered so as to take its delta (fig. 1.9b), theta (fig. 1.9c), alpha (fig. 1.9d), beta (fig. 1.9e) and gamma (fig. 1.9f) activity.





**Fig. 1.9: a) An original eeg signal. b) delta band activity, c) theta band activity, d) alpha band activity, e) beta band activity, f) gamma band activity.**



**SYNOPTIC TABLE FOR THE CHARACTERISTICS OF EACH FREQUENCY BAND**

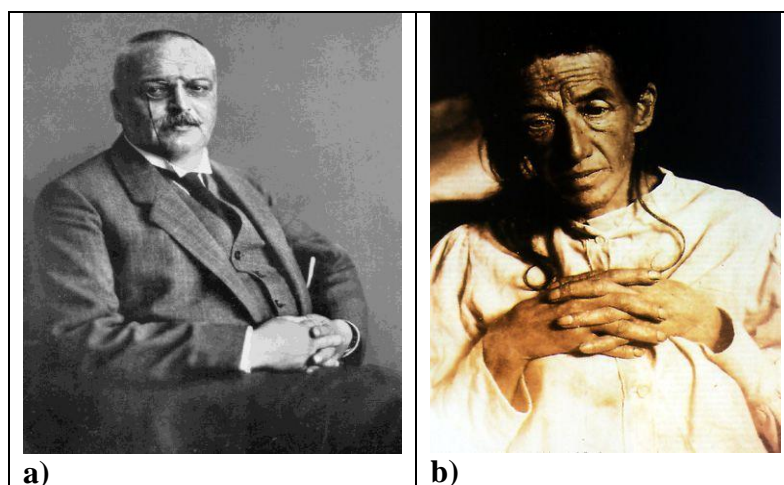
<b>Band</b>	<b>Freq. (Hz)</b>	<b>Location</b>	<b>Normal behaviour</b>	<b>Pathogenic behaviour</b>
Delta	0-3	Frontally in adults, posteriorly in children.	In adults in slow wave sleep, normal in neonates.	Neurological damage, such as cerebral infarct, contusion, local infection, tumor or subdural hematoma
Theta	3-7	In drowsy normal adults, in frontal and temporal regions.	Normal in young children, drowsiness or arousal in older children and adults, idling. Associates with working memory functions, memory load, task demand and episodic (short-term) memory processes, successful encoding of new information, attention.	Focal subcortical lesions, metabolic encephalopathy, deep midline disorders, some instances of hydrocephalus, problems in audio or visual short-term memory encoding, background Theta increases in epilepsy.
Alpha	7-12	Mostly found in occipital and parietal electrodes.	In relaxation, reflecting and closing the eyes. Attention and cognitive processes requiring access and retrieval of information from long term memory. Alpha ERS→ encoding of auditory material. Alpha ERD→ Auditory recognition or comparison (retrieval), cognitive processes (mental arithmetic, reading, tasks).	Coma.
Beta	12-30	Usually on both sides in symmetrical distribution and is most evident frontally.	Observed in alertness, working, active, busy or anxious thinking and in active concentration.	Expectancy tension.
Gamma	30-100	Occipital sites in visual tasks, central sites in motor tasks, posterior sites in reading, occipital, temporal, parietal sites in word reading, frontal, central sites in subtraction /mental tasks.	In certain cognitive or motor functions.	

## 2.6 Alzheimer's Disease

Alzheimer's disease (AD) is a progressive disease, which means that gradually over time, more parts of the brain are damaged. As this happens, the symptoms become more severe. There is often confusion between the relationship of AD and dementia. The word dementia is a term which describes a serious deterioration in mental functions, such as memory, language, orientation and judgment. AD is one cause of dementia, but several other diseases can cause it too. Alzheimer's disease is the most common cause of dementia, accounting for around two thirds of cases in the elderly [42].

In AD, as in other types of dementia, increasing numbers of nerve cells deteriorate and die. A healthy adult brain has 100 billion nerve cells, or neurons, with long branching extensions connected at 100 trillion points. At these connections, called synapses, information flows in tiny chemical pulses released by one neuron and taken up by the receiving cell. Different strengths and patterns of signals move constantly through the brain's circuits, creating the cellular basis of memories, thoughts and skills. In AD, information transfer at the synapses begins to fail, the number of synapses declines and eventually cells die. Brains with advanced Alzheimer's show dramatic shrinkage from cell loss and widespread debris from dead and dying neurons [43], [44].

It was not until 1901 that German psychiatrist Alois Alzheimer identified the first case of what became known as Alzheimer's disease in a fifty-year-old woman he called Auguste D. Alzheimer followed her until she died in 1906, when he first reported the case publicly. During the course of the disease, "plaques" and "tangles" develop in the structure of the brain, leading to the death of brain cells [41]. Plaques build up between nerve cells. They contain deposits of a protein fragment called beta-amyloid. Tangles are twisted fibers of another protein called tau. Tangles form inside dying cells. Though most people develop some plaques and tangles as they age, those with Alzheimer's tend to develop far more. The plaques and tangles tend to form in a predictable pattern, beginning in areas important in learning and memory and then spreading to other regions. People with Alzheimer also have a shortage of some important chemicals in their brains. These chemicals are involved with the transmission of messages within the brain [44].



**Fig. 1.10: a) German psychiatrist Alois Alzheimer, b) The first known AD patient Auguste D. [45].**

### 2.6.1 AD Hallmarks

Some of the most characteristic events taking place in an AD patient's brain are: abundance of two abnormal structures amyloid plaques and neurofibrillary tangles that are made of misfolded proteins especially in certain regions of the brain that are important in memory and loss of connections between cells which leads to diminished cell function and cell death.

- ***Amyloid plaques:*** Amyloid plaques are found in the spaces between the brain's nerve cells. They were first described by Dr. Alois Alzheimer in 1906. Plaques consist of largely insoluble deposits of an apparently toxic protein peptide, or fragment, called beta-amyloid. It is known that some people develop some plaques in their brain tissue as they age. However, the AD brain has many more plaques in particular brain regions. It is under investigation whether amyloid plaques themselves cause AD or whether they are a by-product of the AD process. However it is known that genetic mutations can increase production of beta-amyloid and can cause rare, inherited forms of AD [44].
- ***Neurofibrillary tangles:*** The second hallmark of AD, also described by Dr. Alzheimer, is neurofibrillary tangles. Tangles are abnormal collections of twisted protein threads found inside nerve cells. The chief component of tangles is a protein called tau. Healthy neurons are internally supported in part by structures called microtubules, which help transport nutrients and other cellular components, such as neurotransmitter-containing vesicles, from the cell body down the axon. Tau, which usually has a certain number of phosphate molecules attached to it, binds to microtubules and appears to stabilize them. In AD, an abnormally large number of additional phosphate molecules attach to tau. As a result of this "hyperphosphorylation," tau disengages from the microtubules and begins to come together with other tau threads. These tau threads form structures called paired helical filaments, which can become enmeshed with one another, forming tangles within the cell. The microtubules can disintegrate in the process, collapsing the neuron's internal transport network. This collapse damages the ability of neurons to communicate with each other [44].
- ***Neuron damage:*** The third major feature of AD is the gradual loss of connections between neurons. Neurons live to communicate with each other, and this vital function takes place at the synapse. Since the 1980s, new knowledge about plaques and tangles has provided important insights into their possible damage to synapses and on the development of AD. The AD process not only inhibits communication between neurons but can also damage neurons to the point that they cannot function properly and eventually die. As neurons die throughout the brain, affected regions begin to shrink in a process called brain atrophy. By the final stage of AD, damage is widespread, and brain tissue has shrunk significantly [44].

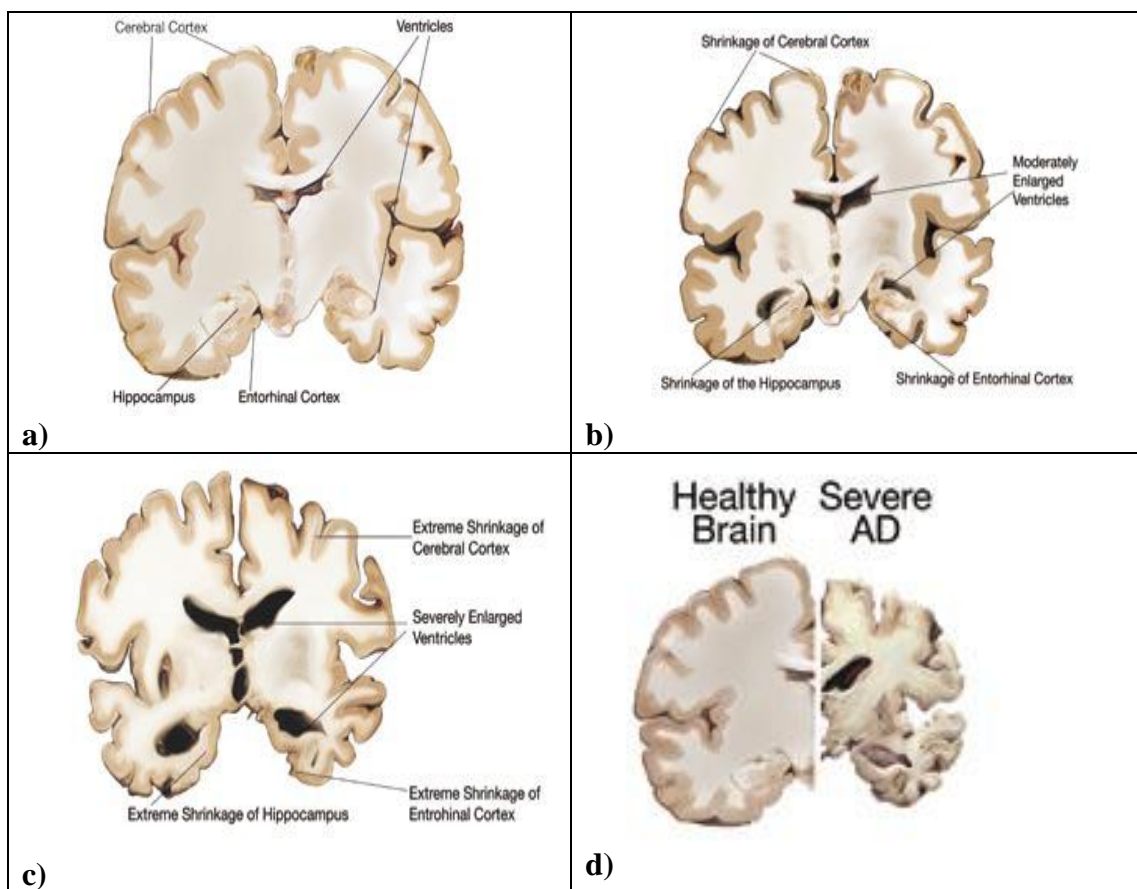
## 2.6.2 AD Stages and Symptoms

The following seven stages were developed by researchers and physicians to describe how a subject having AD changes over time. Many times the seven stages of AD are summed up in three main stages: early, middle, late or mild, moderate and severe.

- **Stage 1(Absence of Impairment):** There are no problems with memory, orientation, judgment, communication, or daily activities [46], [47], [43].
- **Stage 2 (Minimal Impairment):** AD patient might be experiencing some lapses in memory or other cognitive problems not detectably by medical exam [46], [47], [43].
- **Stage 3 (Noticeable Cognitive Decline):** Mild changes in memory, communication patterns, or behaviour can be recognized. Sometimes mild Alzheimer's disease can be diagnosed. Common symptoms in this stage include, problems producing people's names or the right words for objects, noticeable difficulty functioning in employment or social settings, forgetting material that has just been read, misplacing important objects with increasing frequency and decrease in planning or organizational skills [46], [47], [43].
- **Stage 4 (Early-Stage/Mild Alzheimer's):** Cognitive decline is more evident. AD patient becomes forgetful of recent events or personal details. Other problems include impaired mathematical ability (for instance, difficulty counting backwards from 100 by 9s), a diminished ability to carry out complex tasks like throwing a party or managing finances, moodiness, and social withdrawal [46], [47], [43].
- **Stage 5 (Middle-Stage/Moderate Alzheimer's):** At this stage some assistance with daily tasks is required. Problems with memory and thinking are quite noticeable, including symptoms such as an inability to recall one's own contact information or key details about one's history, disorientation to time and/or place and decreased judgment and skills in regard to personal care. Even though symptoms are worsening, people in this stage usually still know their own name and the names of key family members and can eat and use the bathroom without assistance [46], [47], [43].
- **Stage 6 (Middle-Stage/Moderate to Late-Stage/Severe Alzheimer's):** At this stage AD patient has many personality and behaviour changes. In addition, memory continues to decline, and assistance is required for most daily activities. The most common symptoms associated with this stage include: Reduced awareness of one's surroundings and of recent events, problems recognizing one's spouse and other close family members, although faces are still distinguished between familiar and unfamiliar, increased restlessness and agitation in the late afternoon and evening, difficulty using the bathroom independently, bowel and

bladder incontinence, suspicion, repetitive behaviour and wandering [46], [47], [43].

- **Stage 7 (Late-Stage/Severe Alzheimer's):** In the final stage, it is usually no longer possible to respond to the surrounding environment. Basic functions begin to shut down, such as motor coordination and the ability to swallow. The patient starts to have weight loss, seizures and skin infections and reacts by groaning, moaning, or grunting. Although the stages provide a blueprint for the progression of Alzheimer's symptoms, not everyone advances through the stages similarly. Still, the stages are helpful so as to understand AD symptoms [46], [47], [43].



**Fig. 1.12: The figures above show the brain evolution as AD expands. a) Preclinical AD, b) Mild to moderate AD, c) Severe AD, d) Comparison between healthy brain and Severe AD [44].**

In fig. 1.12 we have the brain evolution as AD expands. There are demonstrated the stages from a normal brain (fig. 1.12a) to severe AD's effect on the brain (fig. 1.12c). It has to be mentioned that the last figure (fig. 1.12d) has the real scale correspondence of the healthy brain to the damaged from severe AD brain.

### 2.6.3 AD Causes and Risk Groups

It has not been possible to understand what causes AD, but it is believed that it develops because of a complex series of events that take place in the brain over a long period of time. Many studies are exploring the factors involved in the cause and development of AD. Researchers have discovered that AD:

- Is not a part of normal aging.
- Affects both men and women.
- Is more common in people as they age -- most people with the disease are over 65.
- Is not caused by hardening of the arteries.
- Is not caused by stress.

Those data lead scientists to three different directions, family history (genetic inheritance) and internal environment.

- **Age:** Age is the greatest risk factor for dementia. Dementia affects one in 14 people over the age of 65 and one in six over the age of 80. The likelihood of developing Alzheimer's approximately doubles every five years after age 65. After age 85, the risk reaches nearly 50 percent. However, AD is not restricted to elderly people [48].
- **Genetic inheritance:** In the vast majority of cases, however, effect of inheritance seems to be small. If a parent or other relative has AD, the descendant's chances of developing the disease are only a little higher than someone's whose family has no history of AD [48].
- **Environmental factors:** The environmental factors that may contribute to the onset of Alzheimer's disease have yet to be identified [48].
- **Other factors:** Research has shown that people who smoke, and those who have high blood pressure or high cholesterol levels, increase their risk of developing Alzheimer's [48].

One promising line of research suggests that strategies for overall healthy aging may help keep the brain healthy and may even offer some protection against Alzheimer's. These measures include eating a healthy diet; staying socially active, avoiding tobacco and excess alcohol and exercising both body and mind.

Some reports show that women tend to be infected more frequently than men, though women tend to live longer than men.

## 2.6.4 AD Diagnosis and Treatment

The only definite way to diagnose AD is with an autopsy, which is an examination of the body done after a person dies. However, doctors can determine fairly accurately whether a person who is having memory problems has “possible AD” (the symptoms may be due to another cause) or “probable AD” (no other cause for the symptoms can be found). To diagnose AD, doctors:

- Ask questions about a person’s overall health, past medical problems, ability to carry out daily activities, and changes in behaviour and personality conduct tests of memory, problem solving, attention, counting, and language skills [49].
- Carry out medical tests, such as tests of blood, urine, or spinal fluid [49].
- Perform brain scans, such as a computed tomography (CT) scan or magnetic resonance imaging (MRI) test, single proton emission computed tomography (SPECT) which shows how blood is circulating to the brain, positive electron tomography (PET) which shows how the different areas of the brain respond during certain activities such as reading and talking and other tests such as X-rays and EEG may be used to determine the source of the problem [49].

These tests may be repeated to give doctors information about how the person’s memory is changing over time. Sometimes these tests help doctors find other possible causes of the person’s symptoms. For example, thyroid problems, drug reactions, depression, brain tumors, and blood-vessel disease in the brain can cause AD-like symptoms. Some of these other conditions can be treated successfully.

One of the most exciting areas of ongoing diagnostic research is neuroimaging. Scientists have developed sophisticated imaging systems that may help measure the earliest changes in brain function or structure to identify people in the very first stages of AD well before they develop apparent signs or symptoms.

The National Institute on Aging’s AD Neuroimaging Initiative is a large study that uses MRI and Positron Emission Tomography (PET) scans to learn when and where in the brain changes occur as memory problems develop. These types of neuroimaging scans are still primarily research tools, but one day they may be used more commonly to help physicians diagnose AD at very early stages.

Nevertheless there is currently no cure for AD, but scientific research is bringing us closer to a cure every day. Outstanding progress has already been made in unraveling the mysteries of AD, including what causes it and what happens in the brain as the disease progresses. New understandings about these processes have already provided critical information about how doctors might prevent, delay, stop or even reverse the nerve cell damage that leads to the devastating symptoms of AD. All around the world, scientists and pharmaceutical companies are now racing to develop treatments that address the underlying disease processes, some of which (or a combination of which) might effectively solve the AD puzzle [44].

## 2.7 Event Related Potentials

Originally, Evoked Related Potentials (ERPs) were called Evoked Potentials (EPs) because they were electrical potentials that were evoked by stimuli. The term “EP” is no longer sufficiently general because cerebral processes may be related to voluntary movement or relatively stimulus-independent psychological processes. The term “ERP” was proposed (by Vaughan, 1969) to designate the general class of potentials that display stable time relationships to a definable reference event.

An evoked potential (or “evoked response”) is an electrical potential recorded from a human or animal following presentation of a stimulus, as distinct from spontaneous potentials as detected by electroencephalograms or electromyograms. An event-related potential (ERP) is any stereotyped electrophysiological response to an internal or external stimulus. More simply, it is any measured brain response that is directly the result of a thought or perception.

Evoked potential amplitudes tend to be low, ranging from less than a microvolt to several microvolts, compared to tens of microvolts for EEG. To resolve these low-amplitude potentials against the background of ongoing EEG and other biological signals and ambient noise, signal averaging is usually required. The signal is time-locked to the stimulus and most of the noise occurs randomly, allowing the noise to be averaged out with averaging of repeated responses.

### 2.7.1 Evoked Potential Types

According to the type of stimulation the scientist uses so as to activate the brain circuits of a subject, there are different types of evoked potentials recordings. The most common evoked potentials are sensory and motor evoked potentials. The sensory evoked potentials depending on the kind of the external stimulus are categorized as: visual, auditory and somatosensory.

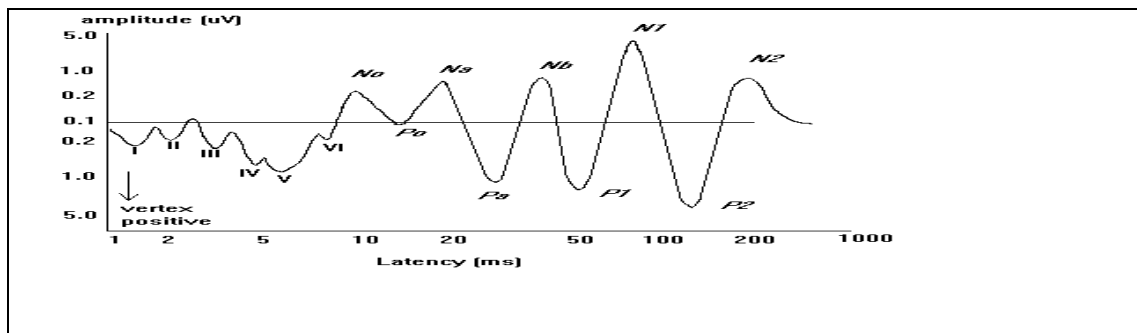
- **Visual Evoked Potentials (VEP):** Visual evoked potentials are very useful in detecting blindness in patients that cannot communicate, such as babies or animals. If repeated stimulation of the visual field causes no changes in EEG potentials, then the subject's brain is probably not receiving any signals from his/her eyes. Other applications include the diagnosis of optic neuritis, which causes the signal to be delayed. Such a delay is also a classic finding in Multiple Sclerosis. Visual evoked potentials are furthermore used in the investigation of basic functions of visual perception. VEPs are also sometimes used to determine if someone is fraudulently alleging blindness. The stimuli are light, pictures projection, colour alternation, etc [64].
- **Auditory Evoked Potentials (AEP):** For AEPs, the “event” is a sound. AEPs (and ERPs) are very small electrical voltage potentials originating from the brain recorded from the scalp in response to an auditory stimulus. The stimuli are different frequency, volume or duration sounds, words, signals, etc [65].
- **Somatosensory Evoked Potentials (SSEP):** Somatosensory Evoked Potentials (SSEPs) are used in neuromonitoring to assess the function



of a patient's spinal cord during surgery. They are recorded by stimulating peripheral nerves, most commonly the posterior tibial nerve, median nerve or ulnar nerve, typically with an electrical stimulus. The stimuli are short duration and amplitude currents which stimulate specific nerves [64].

EPs can also be categorized according to their latency:

- **Early-fast potentials:** They are observed 2-12 msec after the stimulus. The peaks which are observed in this duration are counted as I-VII [66], [67].
- **Middle potentials:** They are observed 12-50 msec after the stimulus. The peaks which are observed in this duration are counted as  $N_0$ ,  $P_0$ ,  $N_a$ ,  $P_a$ ,  $N_b$ ,  $P_b$ . The peaks  $N_0$ ,  $P_0$  are observed before 20 msec and  $N_a$ ,  $P_a$ ,  $N_b$  are observed 30, 40 and 50 msec after the stimulus onset considerably [66], [67].
- **Late potentials:** They are observed 50-800 msec after the stimulus. Those potentials' components are reported as  $N_{100}$ ,  $P_{100}$ ,  $N_{200}$ ,  $P_{200}$ ,  $N_{300}$ ,  $P_{300}$  for 100, 200 and 300 msec after the stimulus onset considerably [66], [67].



**Fig. 1.13: An auditory evoked potential and its components [90].**

In figure 1.13 we illustrate an auditory evoked potential and its components. From 1-10ms we can see the early components I-VI, from 10-50ms the middle potentials  $N_0$ - $N_b$  and from 50-200ms we can see the late potentials  $N_{100}$ - $N_{200}$ .

It should be noticed that as the latency of the potential observation becomes greater, the signal's frequency becomes lower, whereas its amplitude becomes greater.

- **Early-fast potentials:** Amplitude (0.1-0.5  $\mu$ V), frequency (100-1000 Hz) [66], [67].
- **Late potentials:** Amplitude (1-20  $\mu$ V), frequency (0.1-5 Hz) [66], [67].

The early potentials are related to the nerve reduction across the acoustic nerve (or the visual nerve if it is a VEP). On the other hand late potentials show the brain activity's reaction to external information.

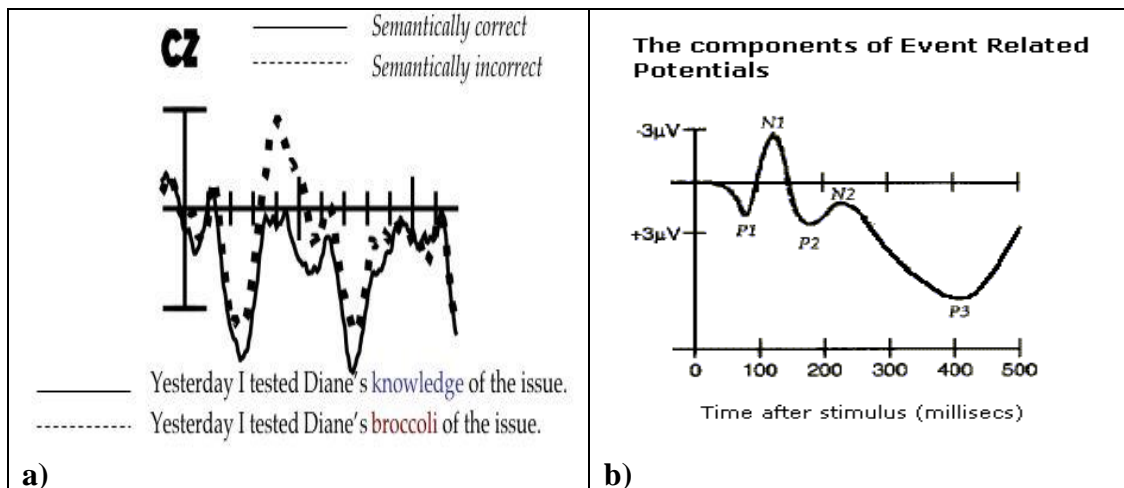
## 2.7.2 Important ERP Components

Most ERP components are referred to by a preceding letter indicating polarity followed by the typical latency in milliseconds. Thus, the P300 component describes a positive voltage peak 300ms after stimulus onset (N300 for a negative voltage peak). The amplitude and the latency of the peaks of the ERP components and also the latency between two different peaks are very important and under research so as to try to identify pathological cases such as Alzheimer, schizophrenia or depression. Some of the most important ERP components are the followings:

- *N<sub>100</sub>*: It is a negative peak usually observed 70-140 msec after the stimulus. While most research focuses on auditory stimuli, the N100 also occurs for visual [68], olfactory [69], heat [70], pain [70], balance [71], respiration blocking [72], and somatosensory [73] stimuli. The N100 is preattentive and involved in perception because its amplitude is strongly dependent upon such things as the rise time of the onset of a sound [74], its loudness [75], interstimulus interval with other sounds [76], and the comparative frequency of a sound as its amplitude increases in proportion to how much a sound differs in frequency from a preceding one [77]. Neuromagnetic research has linked it further to perception by finding that the auditory cortex has a tonotopic organization to N100 [78]. However, it also shows a link to a person's arousal [79] and selective attention [80]. N100 disappears when a person controls the creation of auditory stimuli [81], such as their own voice [82].
- *P<sub>300</sub>*: It is a positive peak observed 240-500 msec after the stimulus. The P300 was discovered originally by Samuel Sutton, Margery Braren, Joseph Zubin, and E. R. John as noted in Science magazine from November 26, 1965 to unpredictable stimuli presented in an oddball paradigm [83]. The P300 wave itself is thought to be comprised of two “wavelets” known as P3a and P3b signals. These components respond individually to different stimuli, and it has been suggested that the P3a wave originates from stimulus-driven frontal attention mechanisms during task processing, whereas P3b originates from temporal-parietal activity associated with attention and appears related to subsequent memory processing. The two wavelets are sometimes referred to as “non-target” (P3a) and “target” (P3b) ERPs [84], [85].
- *N<sub>400</sub>*: It is characterized as a negative deflection, peaking approximately 400ms (300-500ms) after the presentation of the stimulus [86]. The N400 plays a significant role in language processing. It was “discovered” in 1980 by Marta Kutas and Steven Hillyard, in a study that is considered to have first introduced the concept of using ERPs to study language processing, and one of the first studies in what is now the field of neurolinguistics [87]. The N400 response is often elicited by semantically inappropriate words in an otherwise acceptable sentential context [88], and has also been shown to occur in response to words at the end of a sentence when there was a

problem earlier in the sentence. In general, the more difficult it is to “integrate” a word into the preceding context, the bigger an N400 that word will elicit. In models of speech comprehension, N400 is often associated with the semantic integration of words in a sentence context and general “wrap-up” and decision-making processes at the end of a sentence. The N400 remains a common topic of study in neurolinguistics and psycholinguistics [87].

- **P<sub>600</sub>**: It is a positive peak observed peaking around 500ms after the stimulus. It is thought to be elicited by hearing or reading grammatical errors and other syntactic anomalies. Therefore, it is a common topic of study in neurolinguistic experiments investigating sentence processing in the human brain. The P600 can be elicited in both visual (reading) and auditory (listening) experiments and can last several hundred milliseconds. The P600 was first reported by Lee Osterhout and Phillip Holcomb in 1992. It is also sometimes called the Syntactic Positive Shift (SPS), since it has a positive polarity and is usually elicited by syntactic phenomena [89].



**Fig. 1.14: a) A P400 example [92], b) a P300 example [91].**

Except the characterization of the peaks as N<sub>100</sub>, P<sub>300</sub>, etc., there are also specific experiments which evoke the specific peaks. For example auditory oddball experiments can be designed in order to produce P<sub>300</sub> activity. Experiments that evoke such activity take the name of the activity, so an experiment which produces a 300ms peak can be characterized as P<sub>300</sub> experiment. It is not certain that in every experiment we will have all of the peaks we previously considered.

In figure 1.14 we illustrate two such experiments that produce certain peaks. In figure 1.14a there is an experiment which produces a P<sub>400</sub> peak. This experiment shows the response of the brain on electrode Cz in a semantically correct and a semantically wrong phrase. In the semantically wrong phrase we see that we have a peak 400ms after stimulus. This is why the specific experiment can be characterized as a P<sub>400</sub> experiment. In the second figure (fig. 1.14b) there is a corresponding P<sub>300</sub> example. In this case the peak is found around 300ms after stimulus.

# CHAPTER 3: Mathematical Analysis: Background Methods

In this chapter we analyze the mathematical methods and algorithms which we use in our work. Independent component analysis, Partial Directed Coherence, Time-Frequency transform and Event Related Potentials Processing are some of the most important tools in our method. In this chapter, we will give a brief introduction in these methods in order to make some of their main aspects clear.

## 3.1 Blind source separation

Considering the wealth of information embedded in EEG recordings, it is quite important to provide efficient means of decomposing the multi-channel EEG signal into meaningful components by means of a suitable transformation. Let us denote by  $\mathbf{x}$  an  $m$ -dimensional random variable. The problem is then to find a function  $\mathbf{f}$  so that the  $n$ -dimensional transform:

$\mathbf{s} = (s_1, s_2, \dots, s_n)^T$  defined by  $\mathbf{s} = \mathbf{f}(\mathbf{x})$  has some desirable properties.

In most cases, the representation is sought as a linear transform of the observed variables, i.e:  $\mathbf{s} = \mathbf{W}\mathbf{x}$  where  $\mathbf{W}$  is a matrix to be determined.

Using linear transformations makes the problem computationally and conceptually simpler, and facilitates the interpretation of the results. Thus we deal with only linear transformations here.

Several principles and methods have been developed to find a suitable linear transformation. These include principal component analysis, factor analysis, projection pursuit, independent component analysis, and many more. Usually, these methods define a principle that tells which transform is optimal and also the assumptions in each method are important so as to decide which one to use. The optimality may be defined in the sense of optimal dimension reduction, statistical significance of the resulting components  $s_i$ , simplicity of the transformation  $\mathbf{W}$ , or other criteria, including application-oriented ones.

The method of independent component analysis (ICA) provides a tool of EEG decomposition into spatially fixed, timely localized, maximally independent components. As the name implies, the basic goal of ICA is to find a transformation in which the components  $s_i$  are statistically as independent from each other as possible. ICA can be applied, for example in our case, for blind source separation, in which the observed values of  $\mathbf{x}$  correspond to a realization of an  $m$ -dimensional discrete-time signal  $\mathbf{x}(t)$ ,  $t=1, 2, \dots, n$ . Then the components  $s_i(t)$  are called source signals, which are usually original, uncorrupted signals or noise sources. Often such sources are statistically independent from each other, and thus the signals can be recovered from linear mixtures  $x_i$  by finding a transformation in which the transformed signals are as independent as possible, as in ICA. This decomposition is also compliant with the neurophysiological attributes of brain sources and has received significant attention in ERP analysis [50].

Several principles have been developed in statistics, neural computing, and signal processing to find a suitable linear representation of a random variable. Here,

we discuss classical methods for determining the linear transformation. All the methods discussed are based on using centered variables, which means that the mean of the random vector is subtracted. It is henceforth assumed that the variable  $\mathbf{x}$  has already been centered, which means that it has already been transformed by:  $\mathbf{x} = \mathbf{x}_0 - \mathbf{E}\{\mathbf{x}_0\}$ , where  $x_0$  is the original non-centered variable.

### 3.1.1 Second-order Methods

The most popular methods for finding a linear transform are second-order methods. This means methods that find the representation using only the information contained in the covariance matrix of the data vector  $\mathbf{x}$ . The use of second-order techniques is to be understood in the context of the classical assumption of Gaussianity. If the variable  $\mathbf{x}$  has a normal, or Gaussian distribution, its distribution is completely determined by this second-order information. Thus it is useless to include any other information. Another reason for the popularity of the second-order methods is that they are computationally simple, often requiring only classical matrix manipulations.

The two classical second-order methods are principal component analysis and factor analysis. One might roughly characterize the second-order methods by saying that their purpose is to find a faithful representation of the data, in the sense of reconstruction (mean-square) error. This is in contrast to most higher-order methods which try to find a meaningful representation. Of course, meaningfulness is a task-dependent property, but these higher-order methods seem to be able to find meaningful representations in a wide variety of applications.

#### 3.1.1.1 Principal Component Analysis

Principal Component Analysis, or PCA, is widely used in signal processing, statistics, and neural computing. In some application areas, this is also called the (discrete) Karhunen-Loève transform, or the Hotelling transform. PCA is mathematically defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. PCA is theoretically the optimum transform for given data in least square terms [53]. For a data matrix,  $\mathbf{X}^T$ , with zero mean, where each row represents a different repetition of the experiment, and each column gives the results from a particular probe, the PCA transformation is given by:

$$\mathbf{Y}^T = \mathbf{X}^T \mathbf{W} = \mathbf{V} \mathbf{\Sigma}$$

Where  $\mathbf{V} \mathbf{\Sigma} \mathbf{W}^T$  is the singular value decomposition (svd) of  $\mathbf{X}^T$ . Given a set of points in Euclidean space, the first principal component (the eigenvector with the largest eigenvalue) corresponds to a line that passes through the mean and minimizes sum squared error for those points. The second principal component corresponds to the same concept after all correlation with the first principal component has been subtracted out from the points. Each eigenvalue indicates the portion of the variance that is correlated with each eigenvector. Thus, the sum of all the eigenvalues is equal to the sum squared distance of the points with their mean divided by the number of dimensions. PCA essentially rotates the set of points around their mean in order to

align with the first few principal components. This moves as much of the variance as possible (using a linear transformation) into the first few dimensions. The values in the remaining dimensions, therefore, tend to be highly correlated and may be dropped with minimal loss of information. PCA is often used in this manner for dimensionality reduction. PCA has the distinction of being the optimal linear transformation for keeping the subspace that has largest variance. This advantage, however, comes at the price of greater computational requirement if compared, for example, to the discrete cosine transform. Nonlinear dimensionality reduction techniques though, tend to be more computationally demanding than PCA.

The basic idea in PCA is to find the components  $s_1, s_2, \dots, s_n$  so that they explain the maximum amount of variance possible by  $n$ -linearly transformed components. PCA can be defined in an intuitive way using a recursive formulation. Define the direction of the first principal component, say  $w_1$ , by:

$$\mathbf{w}_1 = \arg \max_{\|\mathbf{w}\|=1} E\{(\mathbf{w}^T \mathbf{x})^2\}$$

Where  $\mathbf{w}$  is the matrix of the basis vectors and has one vector per column, and  $\mathbf{w}_1$  is of the same dimension  $m$  as the random data vector  $\mathbf{x}$ . (Let the vectors be column vectors). Thus the first principal component is the projection on the direction in which the variance of the projection is maximized. Having determined the first  $k-1$  principal components, the  $k$ -th principal component is determined as the principal component of the residual:

$$\mathbf{w}_k = \arg \max_{\|\mathbf{w}\|=1} E\left\{\left[\mathbf{w}^T \left(\mathbf{x} - \sum_{i=1}^{k-1} w_i w_i^T \mathbf{x}\right)\right]^2\right\}$$

The principal components are then given by  $s_i = w_i^T \mathbf{x}$ . In practice, the computation of the  $w_i$  can be simply accomplished using the (sample) covariance matrix  $\mathbf{E}\{\mathbf{x}\mathbf{x}^T\} = \mathbf{C}$ . The  $w_i$  are the eigenvectors of  $\mathbf{C}$  that correspond to the  $n$  largest eigenvalues of  $\mathbf{C}$ .

The basic goal in PCA is to reduce the dimension of the data. Thus one usually chooses  $n \ll m$ . Indeed, it can be proven that the representation given by PCA is an optimal linear dimension reduction technique in the mean-square sense. Such a reduction in dimension has important benefits. First, the computational overhead of the subsequent processing stages is reduced. Second, noise may be reduced, as the data not contained in the  $n$  first components may be mostly due to noise. Third, a projection into a subspace of a very low dimension, for example two, is useful for visualizing the data. Note that often it is not necessary to use the  $n$  principal components themselves, since any other orthonormal basis of the subspace spanned by the principal components (called the PCA subspace) has the same data compression or denoising capabilities [53].

### 3.1.1.2 Factor Analysis

A method that is closely related to PCA is factor analysis. In factor analysis, the following generative model for the data is postulated:

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n}$$

where  $\mathbf{x}$  is the vector of the observed variables,  $\mathbf{s}$  is the vector of the latent variables (factors) that cannot be observed,  $\mathbf{A}$  is a constant ( $m \times n$ ) matrix, and the vector  $\mathbf{n}$  is noise, of the same dimension,  $m$ , as  $\mathbf{x}$ . All the variables in  $\mathbf{s}$  and  $\mathbf{n}$  are assumed to be Gaussian. In addition, it is usually assumed that  $\mathbf{s}$  has a lower dimension than  $\mathbf{x}$ . Thus, factor analysis is basically a method of reducing the dimension of the data, in a way similar to PCA.

Suppose we have a set of  $p$  observable random variables,  $x_1, \dots, x_p$  with means  $\mu_1, \dots, \mu_p$ . Suppose for some unknown constants  $l_{ij}$  and  $k$  random variables  $\mathbf{f}$ , where  $i \in 1, \dots, p$  and  $j \in 1, \dots, k$ , where  $k < p$ , we have:

$$x_i - \mu_i = l_{i1}f_1 + \dots + l_{ik}f_k + \varepsilon_i$$

Here  $\varepsilon_i$  is the independently distributed error term with zero mean and finite variance which may not be the same for all of them. In matrix terms, we have

$$\mathbf{x}_{p \times 1} - \boldsymbol{\mu}_{p \times 1} = \mathbf{l}_{p \times k} \mathbf{f}_{k \times 1} + \boldsymbol{\varepsilon}_{p \times 1}, \text{ or omitting the matrix dimensions for clarity, } \mathbf{x} - \boldsymbol{\mu} = \mathbf{L}\mathbf{f} + \boldsymbol{\varepsilon}.$$

Also we will impose the following assumptions on  $\mathbf{F}$ .

1.  $\mathbf{f}$  and  $\boldsymbol{\varepsilon}$  are independent.
2.  $E(\mathbf{f}) = 0$
3.  $Cov(\mathbf{F}) = \mathbf{I}_{k \times k}$

Any solution for the above set of equations following the constraints for  $\mathbf{f}$  is defined as the *factors*, and  $\mathbf{L}$  as the *loading matrix* [135].

There are two main methods for estimating the factor analytic model. The first method is the method of principal factors which is basically a modification of PCA. The idea is here to apply PCA on the data  $\mathbf{x}$  in such a way that the effect of noise is taken into account. In the simplest form, one assumes that the covariance matrix of the noise  $\boldsymbol{\Sigma} = E\{\mathbf{n}\mathbf{n}^T\}$  is known. Then one finds the factors by performing PCA using the modified covariance matrix  $\mathbf{C} - \boldsymbol{\Sigma}$ , where  $\mathbf{C}$  is the covariance matrix of  $\mathbf{x}$ . Thus the vector  $\mathbf{s}$  is simply the vector of the principal components of  $\mathbf{x}$  with noise removed. A second popular method, based on maximum likelihood estimation, can also be reduced to finding the principal components of a modified covariance matrix [51], [52].

Nevertheless, there is an important difference between factor analysis and PCA, though this difference has little to do with the formal definitions of the methods. Equation  $\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n}$  does not define the factors uniquely, but only up to a rotation [51]. This indeterminacy should be compared with the possibility of choosing an arbitrary basis for the PCA subspace, i.e., the subspace spanned by the first  $n$  principal components. Therefore, in factor analysis, it is conventional to search for a 'rotation' of the factors that gives a basis with some interesting properties. The classical criterion is parsimony of representation, which roughly means that the matrix  $\mathbf{A}$  has few significantly non-zero entries. This principle has given rise to such techniques as the varimax, quartimax, and oblimin rotations. Such a rotation has the benefit of

facilitating the interpretation of the results, as the relations between the factors and the observed variables become simpler [51], [52].

### 3.1.2 Higher-order Methods

Higher-order methods use information on the distribution of  $\mathbf{x}$  that is not contained in the covariance matrix. In order for this to be meaningful, the distribution of  $\mathbf{x}$  must not be assumed to be Gaussian, because all the information of (zero-mean) Gaussian variables is contained in the covariance matrix. For more general families of density functions, however, the representation problem has more degrees of freedom. Thus much more sophisticated techniques may be constructed for non-Gaussian random variables. Indeed, the transform defined by second-order methods like PCA is not useful for many purposes where optimal reduction of dimension in the mean-square sense is not needed. This is because PCA neglects such aspects of non-Gaussian data as clustering and independence of the components (which, for non-Gaussian data, is not the same as uncorrelatedness). Three conventional methods based on higher-order statistics are: projection pursuit, redundancy reduction, and Independent Component Analysis (ICA) [50]. Due to its significance ICA is extensively presented in section 3.2.

#### 3.1.2.1 Projection Pursuit

Projection pursuit [136] is a technique developed in statistics for finding “interesting” projections of multidimensional data. Such projections can then be used for optimal visualization of the clustering structure of the data, and for such purposes as density estimation and regression. In one-dimensional projection pursuit, we try to find directions  $\mathbf{w}$  such that the projection of the data in that direction,  $\mathbf{w}^T \mathbf{x}$ , has an “interesting” distribution, for example, it displays some structure. In projection pursuit, one thus wants to reduce the dimension in such a way that some of the “interesting” features of the data are preserved. This is in contrast to PCA where the objective is to reduce the dimension so that the representation is as faithful as possible in the mean-square sense.

The central theoretical problem in projection pursuit is the definition of the projection pursuit index that defines the “interestingness” of a direction. Usually, the index is some measure of non-Gaussianity. A most natural choice is using differential entropy [137], [138]. The differential entropy  $H$  of a random vector  $\mathbf{y}$  whose density is  $f(\cdot)$ , is defined as:

$$H(\mathbf{y}) = -\int f(\mathbf{y}) \log(f(\mathbf{y})) d\mathbf{y}$$

Now, consider zero-mean variables  $\mathbf{y}$  of different densities  $f$ , and constrain the covariance of  $\mathbf{y}$  to be fixed. Then the differential entropy  $H(\mathbf{y})$  is maximized when  $f$  is a Gaussian density. For any other distribution, entropy is strictly smaller. Thus one might try to find projection pursuit directions by minimizing  $H(\mathbf{w}^T \mathbf{x})$  with respect to  $\mathbf{w}$ , constraining the variance of  $\mathbf{w}^T \mathbf{x}$  to be constant.

#### 3.1.2.2 Redundancy Reduction

According to Barlow [140] and several other authors [139], an important characteristic of sensory processing in the brain is redundancy reduction. One aspect



of redundancy reduction is that the input data is represented using components (features) that are as independent from each other as possible (similar to ICA). Such a representation seems to be very useful for later processing stages. Theoretically, the values of the components are given by the activities of the neurons, and  $\mathbf{x}$  is represented as a sum of the weight vectors of the neurons, weighted by their activations. This leads to a linear encoding like the other methods we previously discussed.

One method for performing redundancy reduction is sparse coding [140]. Here the idea is to represent the data  $\mathbf{x}$  using a set of neurons so that only a small number of neurons is activated at the same time. Equivalently, this means that a given neuron is activated only rarely. If the data has certain statistical properties (it is 'sparse'), this kind of coding leads to approximate redundancy reduction. A second method for redundancy reduction is predictability minimization [139]. This is based on the observation that if two random variables are independent, they provide no information that could be used to predict one variable using the other one.

### 3.2 Independent Component Analysis

In the literature, at least three different basic definitions for linear ICA can be found:

- 1) **General definition:** ICA of the random vector  $\mathbf{x}$  consists of finding a linear transform  $\mathbf{s} = \mathbf{W}\mathbf{x}$  so that the components  $s_i$  are as independent as possible, in the sense of maximizing some function  $F(s_1, \dots, s_m)$  that measures independence.
- 2) **Noisy ICA model:** ICA of a random vector  $\mathbf{x}$  consists of estimating the following generative model for the data:  $\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n}$  where the latent variables (components)  $s_i$  in the vector  $\mathbf{s} = (s_1, \dots, s_n)^T$  are assumed independent. The matrix  $\mathbf{A}$  is a constant  $m \times n$  "mixing" matrix, and  $\mathbf{n}$  is a  $m$ -dimensional random noise vector. This definition reduces the ICA problem to ordinary estimation of a latent variable model. However, this estimation problem is not very simple, and therefore the great majority of ICA research has concentrated on the following simplified definition.
- 3) **Noisy-free ICA model:** Nevertheless the great majority of ICA research has concentrated on the following simplified definition [54]: Let  $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)^T$  be the observed  $m$ -dimensional random vector. ICA of  $\mathbf{x}$  consists of estimating the following generative model for the data:  $\mathbf{x} = \mathbf{A}\mathbf{s}$  where the latent variables (components)  $s_i$  in the vector  $\mathbf{s} = (s_1, s_2, \dots, s_n)^T$  are assumed independent and the matrix  $\mathbf{A}$  is a constant ( $m \times n$ ) "mixing" matrix. In this case we do not take into consideration the noise vector, so as to make the estimation problem simpler. This is also the model introduced by Jutten and Héroult in their seminal paper "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture". In this thesis the noise-free ICA model will be considered.

The mutually independent components are produced by:

$$\mathbf{s} = \mathbf{W}\mathbf{x}, \text{ or } \mathbf{x} = \mathbf{W}^{-1}\mathbf{s} = \begin{bmatrix} w'_{11} & \cdots & w'_{1j} & \cdots & w'_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w'_{i1} & \cdots & w'_{ij} & \cdots & w'_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w'_{n1} & \cdots & w'_{nj} & \cdots & w'_{nn} \end{bmatrix} \begin{bmatrix} s_1(t) \\ \vdots \\ s_j(t) \\ \vdots \\ s_n(t) \end{bmatrix}$$

where  $\mathbf{W}$  (“unmixing” matrix) has to be estimated.  $\mathbf{W}$  can be computed as  $\mathbf{W} = \mathbf{A}^{-1}$  when  $\mathbf{A}$  is invertible or else  $\mathbf{W}$  can be estimated as the pseudo-inverse of  $\mathbf{A}$ .

If we are talking about EEG the elements  $w'_{ij}$  of the columns of the inverse matrix of  $\mathbf{W}$  reflect the projection intensity of the  $j^{\text{th}}$  IC  $s_j$  back to each electrode.

It is essential to define the restrictions for the ICA model.

- 1) All the independent components  $s_i$ , with the possible exception of one component, must be non-Gaussian.
- 2) The “sources” have to be truly independent.
- 3) The number of observed linear mixtures  $m$  must be at least as large as the number of independent components  $n$ , i.e.,  $m \geq n$ .
- 4) The matrix  $\mathbf{A}$  must be of full column rank.

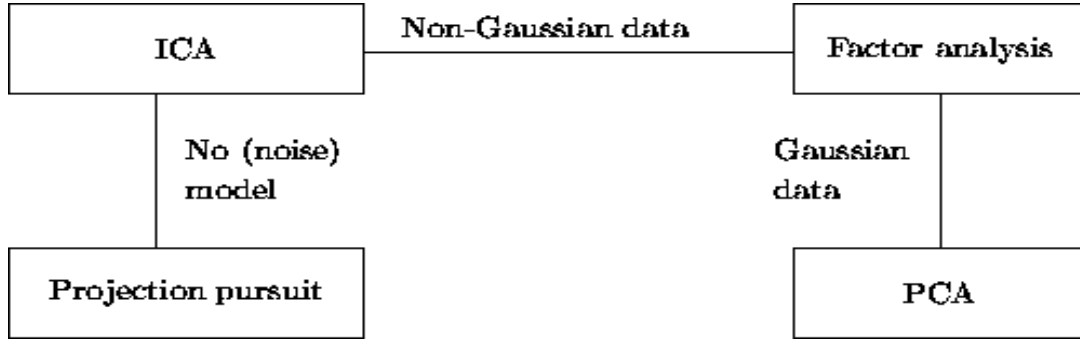
### 3.2.1 ICA Relation to Classical Methods

Previously, we presented some second-order and high-order methods. At this point we will correlate those methods with ICA.

By definition, ICA can be considered a method for achieving redundancy reduction. Indeed, there is experimental evidence that for certain kinds of sensory data, the conventional ICA algorithms do find directions that are compatible with existing neurophysiological data, assumed to reflect *redundancy reduction* [141]. In the noise-free case, the estimation of the ICA model means simply finding certain “interesting” projections, which give estimates of the independent components. Thus ICA can be considered, at least using Definitions 1 and 3, a special case of *projection pursuit*. Comparing Eq.  $\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n}$  in Definition 2 of ICA with the definition of factor analysis, the connection between factor analysis and ICA becomes clear. Indeed, ICA may be considered a *non-Gaussian factor analysis*. The main difference is that usually in ICA, reduction of dimension is considered only as a secondary objective, but this does not need to be the case. Indeed, a simple combination of factor analysis and ICA can be obtained using factor rotations.

Using the first definition we gave for ICA, the relation to *principal component analysis* is also evident. Both methods formulate a general objective function that define the “interestingness” of a linear representation, and then maximize that function. A second relation between PCA and ICA is that both are related to factor analysis, though under the contradictory assumptions of Gaussianity and non-Gaussianity, respectively. The relation between PCA and ICA may be, however, less important than the relation between ICA and the other methods discussed previously. This is because PCA and ICA define their objective functions in quite different ways. PCA uses only second-order statistics, while ICA is impossible using only second-

order statistics. PCA emphasizes dimension reduction, while ICA may reduce the dimension, increase it or leave it unchanged.



**Fig. 3.1: The relations between ICA and some other methods. [50].**

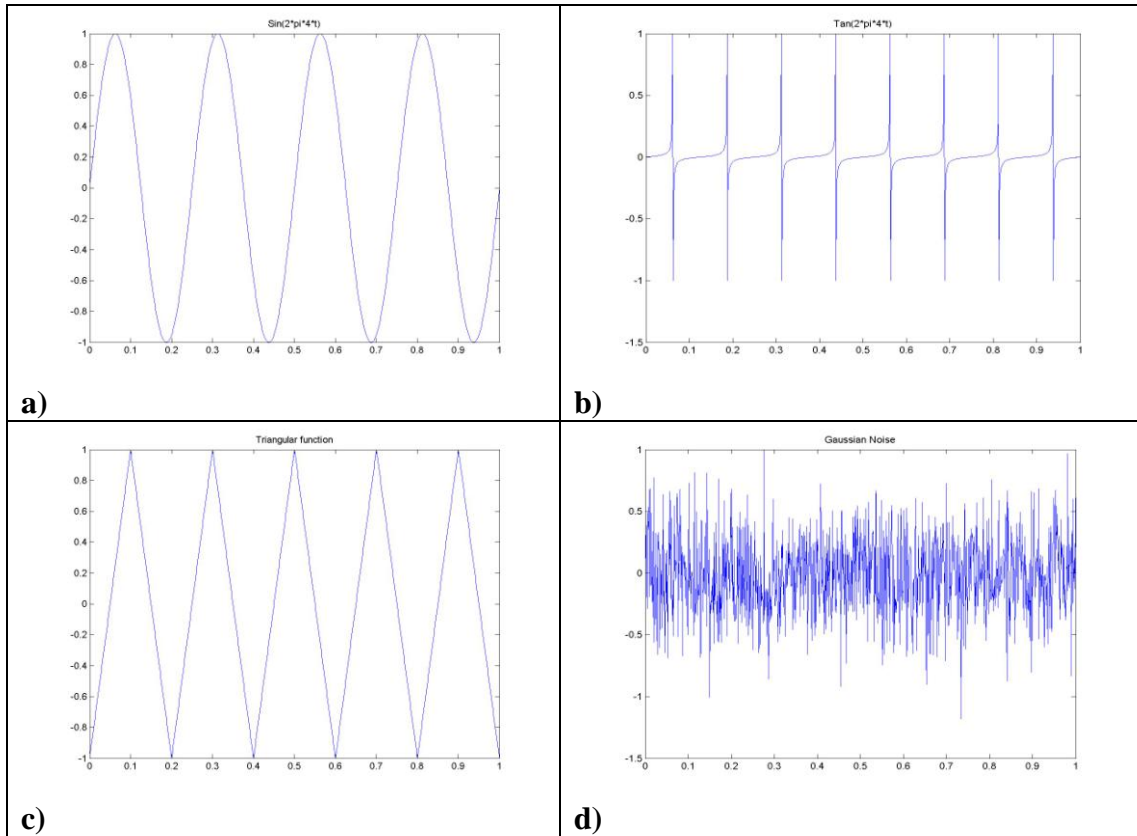
In figure 3.1 we illustrate the relations between the different methods. The lines show close connections, and the texts next to the lines show the assumptions needed for the connections.

### 3.2.2 ICA Applications in BSS

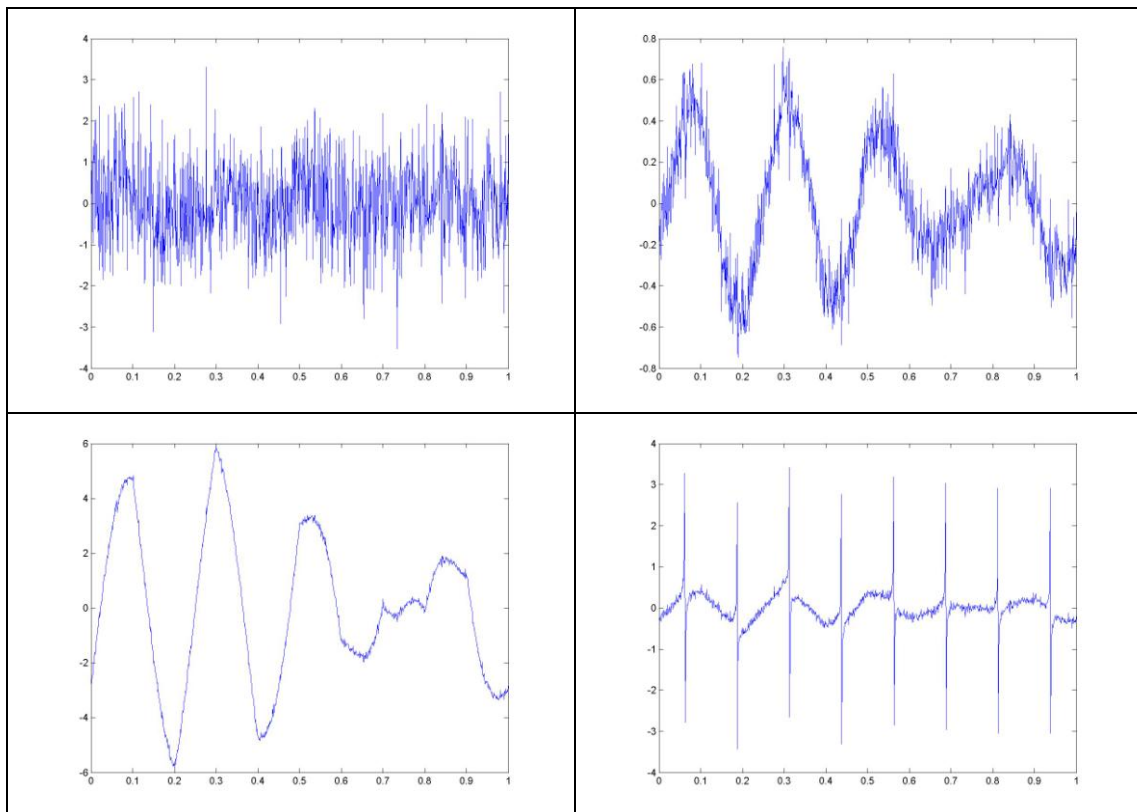
ICA model is used in many different applications, the most classical of which is blind source separation. Other applications in which the ICA model is used are; feature extraction and blind deconvolution. Here we present the aspects of blind source separation:

In blind source separation, the observed values of  $x$  correspond to a realization of an  $m$ -dimensional discrete-time signal  $x(t)$ ,  $t=1,2,\dots,n$ . Then the independent components  $s_i(t)$  are called source signals, which are either uncorrupted signals or noise sources. The cocktail party problem [80] is considered a classical example of blind source separation. Assuming that several people are speaking simultaneously in the same room, as in a cocktail party, the problem is to separate the voices of the different speakers, using recordings of several microphones in the room. In principle, this corresponds to the ICA data model, where  $x$  is the recording of the  $i$ -th microphone, and the  $s_i(t)$  are the waveforms of the voices. A more practical application is noise reduction. If one of the sources is the original, uncorrupted source and the others are noise sources, estimation of the uncorrupted source is in fact a denoising operation [55].

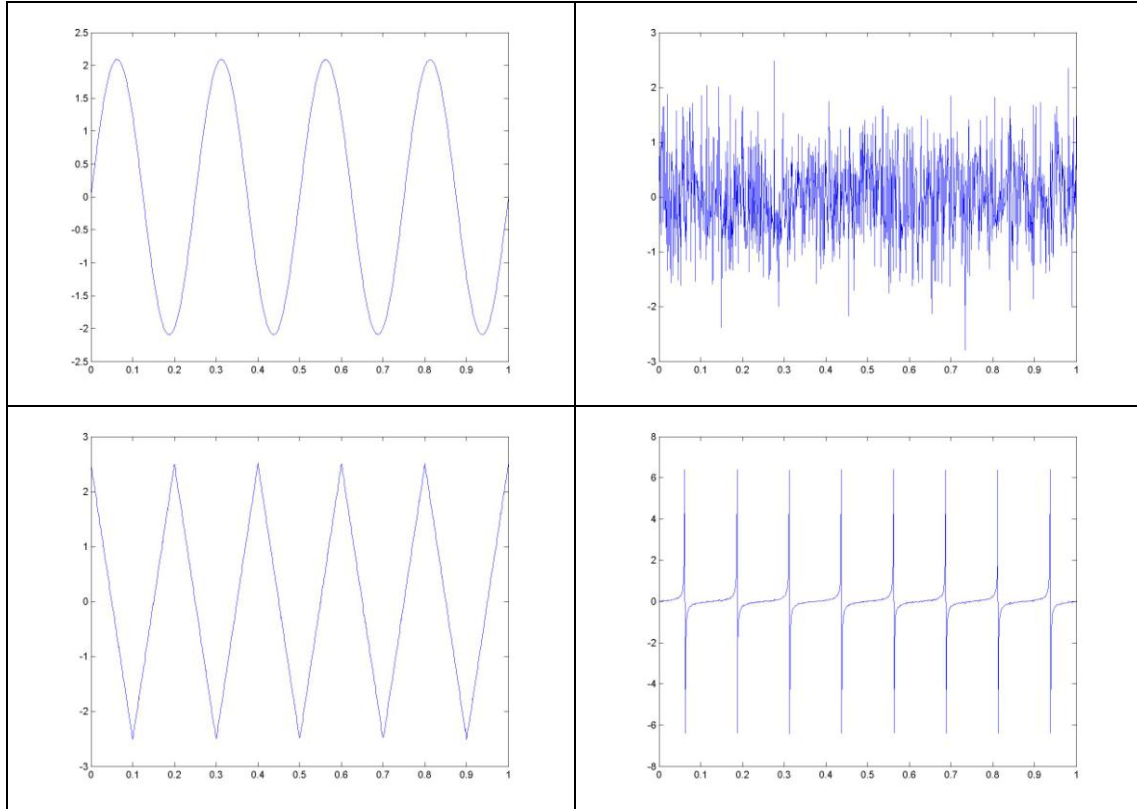
A simulated example on blind source separation of four linearly mixed source signals is presented in the following figures:



**Fig. 3.2: 4 source signals .a) a sinusoidal signal, b) a tangent signal, c) a triangle signal, d) Gaussian noise.**



**Fig. 3.3: Linear mixture of the former source signals.**



**Fig. 3.4: Retrieval of the source signals using ICA.**

In the figures above we have step by step the procedure followed so as to “unmix” linearly “mixed” signals. In figure 3.2 we have the four initially “unmixed” original signals. These signals are a sinusoidal signal, a tangent signal, a triangle signal and Gaussian noise. In the next figure 3.3 we have four new signals coming from the linear mixture of the former signals. Finally in the last figure (fig. 3.4) we have the unmixed signals after applying ICA. The horizontal axis in the figures is for time and the vertical for amplitude.

In this thesis, we exploit the application of ICA on the EEG data, so as to find the unmixed, maximally independent sources which correspond to the observed, mixed EEG data. In this way a decomposition of the content of the EEG data is accomplished. Also by applying ICA on EEG data, artifacts (or noise sources) not corresponding to brain activity are succeeded to be separated.

### 3.2.3 ICA Cost Functions and Algorithms

A cost function for the ICA should be defined as any function whose optimization (by minimizing or maximizing it) enables the estimation of the independent components. The choice of the ICA method should be a parameter of both the cost function which has certain statistical properties (e.g., consistency, asymptotic variance, robustness) as well as the algorithm used to optimize it which also has different properties (e.g., convergence speed, memory requirements, numerical stability) [50].

#### ICA Method $\Leftrightarrow$ Cost Function + Optimization Algorithm

In this essay ICA methods are not the main concern as ICA is used as an intermediate tool for the further analysis. Thus, a thumbnail description of ICA cost functions and algorithms follows.

The ICA cost function could be either multi-unit or one-unit. In the multi-unit cost functions the problem of estimating all the independent components, or the whole data model, at the same time is to be treated. On the other hand in the one-unit cost functions optimization which enables estimation of a single independent component is desired. Thus, instead of estimating the whole ICA model, we try to find here simply one vector, say  $\mathbf{w}$ , so that the linear combination  $\mathbf{w}^T \mathbf{x}$  equals one of the independent components  $s_i$ . This procedure can be iterated to find several independent components.

The most commonly used multi-unit cost functions could be synopsised in: likelihood and network entropy, mutual information and Kullback-Leibler divergence, non-linear cross correlation, non-linear PCA criteria, weighted covariance matrix and high-order cumulant tensors. As far as one-unit cost functions are concerned, they could be synopsised in: negentropy, higher-order cumulants and other general cost functions.

#### 3.2.3.1 Likelihood and network entropy

Denoting by  $\mathbf{w} = (w_1, w_2, \dots, w_n)^T$  the matrix  $\mathbf{A}^{-1}$ , the log-likelihood takes the form [56]:

$$L = \sum_{t=1}^T \sum_{i=1}^m \log f_i(\mathbf{w}_i^T \mathbf{x}(t)) + T \ln |\det \mathbf{W}|$$

where the  $f_i$  are the density functions of the  $s_i$  (here assumed to be known), and the  $\mathbf{x}(t)$ ,  $t=1,2,\dots,T$  are the realizations of  $\mathbf{x}$ .

Another related contrast function was derived from a neural network viewpoint and was based on maximizing the output entropy (or information flow) of a neural network with non-linear outputs [58]. Assume that  $\mathbf{x}$  is the input to the neural network whose outputs are of the form  $g_i(\mathbf{w}_i^T \mathbf{x})$ , where the  $g_i$  are some non-linear scalar functions, and the  $\mathbf{w}_i$  are the weight vectors of the neurons. The goal is to maximize the entropy of the outputs:

$$L_2 = H(g_1(\mathbf{w}_1^T \mathbf{x}), \dots, g_m(\mathbf{w}_m^T \mathbf{x}))$$

If the  $g_i$  are well chosen, this framework also enables the estimation of the ICA model. Indeed, several authors proved the surprising result that the principle of network entropy maximization, or “infomax”, is equivalent to maximum likelihood estimation [57].

### 3.2.4 ICA Optimization Algorithms

After choosing a cost function for ICA, one needs a practical method for its implementation. Some of the most commonly used algorithms for the ICA implementation are: infomax estimation, Jutten-Herault algorithm, non-linear decorrelation algorithms, non-linear PCA algorithms, neural one-unit algorithms, the FastICA algorithm and so forth. In this essay the ICA tool which is used implements the infomax algorithm so some analysis will be given on that particular algorithm.

Usually infomax estimations (or maximum likelihood algorithms) are based on (stochastic) gradient ascent of the cost function. For example the following algorithm:

$$\Delta W \propto [\mathbf{W}^T]^{-1} - 2 \tanh(\mathbf{W}\mathbf{x})\mathbf{x}^T$$

where the tanh function is applied separately on every component of the vector  $\mathbf{W}\mathbf{x}$ . The algorithm in this equation though converges slowly, and it may be improved by whitening the data, and especially by using the natural gradient. The natural (or relative) gradient method simplifies the gradient method considerably, and makes it better conditioned. The principle of the natural gradient is based on geometrical structure of the parameter space, and is related to the principle of relative gradient that uses the Lie structure of the ICA problem. In the case of basic ICA, both of these principles amount to multiplying the right hand of the equation by  $\mathbf{W}^T \mathbf{W}$ . Thus:

$$\Delta W \propto (\mathbf{I} - 2 \tanh(\mathbf{y})\mathbf{y}^T) \mathbf{W}$$

with  $\mathbf{y}=\mathbf{W}\mathbf{x}$  [58].

### 3.3 Time-Frequency Representation

A signal, as a function of time, may be considered as a representation with perfect time resolution. In contrast, the magnitude of the Fourier transform (FT) of the signal may be considered as a representation with perfect spectral resolution but with no time information because the magnitude of the FT conveys frequency content but it fails to convey when, in time, different events occur in the signal. TF analysis provides a bridge between these two representations in that they provide some temporal information and some spectral information simultaneously. Thus, TF is useful for the analysis of signals containing multiple time-varying frequencies.

There are many ways in which we can have a time-frequency representation such as the short time FT and the wavelet transforms. In general, such representations are implemented as linear transforms. They include and generalize the Fourier transform, fractional Fourier transform, and others, thus providing a unified view of these transforms in terms of their action on the time–frequency domain [59]. Furthermore, there are other ways of representing a time-frequency mapping such as the popular quadratic form and the Cohen’s class of functions, with most commonly used the Wigner-Ville transform.

- **Linear forms:** In linear forms we have comparison of the signal with a different function. Such representations are known as linear TF because the representation is linear in the signal. The short-time Fourier transform localizes the signal by modulating it with a window function, before performing the Fourier transform to obtain the frequency content of the signal in the region of the window [59].

$$\text{Mathematically, this is written as: } STFT(\mathbf{t}, \boldsymbol{\omega}) = \int_{-\infty}^{\infty} f(\mathbf{t})w(\mathbf{t} - \tau)e^{-j\boldsymbol{\omega}\tau} d\tau$$

The STFT implements signal projection on the time-frequency domain. The detailed TF energy distribution function is referred to as “Spectrogram”.

- **Wavelet transforms:** Wavelet transforms, in particular the continuous wavelet transform, expand the signal in terms of wavelet functions which are localized in both time and frequency. Thus the wavelet transform of a signal may be represented in terms of both time and frequency [59]. The WT can be defined as:

$$WT(s, \tau) = \int f(t)\psi_{s,\tau}^*(t)dt$$

. The WT implements signal projection on the time-scale domain, where the detailed TS energy distribution function is referred to as “Scalogram”. Both the spectrogram and the scalogram distributions are effectively approximated by the WV transform.

- **Quadratic form:** One form of TF can be formulated by the multiplicative comparison of a signal with itself, expanded in different directions about each point in time. Such formulations are known as quadratic TF representations because the representation is quadratic in the signal. This formulation was first described by Eugene Wigner in 1932 in the context of quantum mechanics and, later, reformulated as a general TF representation by Ville in 1948 to form what is now known



as the Wigner–Ville distribution [59]. The Wigner distribution

$$WV(\mathbf{t}, \boldsymbol{\omega}) \text{ is defined as: } WV(\mathbf{t}, \boldsymbol{\omega}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\mathbf{t} + \frac{\tau}{2}) f^*(\mathbf{t} - \frac{\tau}{2}) e^{-j\boldsymbol{\omega}\tau} d\tau$$

where  $f$  is the original time signal and  $\mathbf{t}$  and  $\boldsymbol{\omega}$  are the time and the frequency realizations. The WV transform has been extensively used for the approximation of the time-frequency and time-scale energy distributions.

### 3.4 Wavelets

In this essay the wavelet transform (WT) is used due to its better frequency resolution compared to Wigner –Ville distribution, as it can be observed in the previous figures and we discussed. The word *wavelet* is due to Morlet and Grossmann in the early 1980s. They used the French word *ondelette*, meaning “small wave”. Soon it was transferred to English by translating “onde” into “wave”, giving “wavelet”. Wavelets are purposefully crafted to have specific properties that make them useful for signal processing. A wavelet transform is the representation of a function by wavelets. The wavelets are scaled and translated copies (known as “daughter wavelets”) of a finite-length or fast-decaying oscillating waveform (known as the “mother wavelet”). Wavelet transforms have advantages over traditional Fourier transforms for representing functions that have discontinuities and sharp peaks, and for accurately deconstructing and reconstructing finite, non-periodic and/or non-stationary signals. Such signals are the EEG data, for which wavelet transform is a very efficient way to show their frequency content during time pass.

Wavelet transforms are broadly divided into three classes: continuous, discrete and multiresolution-based. We focus on continuous wavelet transform.

#### 3.4.1 Continuous Wavelet Transform

The wavelet transform (WT) or wavelet analysis is probably the most recent solution to overcome the shortcomings of the Fourier transform. In wavelet analysis the use of a fully scalable modulated window solves the signal-cutting problem. The window is shifted along the signal and for every position the spectrum is calculated. Then this process is repeated many times with a slightly shorter (or longer) window for every new cycle. In the end the result will be a collection of time-frequency representations of the signal, all with different resolutions. Because of this collection of representations we can speak of a multiresolution analysis. In the case of wavelets we normally do not speak about time-frequency representations but about time-scale representations, scale being in a way the opposite of frequency, because the term frequency is reserved for the Fourier transform [61].

The WT can be defined as:

$$WT(s, \tau) = \int f(t) \psi_{s,\tau}^*(t) dt$$

where  $*$  denotes complex conjugation. This equation shows how a function  $f(t)$  is decomposed into a set of basis functions  $\psi_{s,\tau}(t)$ , called the wavelets. The variables  $s$  and  $\tau$  are the new dimensions, scale and translation, after the wavelet transform.

The wavelets are generated from a single basic wavelet  $\psi(t)$ , the so-called mother wavelet, by scaling and translation:

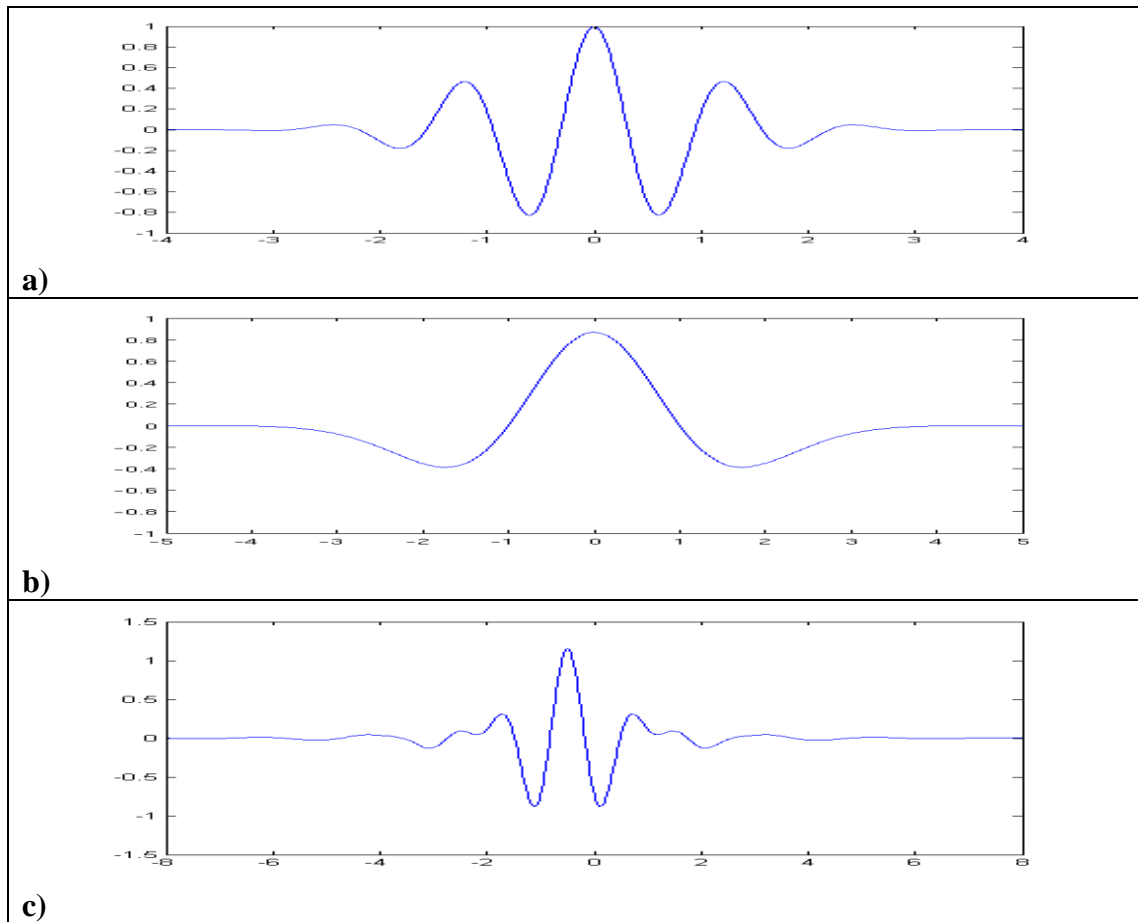
$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right)$$

In this equation  $s$  is the scale factor  $\tau$ , is the translation factor and the factor  $s^{-1/2}$  is for energy normalization across the different scales.

For the Complex Morlet Wavelet we use in our analysis we have the following wavelet function:

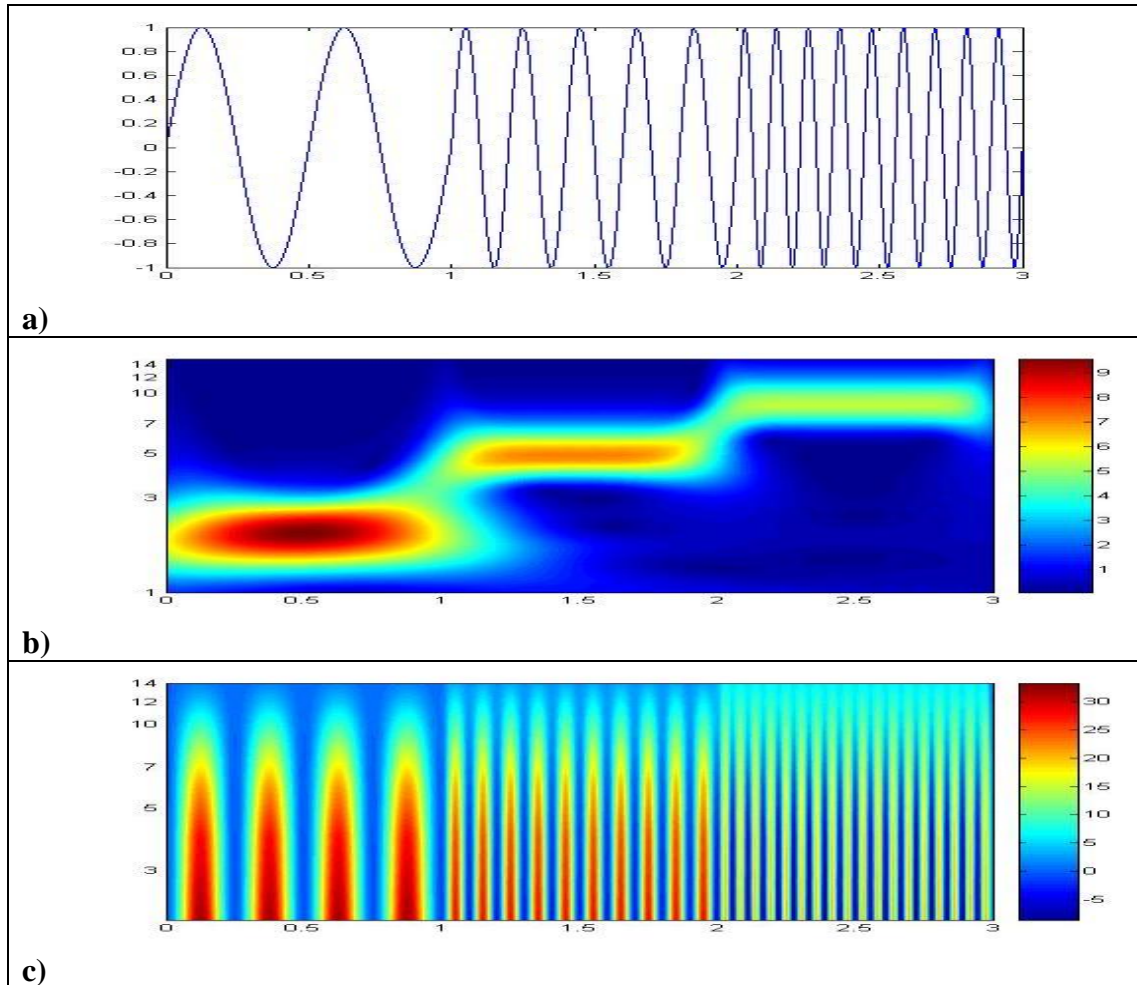
$$CMWT(t) = (\pi F_b)^{-\frac{1}{2}} e^{2i\pi F_c t} e^{-\frac{t^2}{F_b}}$$

depending on two parameters:  $F_b$  is a bandwidth parameter,  $F_c$  is a wavelet center frequency.



**Fig. 3.5: Three examples for “mother wavelets”. a) Morlet wavelet, b) Mexican hat, c)Meyer wavelet.**

In the figures above we illustrate some examples of popular functions which can be used as mother wavelets. Those are the Morlet wavelets, the Mexican hat and the Meyer wavelets.



**Fig. 3.6: Two different TF representations of a) a sinusoidal signal : b) Wavelet, c) Wigner-Ville.**

In figure 3.6 we have a sinusoidal signal which is consisted of  $\sin(2\pi f_i t)$   $i=1, \dots, 3$  where  $f_1=2\text{Hz}$ ,  $f_2=5\text{Hz}$  and  $f_3=9\text{Hz}$ . Those sinusoids are concatenated in time in figure 3.5a and then in figures 3.6b and 3.6c we have their Morlet Complex Wavelet TF representation and Wigner-Ville TF representation respectively. We can see that each transform has other advantages. For the wavelet we see that it has leaks in time though shows more efficiently and limited the frequency content of the signal. We can also see that the higher the frequency is, the more efficient the result is. On the other hand, Wigner-Ville has an excellent time analysis as it “catches” every peak of the sinusoidal signal, though it does not have a good frequency analysis.

### 3.4.2 Continuous Wavelets vs Fourier Transform

The wavelet transform is often compared with the Fourier transform, in which signals are represented as a sum of sinusoids. The main difference is that wavelets are localized in both time and frequency whereas the standard Fourier transform is only localized in frequency. The Short-time Fourier transform (STFT) is also time and frequency localized but there are issues with the frequency time resolution and wavelets often give a better signal representation using Multiresolution analysis.

An advantage of wavelet transforms is that the windows vary. In order to isolate signal discontinuities, one would like to have some very short basis functions.

At the same time, in order to obtain detailed frequency analysis, one would like to have some very long basis functions. A way to achieve this is to have short high-frequency basis functions and long low-frequency ones. This happy medium is exactly what we get with wavelet transforms [63].

The discrete wavelet transform is also less computationally complex, taking  $O(N)$  time as compared to  $O(N \log N)$  for the fast Fourier transform. This computational advantage is not inherent to the transform, but reflects the choice of a logarithmic division of frequency, in contrast to the equally spaced frequency divisions of the FFT [62].

### 3.5 Partial Directed Coherence

In order to find the relations between different independent components we use the method of partial directed coherence. Partial directed coherence (PDC) was recently introduced as a linear frequency-domain quantifier of the multivariate relationship between simultaneously observed timeseries for application in functional connectivity inference in neuroscience. Conceptually, PDC is a generalization to the case of multiple time-series of Saito and Harashima's "directed coherence" (DC) which was also introduced in the context of analysing neural data and which was aimed at pinpointing the direction of information flow between isolated pairs of time series in a frequency domain representation of the notion of Granger Causality to the scenario when more than just pairs of time series are simultaneously analyzed. A time-series  $x_2(n)$  is said to Granger cause  $x_1(n)$  if consideration of  $x_2(n)$ 's past implies significant improvement in the mean-squared prediction error  $x_1(n)$ . Granger causality is not reciprocal. If  $x_1(k)$  Granger causes  $x_2(k)$  then it does not necessarily follow that  $x_2(k)$  Granger causes  $x_1(k)$ . Or equivalently the distinctive fact about Granger causality is its unreciprocal nature [96], [97]:

$$x_1(k) \xrightarrow{\text{Granger}} x_2(k)$$

does not imply

$$x_2(k) \xrightarrow{\text{Granger}} x_1(k)$$

Because PDC is based on the notion of partial coherence its chief property is to provide a description of the mutual interaction between pairs of time series after deducting the effect of other simultaneously observed time series.

If one assumes that a set of simultaneously observed time series:

$$\mathbf{x}(n) = [x_1(n) \dots x_N(n)]^T$$

is adequately represented by a Multivariate Autoregressive Model of order  $p$  (MVAR( $p$ )):

$$\mathbf{x}(n) = \sum_{k=1}^p \mathbf{A}_k \mathbf{x}(n-k) + \mathbf{w}(n)$$

,where  $A_k$  comprise the coefficients  $a_{ij}(k)$  that link the signal (time series)  $x_i(n)$  with  $x_j(n)$  at lag  $k$  (describing the interactions between time series pairs over time) and where:

$$\mathbf{w}(n) = [w_1(n) \dots w_N(n)]^T$$

,is the vector of model innovations (zero mean and with covariance matrix  $\Sigma_w$ ). This leads to PDC between signals  $i$  and  $j$  expressed as:

$$\pi_{ij}(f) = \frac{\bar{A}_{ij}(f)}{\sqrt{\sum_{k=1}^N \bar{A}_{kj}(f) \bar{A}_{k^*j}(f)}}$$

,where  $f$  is the normalized frequency in the interval  $[-0.5, 0.5]$  where:

$$\bar{A}_{ij}(f) = \delta_{ij} - \sum_{k=1}^p a_{ij}(k) e^{-2j\pi f k}$$

,for  $\delta_{ij}=1$  whenever  $i=j$  and  $\delta_{ij}=0$  otherwise.

To circumvent the numerical problem associated with time series scaling, a new partial directed coherence estimator is defined as:

$$\pi_{ij}^{(w)}(f) = \frac{\frac{1}{\sigma_i} \bar{A}_{ij}(f)}{\sqrt{\sum_{k=1}^N \frac{1}{\sigma_k^2} \bar{A}_{kj}(f) \bar{A}_{k^*j}(f)}}$$

whence it follows that

$$|\pi_{ij}^{(w)}(f)|^2 \leq 1$$

and

$$\sum_{i=1}^N |\pi_{ij}^{(w)}(f)|^2 = 1$$

### 3.6 Common Processing of ERP

As the EEG reflects thousands of simultaneously ongoing brain processes, the brain response to a single stimulus or event of interest is not usually visible in the EEG recording of a single trial. To see the brain response to the stimulus, the experimenter must conduct many trials and average the results together, causing random brain activity to be averaged out and the relevant ERP to remain. This averaging procedure enhances the signal to noise ratio (SNR).

Suppose, we have the  $i$ -th trial of an EEG signal  $r_i(t)$ . It can be considered that this signal is consisted of the desirable signal and noise.

$$r_i(t) = s_i(t) + n_i(t) \quad i = 1, \dots, N$$

where  $N$  is the number of iterations. Taking the average of all the recordings we have:

$$E(r_i(t)) = \frac{\sum_{i=1}^N r_i(t)}{N} = \frac{1}{N} \left( \sum_{i=1}^N s_i(t) + \sum_{i=1}^N n_i(t) \right)$$

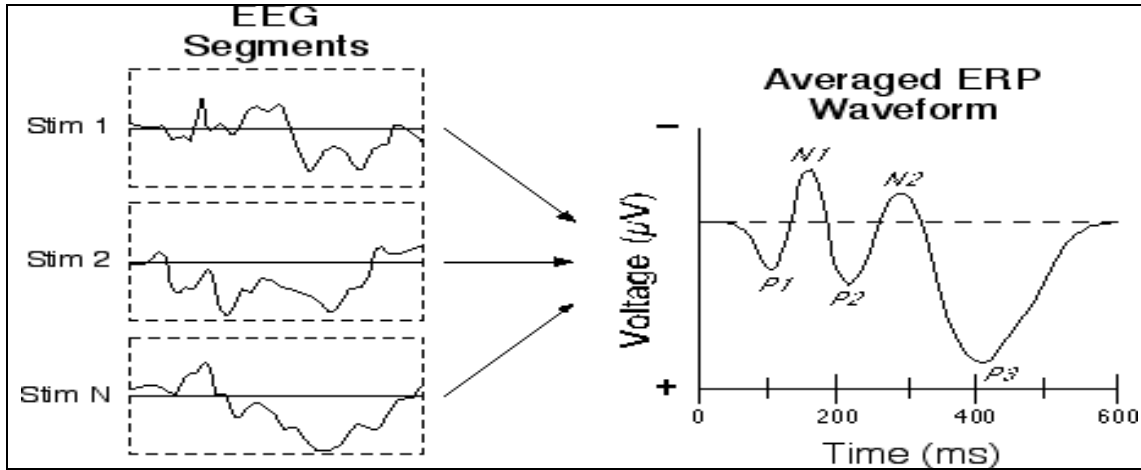
Supposing that the evoked potential is produced the same way in every trial we have:

$$s_i(t) = s_j(t) = s(t) \quad \forall i, j \in \{1, \dots, N\}$$

On the other hand noise can be considered as uncorrelated white Gaussian noise so:

$$E(r_i(t)) = \frac{1}{N} \sum_{i=1}^N s_i(t) + \frac{1}{N} 0 = \frac{1}{N} N s(t) \Rightarrow E(r_i(t)) = s(t)$$

In that way the signal related to the stimulus can be retrieved.

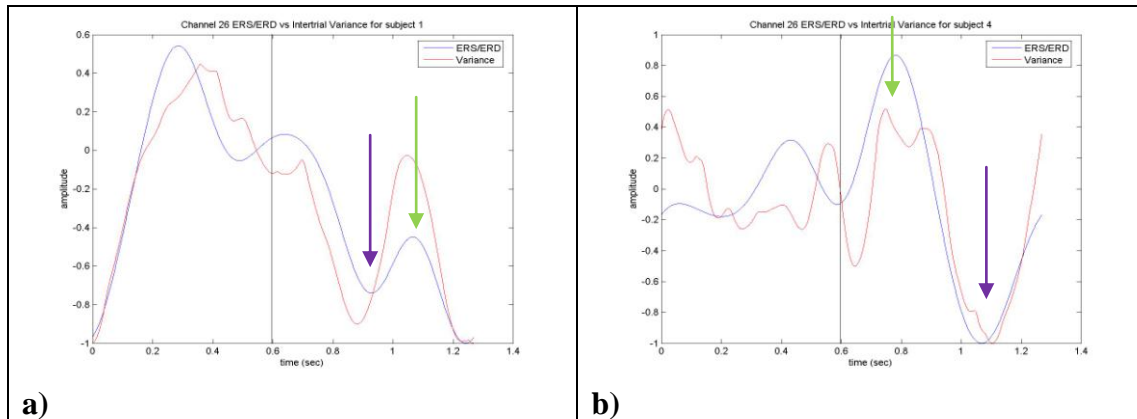


**Fig. 3.7: ERP averaging procedure.**

### 3.6.1 ERD-ERS

Several kinds of events, the most notably being sensory stimuli, can induce time-locked changes in the activity of neuronal populations that are generally called event-related potentials (ERPs). We previously discussed that in order to detect such ERPs, averaging techniques are commonly used. The basic assumption is that the evoked activity, or signal of interest, has a more or less fixed time-delay to the stimulus, while the ongoing EEG activity behaves as additive noise. The averaging procedure will enhance the signal-to-noise ratio. However, this simple and widely used model is just an approximation of the real situation. Indeed it is known that evoked potentials (EPs) can be considered to result from a reorganization of the phases of the ongoing EEG signals. In addition it was also shown that visual stimuli can reduce the amplitude of the ongoing EEG amplitude, thus demonstrating that the model assuming that an ERP can be represented by a signal added to uncorrelated noise does not hold in general. Furthermore, it is known since Berger (1930) that certain events can block or desynchronize the ongoing alpha activity. These types of changes are time-locked to the event but not phase-locked, and thus cannot be extracted by a simple linear method, such as averaging, but may be detected by frequency analysis. This means that these event-related phenomena represent frequency specific changes of the ongoing EEG activity and may consist, in general terms, either of decreases or of increases of power in given frequency bands. This may be considered to be due to a decrease or an increase in synchrony of the underlying neuronal populations, respectively. The former case is called event-related

desynchronization or ERD, and the latter event-related synchronization (ERS) [93], [142].



**Fig. 3.8: ERD/ERS on two EEG signals for alpha band.**

In figure 3.8 we can see two different signals analysis of their ERD/ERS activity in alpha band. So as to take the signals' alpha activity we bandpass filter the signal, because an EEG signal has many frequencies within it. The ERD/ERS measure actually estimates the power of the signal first and then takes its intertrial average so as to preserve its non-phase locked activity. The ERD/ERS measure is shown by the blue lines in the two figures and the black vertical line shows the stimulus onset. In figure 3.8a we can see that we first have an ERD (purple arrow) and then an ERS (green arrow) in contrast to figure 3.8b where we have the exact opposite. The red signal in the figures above shows the intertrial variance of the EEG signal and what we can see is that it follows the ERD/ERS measure. What we do in order to estimate the intertrial average of the signal is that we take for every signal trial its variance from the average of a chosen channel's mean value and then we add the trials.



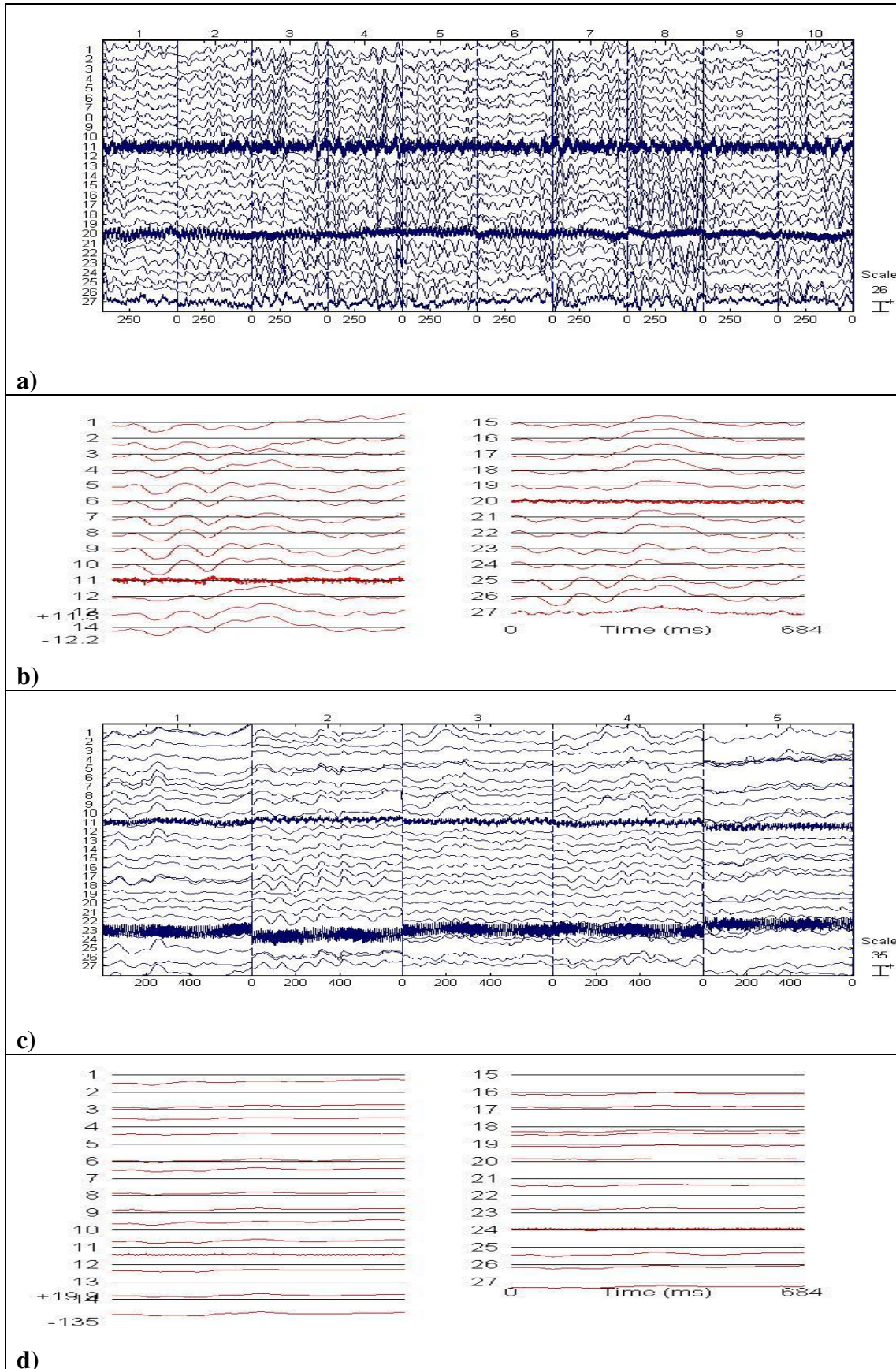


Fig. 3.9: In 3.9a,c we have an EEG's epochs and in 3.9b,d we have the corresponding ERPs.



In figure 3.9 we demonstrate the trials and the ERPs of the two subjects EEG's whose ERD/ERS we estimated in figure 3.8. In 3.9a we show ten trials (out of forty) one after the other (concatenated) for the 27 seven channels of the EEG and in 3.9c we show the first five trials, in order to have a clearer result. In the other two figures we have the ERPs of the corresponding subjects' EEGs on the 27 channels. In 3.9b the results can be seen clearly because the range of the values is small (minus twelve to eleven) while in 3.9d the ERPs cannot be seen clearly because we have a big range of values (minus one hundred and thirty five to twenty). From those figures one can understand that the EEG analysis is a complex and demanding procedure because we have many channels, many trials and the signals are not simple to analyze.

# CHAPTER 4: Proposed Methodology

Over the past few years there has been an increased interest in studying the underlying neural mechanism of cognitive brain activity related to memory. In this direction, we study the brain activity based on its independent components instead of the EEG signal itself aiming towards identifying and analyzing induced responses being attributed to oscillatory bursts from local or distant neural assemblies, with variable latency and frequency, in an auditory working memory paradigm. The contribution and functional coupling of independent components to evoked and/or induced oscillatory activities is investigated through the concept of the recently introduced partial directed coherence method, which can also reveal the direction of the statistically significant relationships. The results on read data from an oddball experiment are in accordance with previous psychophysiology studies suggesting increased phase locked activity most prominently in the delta/ theta band, while alpha is also apparent in measures of non phase-locked activity. Dynamic synchronization is inferred between the alpha and delta bands, whereas some influence of the theta band is also detected. This study indicates that functional connectivity during cognitive processes may be successfully assessed using spectral power measures applied on independent components, which reflect distinct spatial patterns of activity.

Event related brain dynamics entail a variety of activations and oscillations, from phase resetting of ongoing EEG activity in the alpha and theta bands [99] to phase-locked evoked and non phase-locked induced oscillations especially in delta, theta and gamma bands [36], [102]. Their origins relate to multiple task conditions and many stimulus types engaged during the event presentation and execution of its consequent actions [100], which define distinct brain functions, some operating independently and some being coupled [101].

Since we are interested in identifying distinct signal components and analyzing their coupling, we focus on decomposing the EEG signal into ICA components stemming from different brain regions and then analyzing their time-frequency content throughout multiple trials. The first aspect of analysis has been partially addressed with measures that can reveal phase locking effects [39], [98]. Besides these measures, we introduce a metric for considering stimulus-locked but not phase-locked activity. The second aspect related to synchronization is addressed through the PDC measure to reveal coupling characteristics.

## 4.1 Data acquisition and test description

The EEG signals used in this work arise from ten representative subjects (5 healthy, 5 AD) out of eighteen healthy and AD participants (age: 37-74, 9 healthy, 9 AD). The measurement involved 27 channels with linked ears (A1-A2) as the recording reference and electrode AFz as Ground. The signals were digitally sampled at 1024Hz, with a high pass filter of cut-off frequency 0.016Hz, a low pass filter of cut-off frequency 60Hz, and a notch filter at 50 Hz. Recordings were acquired from an auditory oddball experiment, where a stimulator provided 40 2 kHz target tones (20%) and 160 1 kHz non-target tones (80%). The inter-stimuli interval (ISI) was 1.29s.

## 4.2 Independent Component Extraction

Instead of directly measuring the synchronization using the actual EEG traces, independent components (ICs) were first obtained and then identified based on their spatial and frequency properties. To decompose the data into brain source activities, we used an Infomax ICA algorithm which minimizes mutual information among the data projections in order to achieve independence. Considering the multiple trials of an evoked response experiment, the decomposition schemes should be extended to reflect some form of consistency throughout the trials. Under the assumption of spatially consistent sources, the ICA decomposition can be performed in a concatenated trials scheme, with the EEG signal extended by one trial following the other, in the same way for each channel. Besides its increased stability and generalization capabilities, the concatenated trials approach has the add-on advantage of preserving the correspondence of components throughout the trials. Thus, the content of each ICA component can be analyzed in several perspectives including its topological origin, the time and frequency distribution, as well as its coherence over trials.

## 4.3 Independent Component Selection

Following their derivation, the components of ICA are often organized (or clustered) by means of multiple spatiotemporal constraints on their structure [99], [100]. Under the assumption of spatially consistent sources, we perform Infomax ICA [98] decomposition in a concatenated trials scheme, with the EEG signal extended by one trial following the other, in the same way for each channel. Besides its increased stability and generalization capabilities, the concatenated trials approach has the add-on advantage of preserving the correspondence of components throughout the trials. . In this way, the content of each ICA component can be analyzed in several perspectives including its topological origin, the time and frequency distribution, as well as its coherence over trials.

The ICA decomposition process results in EEG composite signals stemming from fixed locations and expressing a distinct temporal activation. Some of these components are expected to reflect phase locked activity, some others stimulus or response but not phase-locked activity and some others just express noise processes. The similarity of components across trials has been initially studied through clustering of the latency of maximum activation and identification of similarities in the “brain maps” of individual components [98], [126]. Alternatively, similarity of the nature of components across trials can be studied using their time-frequency energy distribution maps throughout a large number of trials often recorded from the repetition of the same experiment. Indeed, the time frequency spectrum of the average component/signal over trials reflects a measure of consistency in the phase-locked energy. In a related single-trial approach, the authors in [36], [39] search for similar patterns of activation (clusters) in the PCA decompositions of all single-trial time-frequency spectra of components, in order to identify strong and consistent patterns along trials. These methods form indirect approaches to measuring intertrial coherence through the spectral energy of an average signal, which might be justified only for phase-locked activity. In any case, the measure of spectral energy does not necessarily reflect signal synchronization across trials. Indeed, a strong activity in just

a few trials can induce significant spectral energy, but without providing any indication of synchronization among trials.

In this thesis, we employ yet another scheme which considers phase synchronization of each component across trials. For phase locked synchronization we use the intertrial coherence measure [39], whereas for non-phase locked activity we introduce the so-called shift-phase intertrial coherence. The former is based on the phase similarity of phase-locked components, whereas the latter is based on the power similarity of same structure but not phase-locked components across trials, which is an extension of the power measure used in ERD/ERS detection [30]. Thus, the proposed shift-phase intertrial coherence is based on the energy of single-trial decompositions and highlights frequency bands of increased energy in all trials. The intertrial coherence measures can be computed for the signal itself, or its time-frequency decomposition, deriving a trial-synchronization map complementary to that of the time-frequency spectrum. In fact the intertrial coherence map is a direct reflection of the synchronization pattern of the signal in time and frequency, without resorting to average energy measures.

In our approach, we first label components based on the phase locking attributes over trials. Some components may involve phase-locked evoked responses, others may capture time but not phase-locked induced responses, while others could engage both types. The similarity of components across trials has been studied using their time-frequency energy distribution maps. Indeed, the time frequency spectrum of the average component over trials reflects a measure of consistency in the phase-locked energy.

#### 4.3.1 Phase Intertrial Coherence (PIC)

For the  $i$ -th trial, the phase shift is reflected as an exponential term in the Fourier transform representation, i.e.  $x_i[n] = x[n - n_0] \leftrightarrow X_i[k] = X[k]e^{-j\frac{2\pi k}{N}n_0}$ . If all trials are phase locked to the same  $n_0$ , then we can define the metric referred to as phase-locking factor or Phase Intertrial Coherence (PIC):

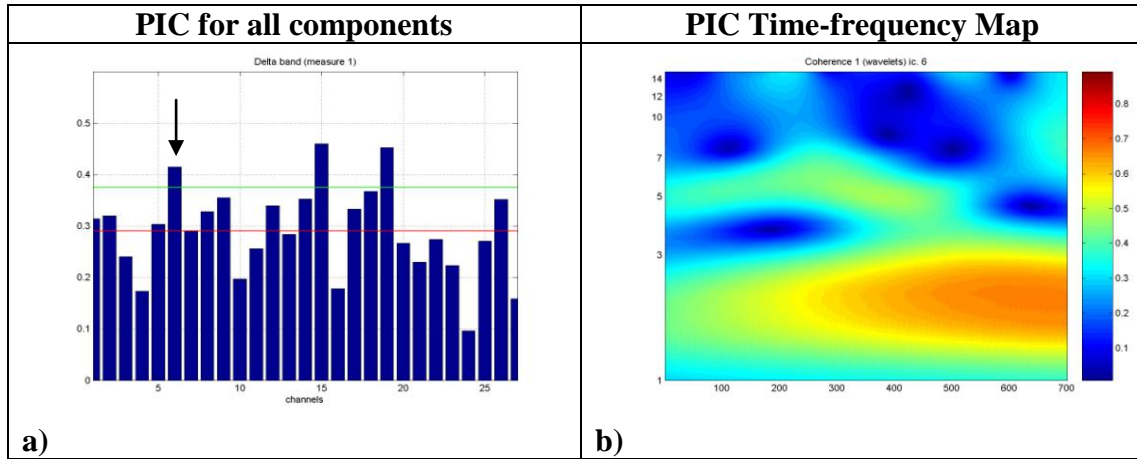
$$c[k] = \frac{\left| \sum_i X_i[k] \right|}{\sum_i |X_i[k]|} \leq 1$$

with equality holding if and only if the trials involve the same basic signal with the same shift. This metric is expanded to the time-frequency representation of a signal  $X_i[k, n]$ , with  $k$  and  $n$  indicating the frequency and time ticks, respectively.

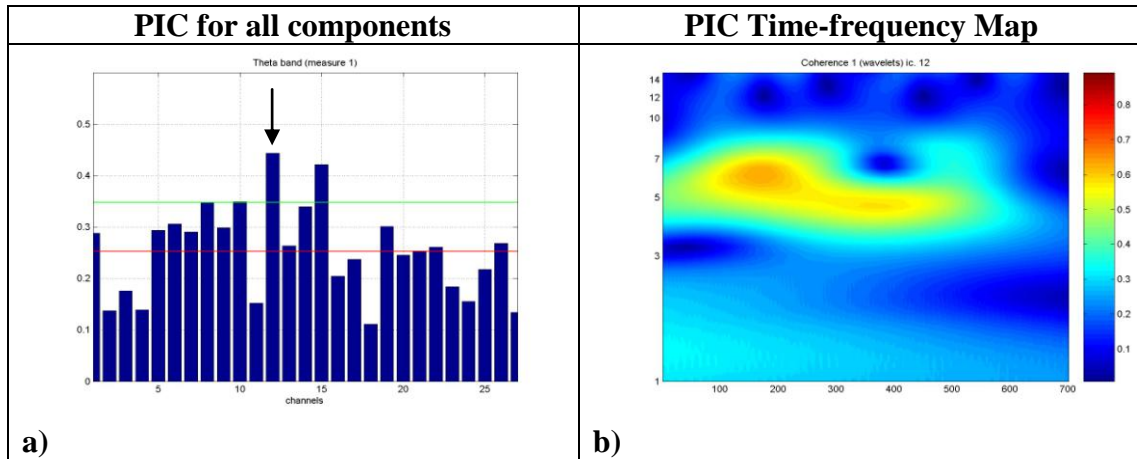
The intertrial coherence index for a specific band  $b$  is defined as the average in time

and scale of  $c[k, n]$ , i.e.  $c_b = \frac{1}{N_b} \sum_{k=1}^{N_b} c[k] \leq 1$ , where  $N_b$  the number of frequency (or

scale) ticks of interest, with equality implying perfect phase locking within the band. Summarizing, this measure reaches large values (close to 1) either in its frequency, or time-frequency application for phase-locked activity.



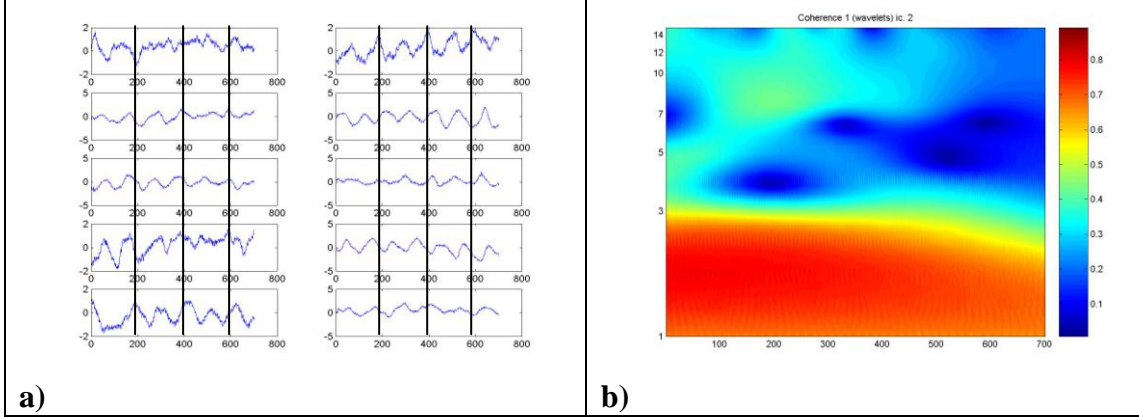
**Fig. 4.1:** a) the overall PIC for all the components in delta band, b) the TF map of component 6.



**Fig. 4.2:** a) the overall PIC for all the components in theta band, b) the TF map of component 12.

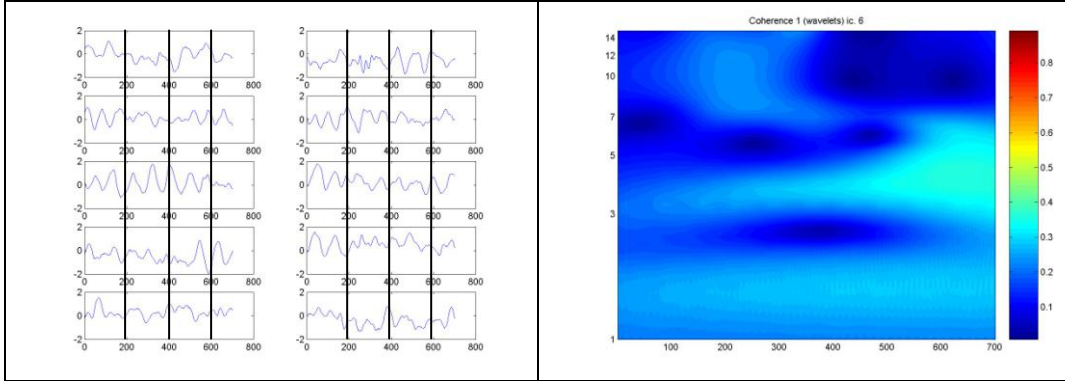
In figures 4.1a and 4.2a above we can see the PIC measure utilized as global metrics on 27 components, measuring their intertrial coherence (they are demonstrated as bars) in delta and theta bands correspondingly. The horizontal axis in TF maps goes up to 700ms after stimulus and the vertical axis goes from 1 to 14Hz of the logarithmic scale. It has to be mentioned that the bars represent certain frequency bands, though the frequency bands have many frequencies within them. In order to have one bar for each component we use the average of the corresponding frequency bins. Alternatively, PIC metric can be estimated for each tick in the time-frequency representation for specific components in figures 4.1b and 4.2b, in order to provide timely localized maps of the component coherence over trials. Thus, it is a reasonable thought to depend on the bar diagram (fig. 4.1a and 4.2a, the black arrow shows the component which have been chosen to make their TF map) so as to make a first choice of the components and then make their TF maps in order to refine this selection. In the figures above we can see that the information given by the bar diagram that the specific components have a strong activation in the corresponding frequency bands was accurate. We decide to choose these components as they have an

activation limited to the frequency band limits, in other words they have quite unmixed content.



**Fig. 4.3: a) Ten trials of an ICA component, b) TF-map of PIC for corresponding component.**

In figure 4.3a there are illustrated ten out of forty samples for the component we made the TF-map of PIC in figure 4.3b. Although the signals in figure is not band passed so as to see only the delta band, it can be observed that the slow waves of the trials are basically synchronized, which confirms the TF-map result of figure 4.3b.



**Fig. 4.4: a) Ten trials of an ICA component, b) TF-map of PIC for corresponding component.**

In contrast to figure 4.4 here the slow waves of the trials are not synchronized (fig. 4.4a), which confirms the TF-map result of figure 4.3b where we do not seem to have strong phase locked activity.

### 4.3.2 Phase-shift Intertrial Coherence (PsIC)

In case of the same basic signal with different shifts from trial to trial, there is a different exponential term remaining in the Fourier transform of each trial, i.e.

$x_i[n] = x[n - n_i] \leftrightarrow X_i[k] = e^{-j\frac{2\pi k}{N}n_i} X[k]$ . Similar to metrics defined for ERD/ERS activity [2] in time-locked responses, the proposed metric utilizes the power of activity instead of the signal's value. For phase-shift responses, this metric eliminates the complex phase effects and compares the intertrial content only based on its power in specific frequency bands.

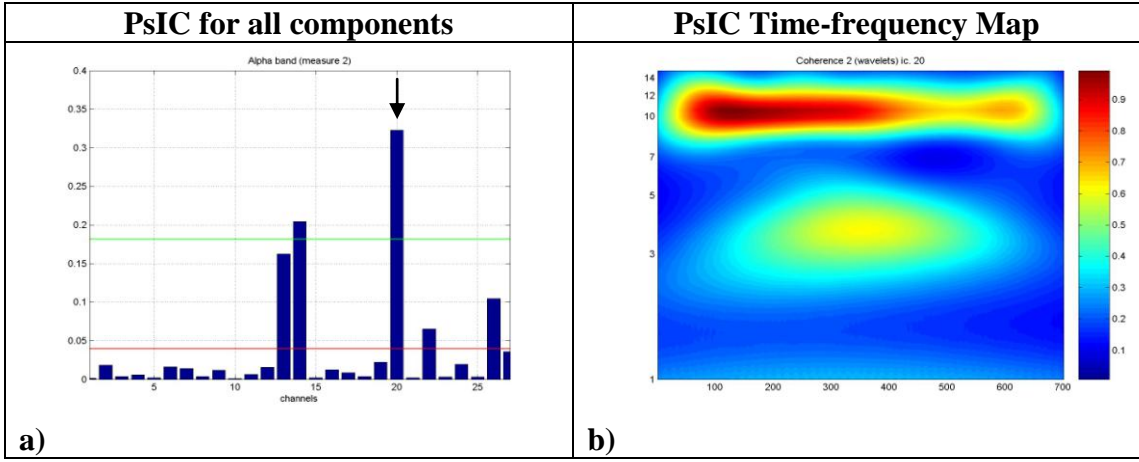
Initially, we tested the second-order measure similar to PIC above in the form of:

$$c[k] = \frac{\left[ \sum_i |X_i[k]| \right]^2}{\sum_i |X_i[k]|^2} \leq 1, \text{ which holds true due to the triangle inequality. The}$$

discrimination of this metric, however, is very low for a wide range of signal coherence over the trials. In order to utilize the power of the signal over the trials, the so-called phase-shift intertrial coherence (PsIC) metric is thus defined as:

$$c[k] = \frac{\sum_i |X_i[k]|^2}{\max_k \sum_i |X_i[k]|^2} \leq 1.$$

Similar to PIC, this metric can be expanded to the time frequency domain as  $c[k,n]$  and for a specific band  $b$  the PsIC measure is defined as the average in time and scale of  $c[k,n]$ .



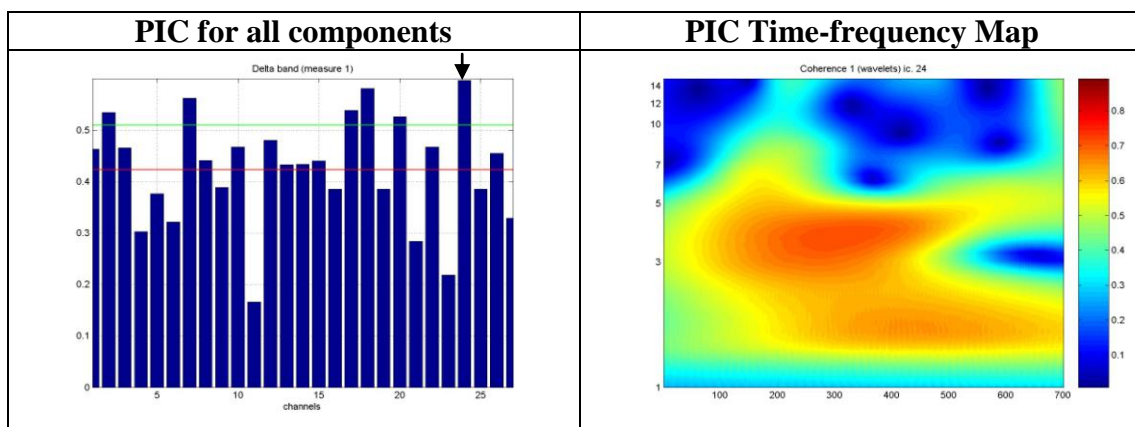
**Fig. 4.5: a) the overall PsIC for all the components in alpha band, b) the TF map of component 20.**

In figures 4.5a above we can see the PsIC measure utilized as global metrics on 27 components, measuring their intertrial coherence (they are demonstrated as bars) in alpha band. Alternatively, PsIC metric can be estimated for each tick in the time-frequency representation for specific components (figure 4.5b), in order to provide timely localized maps of the component coherence over trials. Thus, it is a reasonable thought to depend on the bar diagram so as to make a first choice of the components and then make their TF maps in order to refine this selection. In the figures above we can see that the information given by the bar diagram that the specific component has a strong activation in the corresponding frequency band was accurate.

### 4.3.3 Utilization of Intertrial Coherence Metrics

Both the phase and the shift-phase locking factors can be utilized as global metrics on a component, measuring its intertrial coherence in a specific band. Alternatively, these metrics can be computed for each tick in the time-frequency representation of the component, in order to provide timely localized maps of the

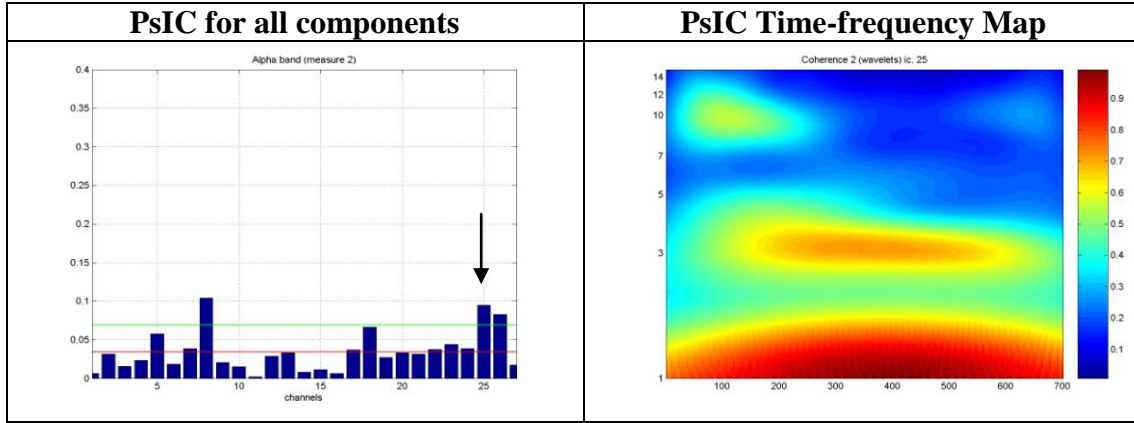
component coherence over trials. In the second form, the intertrial coherence map conveys information complementary to the time-frequency spectrum of the average component. Indeed, the latter preserves the phase-locked energy of the component over all trials, whereas the former reflect the time-locked activity with or without phasic coherence across trials. All three maps will be used for the characterization of relevant components as each one emphasizes on different aspects of synchronous activity. For the selection of components with relevant activity in each band we will resort to the global indices of intertrial coherence, which can be plot and compared in a single diagram for all derived components. For the detailed analysis of selected components, we will consider the two-dimensional maps of the three measures that reflect the component consistency (in phase and/or power) along trials. An example follows showing inappropriate components, which even though they seem to have strong activity in the bar diagram of the global metrics, the TF-map shows mixed activity which is not something we want in a component for our analysis.



**Fig. 4.6: a) the overall PsIC for all the components in delta band, b) the TF map of component 24.**

In the figure above (4.6a,b) we can see that component 24 has a strong phase-locked delta activity, so a first thought is that it is ideal to be a delta band representative. Though, we can see in its TF map that it also has theta phase locked activity. This fact makes this component improper for our analysis as it has mixed activity.





**Fig. 4.7: a) the overall PsIC for all the components in alpha band, b) the TF map of component 25.**

In figure 4.7a,b we can see that component 25 has a strong non-phase-locked alpha activity, so a first thought is that it is ideal to be an alpha band representative. Though, we can see in its TF map that it also has theta and delta non-phase-locked activity. This fact makes this component improper for our analysis as it has mixed activity.

#### 4.4 Functional Synchronization of Components

Integrated and complex brain activities have been considered as supported by single composite tasks, which include anticipation of, attention to the stimulus and preparation for its associated actions. Examples include the consideration of bottom-up visual perception with binding delta activities, visual or auditory oddball experiments, motor imagery following a visual trigger, etc. [36], [102], [126]. In general, neural assemblies synchronize and interact dynamically in local or distant regions in order to accomplish perceptual, motor or cognitive functions [128]. The dynamics underlying the coupling of the oscillatory activities among such assemblies can highlight the synchronization mechanisms among brain sources during the performance of a certain activity, such as visual perception, object recognition and categorization, motor activity or motor imagery (as applied to BCI). Studies with more detailed MEG signals have revealed local synchronization patterns and cortico-cortical interactions involved in several cognitive operations [38]. Since composite subtasks are triggered by different brain sources and subsequently synchronize to complete the task, the dynamics of interaction among independent components may be used for indexing neural synchrony of such local or distant brain sources [128].

In this thesis, components of particular interest are examined for their functional synchronization within the brain neuronal structure in quest of revealing a sequence of distinct brain operations closely coupled in the execution of a single task (cognitive, mental or motor, internally or externally triggered). For instance, in the auditory oddball task under consideration we expect to distinguish phase locked theta and delta activity, as well as non-phase locked alpha activity with temporal activation at both P3a and P3b responses.

### 4.4.1 Partial Directed Coherence (PDC)

Only those components associated with event-related activity (evoked or induced) are further analyzed in terms of their dynamic coupling through PDC, which is based on the commonsense idea that causes precede their effects in time and is formulated in terms of predictability. In a linear framework, Granger-causality is commonly evaluated by fitting Vector Autoregressive Models. Suppose that a set of  $N$  simultaneously observed time series:

$$\mathbf{x}(t) = [x_1(t), \dots, x_N(t)]^T$$

is adequately represented by a Vector Autoregressive Model of order  $p$  (MVAR(p)):

$$\mathbf{x}(t) = \sum_{k=1}^p \mathbf{A}_k \mathbf{x}(t-k) + \mathbf{w}(t)$$

where  $\mathbf{A}_k = \begin{bmatrix} a_{11}(k) & \cdots & a_{1N}(k) \\ \vdots & \ddots & \vdots \\ a_{N1}(k) & \cdots & a_{NN}(k) \end{bmatrix}$  is the coefficient matrix at time lag  $k$ , and

$$\mathbf{w}(t) = [w_1(t), \dots, w_N(t)]^T$$

is the vector of model innovations. Let

$$A(\lambda) = \sum_{k=1}^p \mathbf{A}_k e^{-i2\pi\lambda k}$$

be the Fourier transform of the coefficient matrices, where  $\lambda$  is the normalized frequency in the interval  $[-0.5, 0.5]$ . Then the PDC is defined as [96]:

$$|\pi_{i \leftarrow j}(\lambda)| = \frac{1}{\sigma_i} |\bar{A}_{ij}(\lambda)| \left/ \left( \sum_{m=1}^p \frac{1}{\sigma_m^2} \bar{A}_{mj}(\lambda) * \bar{A}_{mj}^H(\lambda) \right)^{1/2} \right.$$

Where

$$\bar{A}(\lambda) = I - A(\lambda)$$

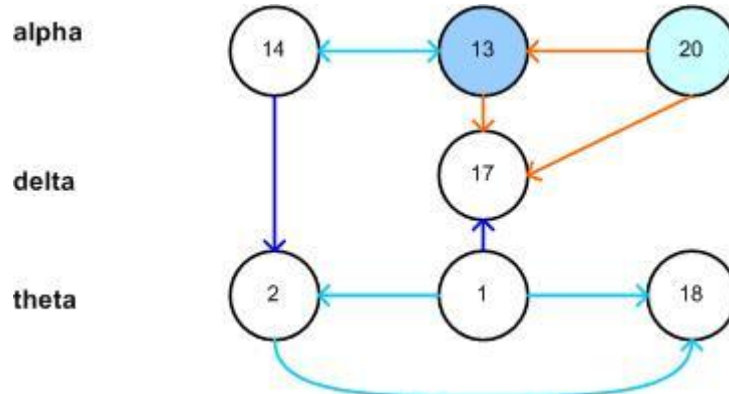
and

$$\sigma_i^2$$

refers to the variance of the innovation processes

$$w_i(t)$$

PDC ranges between 0 (indicating independence) and 1 (maximum coherence).



**Fig. 4.8: The PDC result graph.**

We can see in the figure above the result of PDC. There can be seen the “driver and response” relations between components of different or same frequency bands. The light blue lines imply relations between components of the same frequency bands and dark blue lines imply relations between components of different frequency bands. The light blue nodes imply early alpha activity and the darkest blue ones imply late alpha activity. The orange lines imply relations between delta and alpha band which correlate late and early alpha components with delta components.

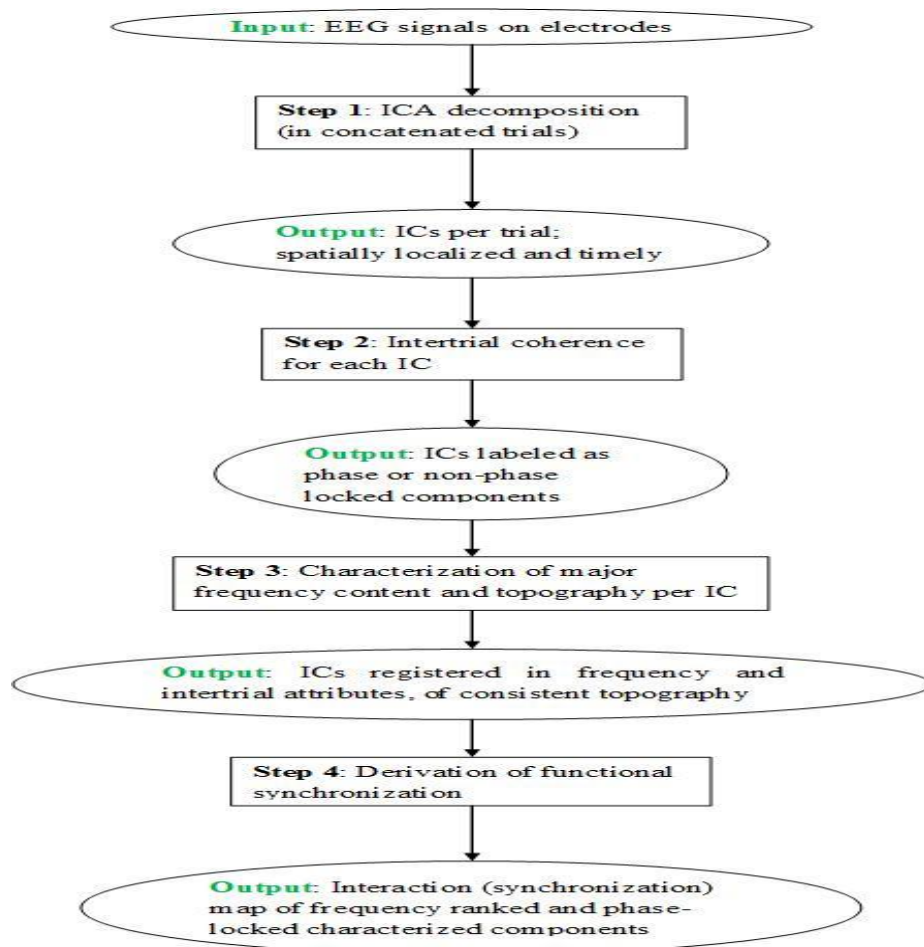
## 4.5 Summary of Algorithmic Steps

The EEG signal processing is an extremely complicated procedure itself. In our work we tried to organize the EEG principal components according to their frequency content and reveal the relationships between them via PDC method. The goal of organizing the principal components was succeeded by applying the coherence measures we suggested, on the time-frequency content of the components. The TF maps of the components gave us the opportunity to choose the components which had valuable information for our analysis, cluster them as delta, theta or alpha components and then be able to describe and explain the relations which came up after applying PDC.

The previous studies on ERP activations reveal that they are composed of different components (others coupled and others oscillating independently) which originate from specific brain regions and oscillate at a dominant frequency band, which are characteristics of the particular (sub)task that this component is enabling. Single-trial ERP component analysis has been used primarily for relating a specific task (sensory, motor or cognitive) with the topography and/or frequency content of its components [126], rather than associating those components in terms of their phase-locking or synchronization attributes.

This thesis attempt was to provide an integrated view of ICA components in terms of their organization, their involvement in brain processes and their coupling in the time course of activities. Along this direction, the proposed methodology involved derivation of ICs, selection of interesting components that preserve either phase-locked or non phase-locked characteristics related to the event, analysis of content and organization of these components in terms of their frequency attributes and, finally, study of their dynamic coupling (synchronization) and derivation of an interaction map of components per frequency band, which is a new approach in contrast to previous studies which only derive components and analyze their content [39], [98], [127].

Thus, the overall contribution of our study was twofold: i) to identify components that express consistent pattern of activation related to the event and ii) to derive the synchronization pattern of components oscillating at three specific frequency bands (delta, theta, alpha). The first aspect of analysis has been partially addressed with measures that can reveal phase locking effects. Besides these measures, we introduced a metric for considering stimulus-locked but not phase-locked activity. The second aspect related to synchronization is addressed through the PDC measure to reveal coupling characteristics. In characterizing the components, we emphasized more on their frequency attributes than the topography, since the latter expresses large variation among subjects and trials. Nevertheless, the general location of components in terms of wider brain lobes, and the spatial localization of per-band oscillations was also addressed in this work. The various steps of analysis are briefly outlined here:



**Fig. 4.9: The analysis steps.**

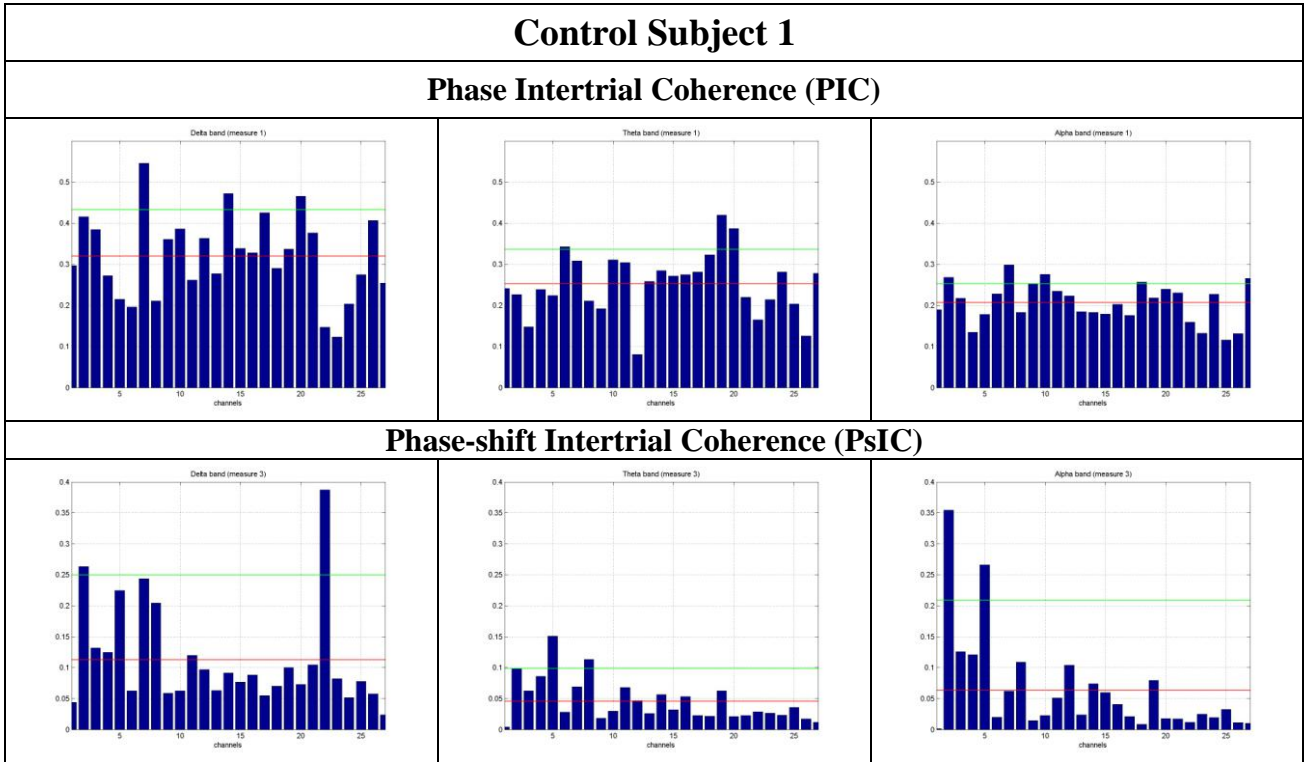
# CHAPTER 5: Results

So as to demonstrate the results of our method and the steps of the methodology itself, the results of four people will be analytically presented. Two of those will be normal people and the other two will be AD patients. The rest of subject analysis is presented in the appendix. The statistics for all subjects are presented in section 5.2 and the the results analysis in section 5.3 (this analysis refers to all the subjects).

We will try to show the phase and non-phase locked activation of the frequency bands of our interest (delta, theta and alpha band) using the coherence measures we proposed on the time-frequency maps of the components. According to that information we will organize the components and find their relations using the PDC method. The procedure which follows shows step by step the process of our methodology.

## 5.1 Application on Real Data

### 5.1.1 First Control Subject Analysis

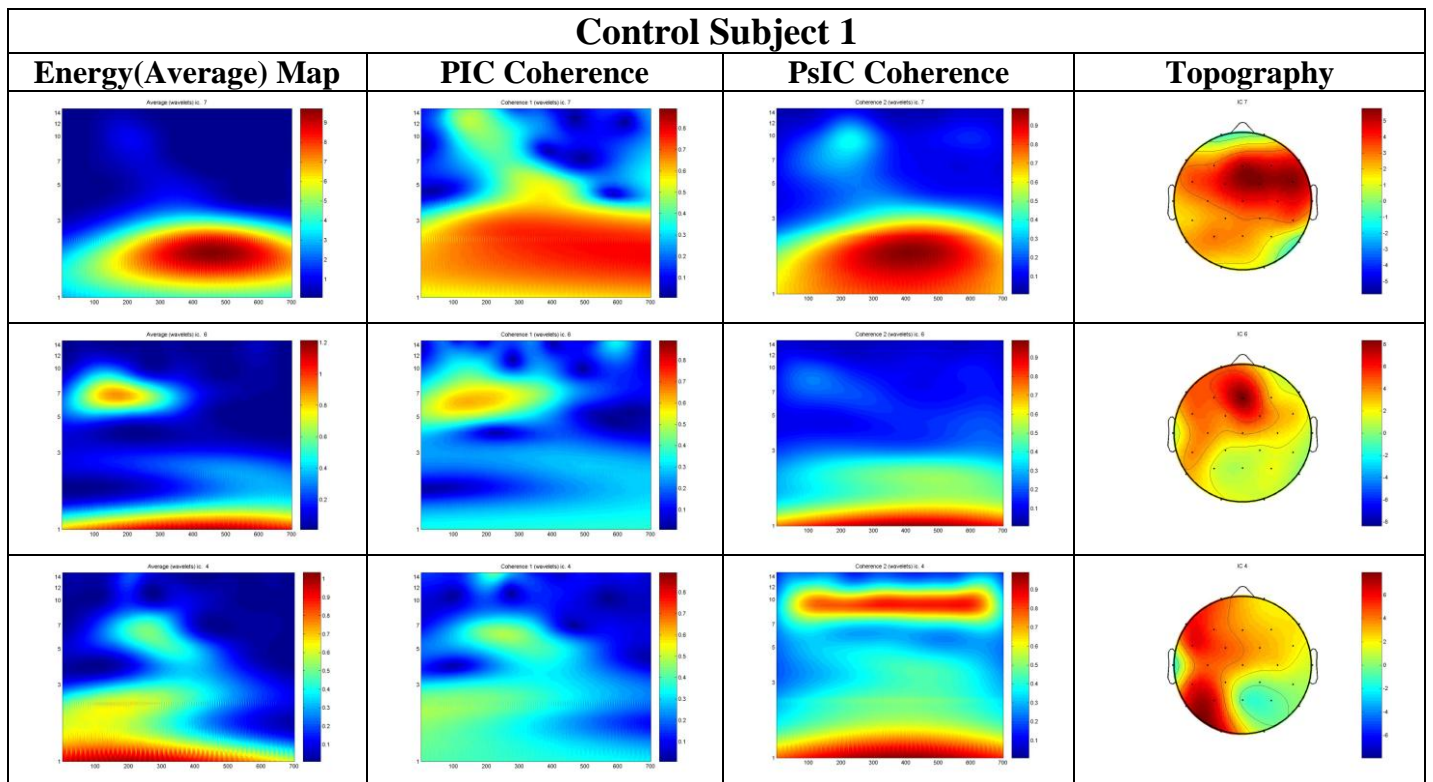


**Fig. 5.1: Components bar diagrams for delta, theta and alpha band.**

After the ICA is applied on the concatenated trials of the eeg signal for each person, we have to decide which components of the signal are important for further analysis (TF-representation, PDC). In this direction the measures we previously proposed are very helpful. We apply these measures on the FT of the signal of the independent components so as to have a view of which components have greater phase or non-phase locked activity in the frequency bands of our interest which in particular are delta, theta and alpha band.

In figure 5.1 we can see which components are phase or non-phase locked in the particular frequency band of our interest. The first row shows Phase Intertrial Coherence (PIC) for every channel in delta, theta and alpha frequency band and the second row shows Phase-shift Intertrial Coherence (PsIC) for every channel in delta, theta and alpha frequency band respectively. In the horizontal axis we can see the component and in the vertical axis there is the coherence. By observing the bar diagrams we see that some of the components happen to have activation in more than one frequency bands such as the second components which has an alpha non-phase locked high activation but it also has a big enough face-locked delta activation. Such cases, give an extra difficulty in the organization of the components in delta, theta and alpha band and rises up the disadvantage of ICA which is not always successful in separating different frequency bands.

By observing the bar diagrams above, we can see that we have plenty of powerful components in PIC for delta and theta band, and not many powerful components in alpha band. On the other hand, we see that one component (component no. 22) has an efficient delta non-phase locked activity (PsIC measure) without having a corresponding delta phase locked activity (PIC) which refers to an artifact component. We do not seem to have special non-phase locked activity in theta band, though we have some important non-phase locked components in alpha band.



**Fig. 5.2:** The maps of three intertrial coherence measures and their scalp topography. First row (cmp. 7), second row (cmp. 6) and last row (cmp. 4).

For the components selected from the bar diagram, the next step is to find their time-frequency energy distribution maps and use the PIC and PsIC measures on them so as to show their activation in the frequency bands of our interest. We have to make clear that for the Energy(Average) measure in the first column of the figure above we

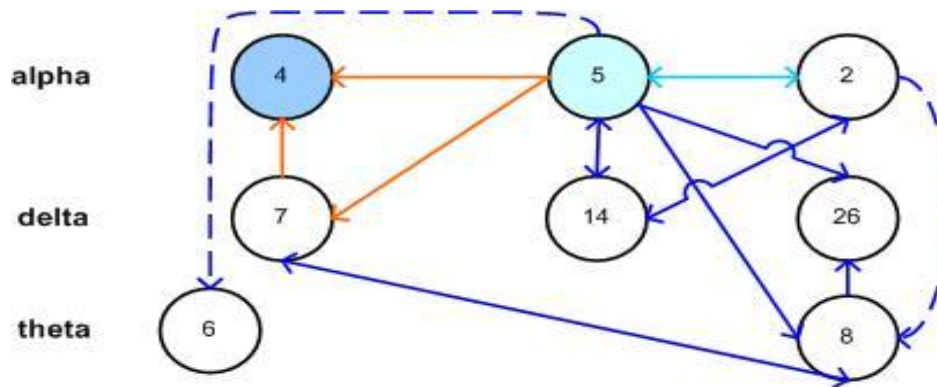


mean  $c[k] = \left| \sum_i X_i[k] \right|^2$  for a signal  $X_i[k]$ , which does not preserve the non phase

locked activity. As we see in the figure 5.2 the components have a quite clear activation in the frequency bands they belong to. Though, this is not always the case as many times the components have mixed activation in different frequency bands.

As far as the colour-maps are concerned for PIC and PsIC coherence they range from zero to one, the Average Energy Map colour-map ranges from the minimum to its maximum value and the brain topography which represents the activation in the brain using the ICA unmixing matrix  $W^{-1}$  has a colour-map which depending to the better visualization result it is from the maximum to the minimum value of the component or from  $-\text{maximum}$  to the maximum value. In every case, though the deepest red represents the maximum values and the deepest blue the minimum ones. The vertical axis of the frequency maps has the frequencies from one to fourteen in logarithmic scale and the horizontal axis has the samples beginning from the stimulus onset. Though, because our sample rate is 1024Hz, we can that for example the sample 700 corresponds to 700ms.

In the time-frequency maps represented in figure 5.2, we can see the three frequency bands of our interest (delta, theta, alpha) respectively. In the first row we can see an extensive phase locked delta activation, though it seems to be stronger from 300ms to 700ms which verifies that delta activity follows in time theta activity. We can see from the second row from the measure of PIC, that theta activity is mostly observed in the first 300ms. In the third row the PsIC measure shows the non-phase locked activation of the component, which in the first 300ms seems to be between 7 to 10Hz (early alpha activity  $\rightarrow$  attention) and after that it goes up to 12Hz (late alpha activity  $\rightarrow$  retrieval).



**Fig. 5.3: The PDC result graph which shows the relations between the chosen components.**

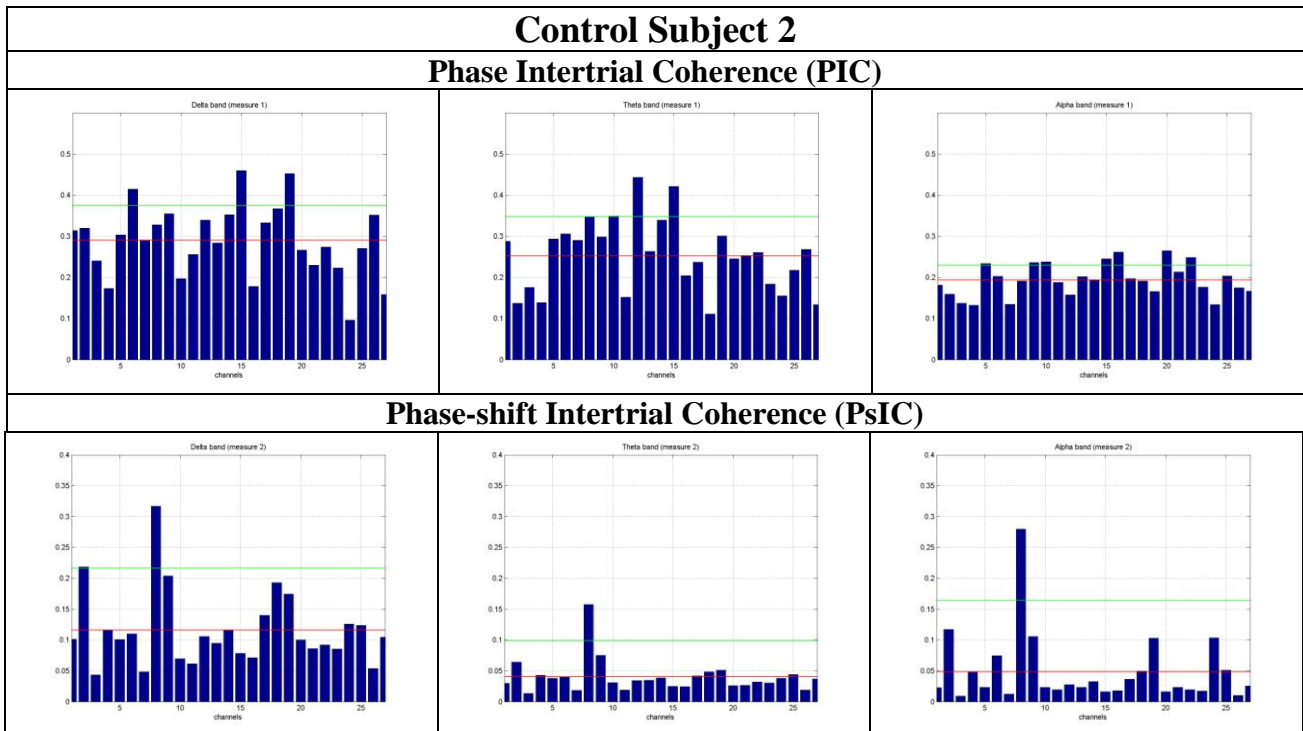
The components which are chosen according to their frequency content are further analyzed in terms of their dynamic coupling through PDC. The concluding result can be seen in the figure 5.3. In the graph there can be seen the selected components as nodes and the arrows between them represent their relations either one way or two-ways. The light blue arrows are between components of the same band and the darkest blue arrows are between components of different bands. The orange arrows show relations between delta and late-early alpha components. The light blue nodes are the early alpha components and the darkest blue nodes are the late alpha

components. Early and late characterizations concern the time when the activity seems to appear. The dashed lines show weak relations between two components.

By observing the graph we can see that consistently, early alpha components lead some theta components and, in turn, theta lead delta components. There are also two-way interactions between early alpha and delta pairs, as well as theta and delta pairs of components. Furthermore, there is a consistent tendency of early alpha and delta components to lead a late alpha component. This pattern is in close agreement with the proposed involvement of early alpha and theta bands in alertness and attention, followed by the subsequent stimulation of delta band and late alpha activations for the completion of cognitive tasks.

Some more results for three more persons are demonstrated on the following figures.

### 5.1.2 Second Control Subject Analysis

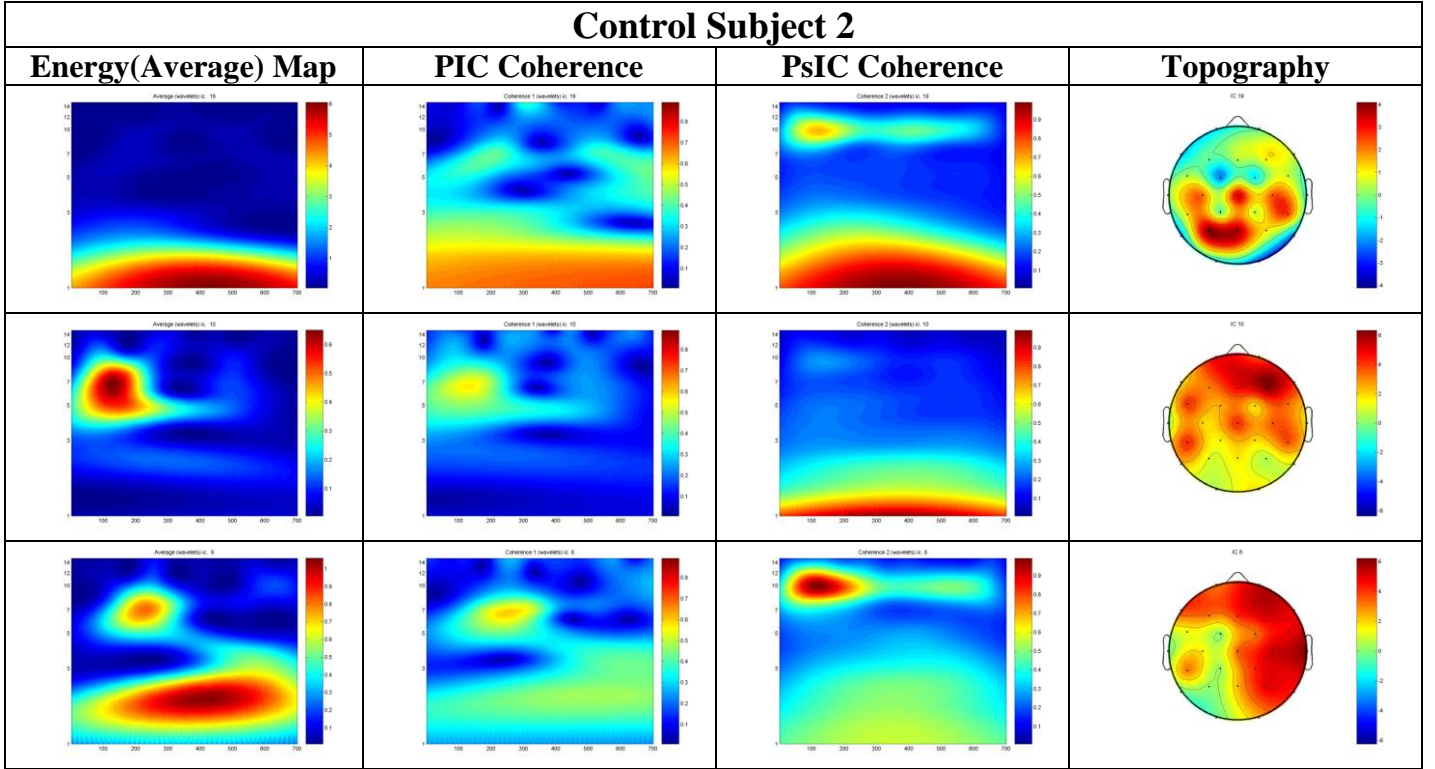


**Fig. 5.4: Components bar diagrams for delta, theta and alpha band.**

By observing the bar diagrams above, we can see that we have plenty of phase-locked components in PIC for delta and theta band, and less phase locked components in alpha band. On the other hand, we see that one component (component no. 8) has an efficient non-phase locked delta activity (PsIC measure) though it also has a corresponding phase-locked activity which leads us to think that it is not an artifact component. We do not seem to have special non-phase locked activity in theta band, though we have some important non-phase locked components in alpha band.

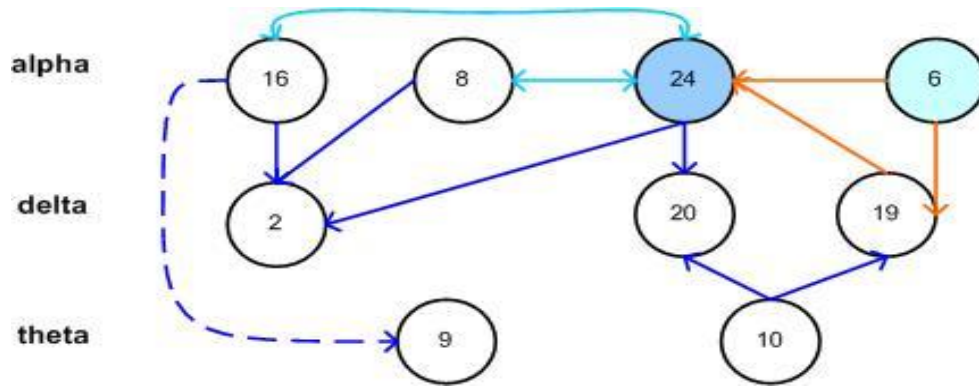


## Control Subject 2



**Fig. 5.5: The maps of three intertrial coherence measures and their scalp topography. First row (cmp. 19), second row (cmp. 10) and third row (cmp. 8).**

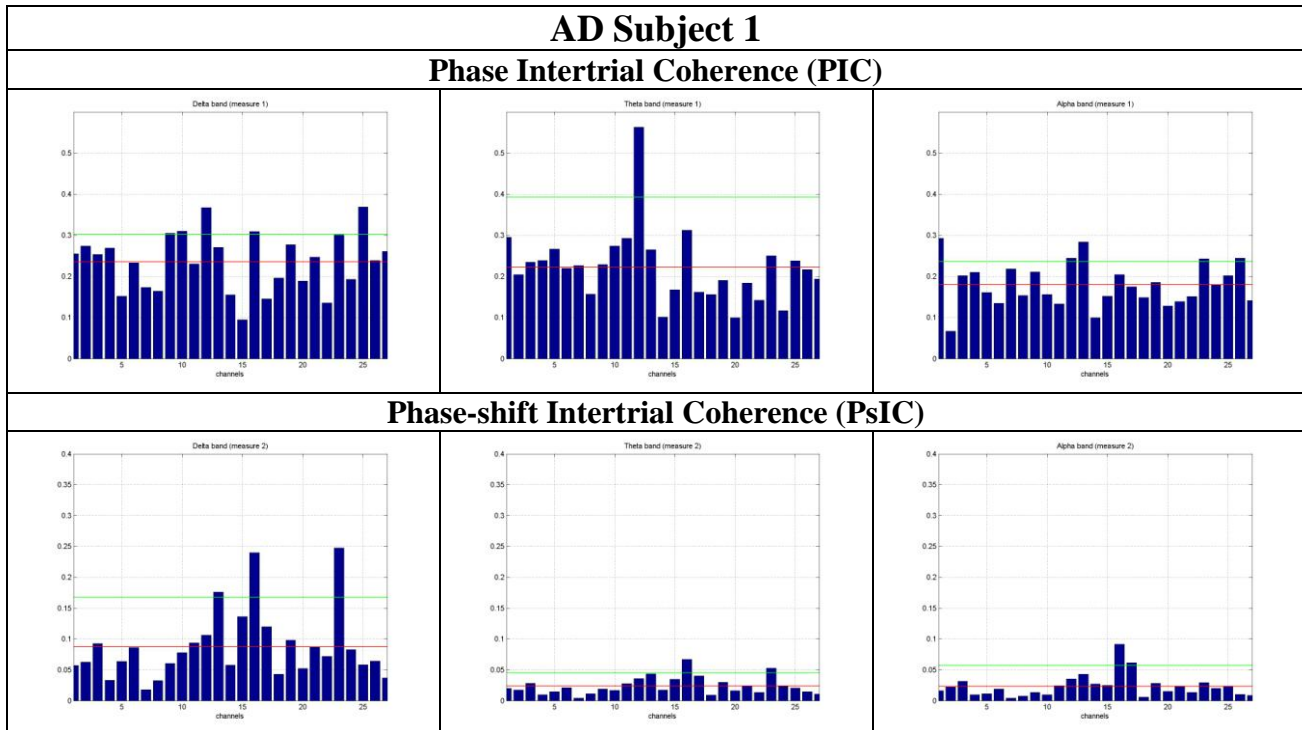
From the figure above we can see that each component was selected according to its frequency content so as to have one representative for each frequency band of our interest. In the first row the component we chose has an efficient delta phase locked activity which can be seen in the second column of PIC Coherence. Though, as we have already mentioned, this component does not only have delta activity as we ideally would like. It also has a weak theta phase locked activity (PIC Coherence) as well as non phase locked delta and alpha activity. In such cases the topography is also very helpful as we can see that it shows a posterior activation which is something that characterizes delta band activity. In the second row we have a theta band representative component, as both the PIC Coherence map and the topography stand to reason this clustering. The topography more specifically is anterior as expected for theta band activation. The last row shows an alpha non-phase locked representative component, something which is verified by the PsIC Coherence map mostly and the topography. It has to be mentioned that we have a quite mixed activation in the topography because as we see, there exists a weak activity of delta and theta band which is demonstrated by the PIC Coherence map.



**Fig. 5.6: The PDC result graph which shows the relations between the chosen components.**

Just like the synchronization graph of the first control subject in this graph we can see that consistently, early alpha components lead some theta components and, in turn, theta lead delta components. There are also two-way interactions between early alpha and delta pairs, as well as theta and delta pairs of components. Furthermore, there is a consistent tendency of early alpha and delta components to lead a late alpha component. This pattern is in close agreement with the proposed involvement of early alpha and theta bands in alertness and attention, followed by the subsequent stimulation of delta band and late alpha activations for the completion of cognitive tasks.

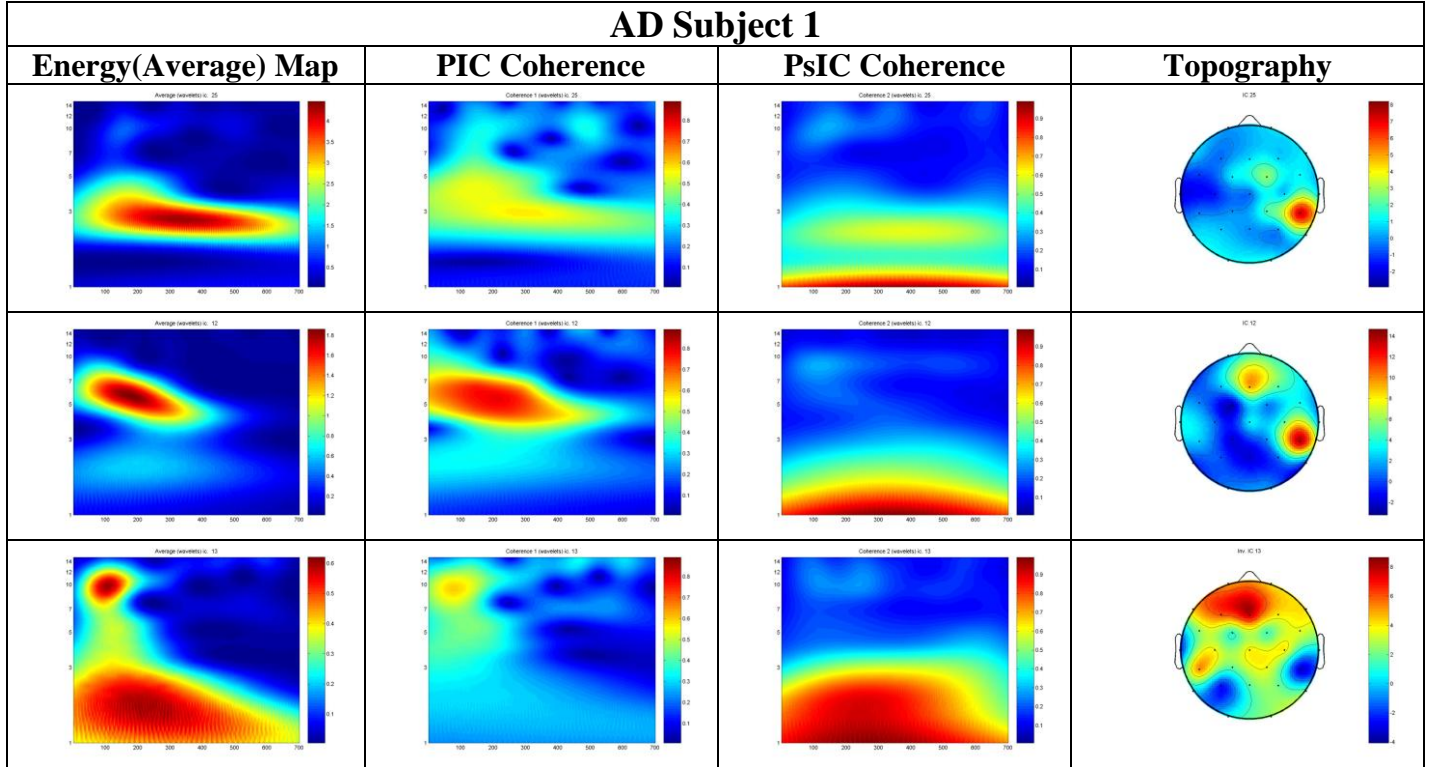
### 5.1.3 First AD Subject Analysis



**Fig. 5.7: Components bar diagrams for delta, theta and alpha band.**

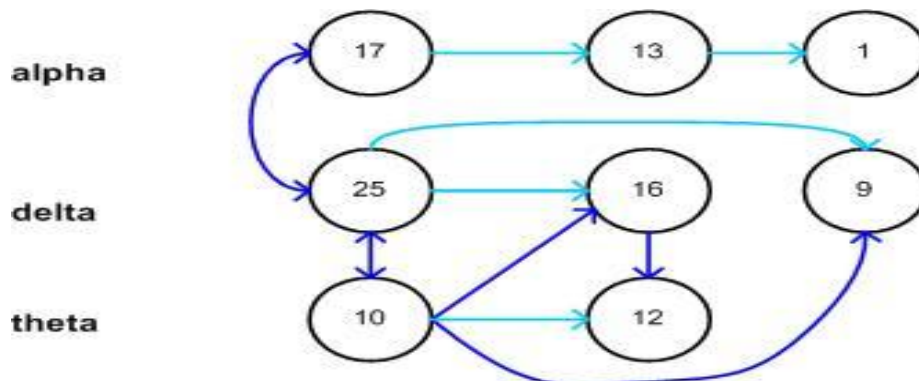
In contrast to the control subjects from the bar diagram of the AD patient above we can see that very few and weak alpha non phase locked components (PsIC). Also we seem to have quite a few theta phase-locked powerful components and less delta or alpha phase locked ones.

## AD Subject 1



**Fig. 5.8:** The maps of three intertrial coherence measures and their scalp topography. First row (cmp. 25), second row (cmp. 12) and third row (cmp. 13).

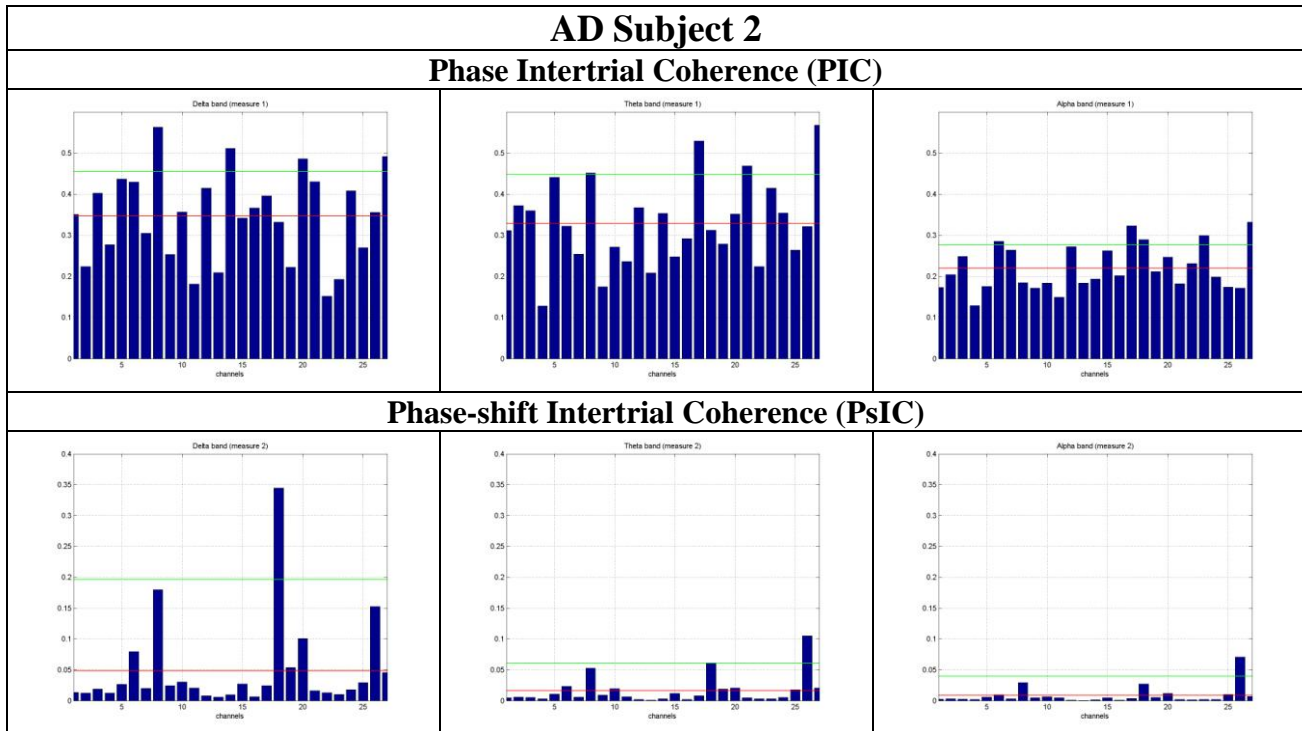
From the figure above we can see that each component was selected according to its frequency content so as to have one representative for each frequency band of our interest. In the first row the component we chose has an efficient delta phase locked activity which can be seen in the second column of PIC Coherence. Its topography seems to be posterior, which is something that characterizes delta activity. In the second row we have a theta band representative component, with an anterior topography activation. The last row shows an alpha phase locked representative component, as we do not have a strong enough component to demonstrate (bar diagram of alpha band in PsIC measure). In this components topography we also have a strong anterior activation which could be an eye artifact.



**Fig. 5.9:** The PDC result graph which shows the relations between the chosen components.

We can observe a close connectivity of theta components and a two-way interaction between the delta and theta components, with the theta band mostly driving the delta band. The alpha activity appears in this AD subject and though it only shows early alpha (lower alpha) oscillations or phase locked alpha activity. The alpha band oscillations demonstrate a large phase locked portions, at least appearing in the power of the average trial signal, as opposed to the control subjects that typically demonstrate non-phase locked oscillations in alpha band.

#### 5.1.4 Second AD Subject Analysis

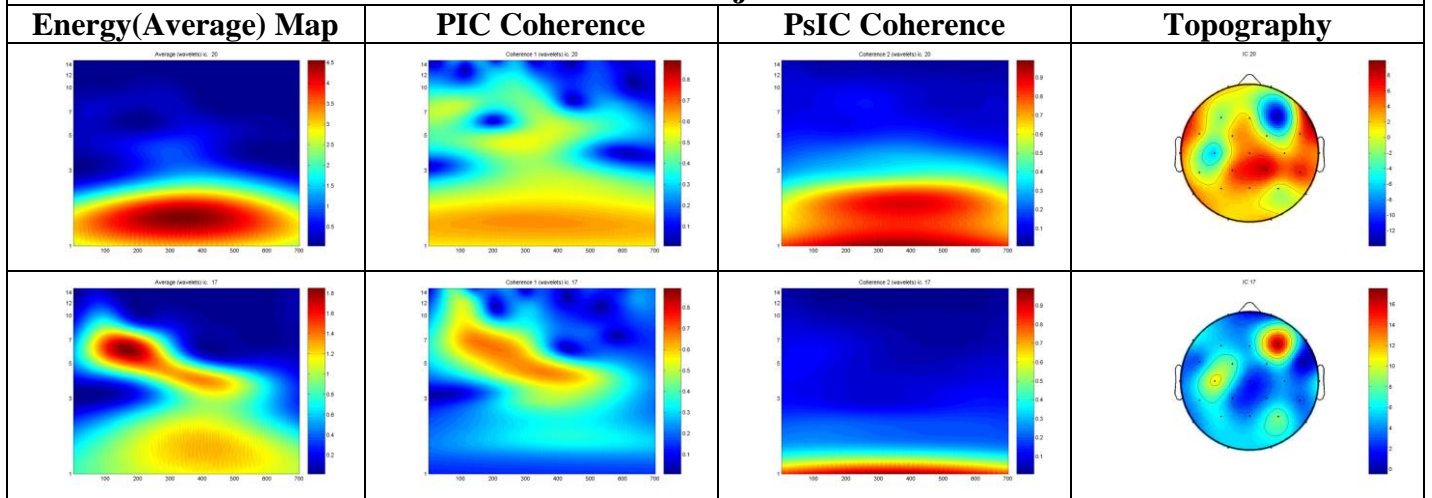


**Fig. 5.10: Components bar diagrams for delta, theta and alpha band.**

For this AD subject we can see that we have no components having efficient alpha non phase locked (PsIC) activation. Also we seem to have many theta and delta phase-locked powerful components and less alpha phase locked ones.

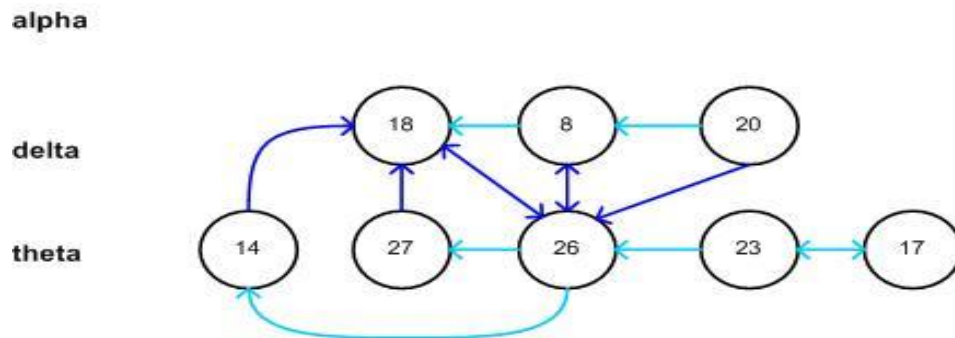


## AD Subject 2



**Fig. 5.11:** The maps of three intertrial coherence measures and their scalp topography. First row (cmp. 15), and second row (cmp. 12).

Two components were chosen for this subject as it did not seem to have a good enough alpha representative component. In the first row there is a delta band component with a posterior activation in its topography. The second row shows a component selected for its theta band activity, having a corresponding anterior topography.



**Fig. 5.12:** The PDC result graph which shows the relations between the chosen components.

In the graph of the second AD subject we can observe a close connectivity of theta components and a two-way interaction between the delta and theta components, with the theta band mostly driving the delta band. We do not have any nodes in the graph for alpha band as there no sufficient alpha activation for this patient.

## 5.2 Statistical Analysis of Components

Here is given a more detailed analysis of the results of the graphs. We try to take into consideration different parameters such as the number of components for each band, for each subject, the differences between normal and AD subjects and the different kinds of relations between components.

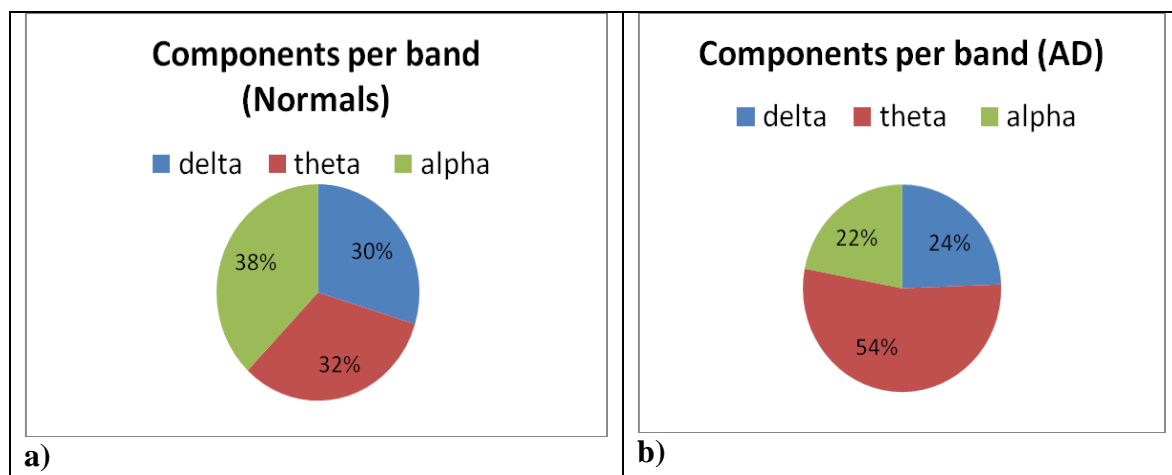
### 5.2.1 Components

For every subject used in the analysis we chose the most important components, each expressing (mostly) one of delta, theta or alpha frequency bands. The number of components is not the same for each subject because not all the subjects have the same activation in the different frequency bands. Also there are differences between the activations of an AD patient and a normal subject. The synoptic table which follows shows the number of components taken for each subject in each frequency band.

		Components		
Bands		delta	theta	alpha
Normal	Subject 1	3	2	3
	Subject 2	1	3	3
	Subject 3	2	3	1
	Subject 4	3	2	4
	Subject 5	2	2	3
AD	Subject 6	3	5	0
	Subject 7	1	7	1
	Subject 8	3	2	3
	Subject 9	1	3	5
	Subject 10	2	5	0
SUM	ALL	21	34	23
	CONTROL SUBJECTS	11	12	14
	AD	10	22	9

**Table 5.1**

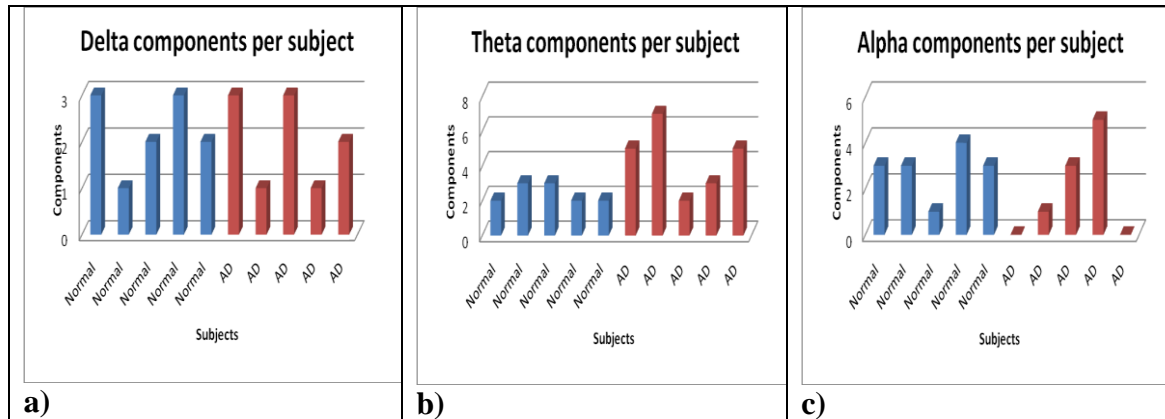
More specifically, in the following figures we have the percentage of delta, theta and alpha components for normal people and AD patients, respectively:



**Fig. 5.13: The percentage of components chosen for each frequency band for a) Control subjects and b) AD patients**

For the control subjects (subjects who haven't got AD) we have almost the same percentage of different frequency band components (30% delta components, 32% theta components, 38% alpha components). On the other hand, for the AD patients we have a great deal of theta band components (54% theta components).

In the following figures there is an analytic visualization of how many components we have for each subject (normal or AD) in each frequency band (delta, theta and alpha).



**Fig. 5.14: In these figures we have a visualization of the number of components per subject in a) delta band, b) theta band and c) alpha band.**

In the figures above we can observe that for delta band, for both normal and AD subjects we have from 1 to 3 components. In theta band, for control subjects we have 2 to 3 components and for AD patients we have 2 to 7 components. Here we observe that 3 out of 5 AD patients have more than 3 theta components which shows that they have a greater theta activation than the control subjects. In alpha band, for control subjects we have 1 to 4 components and for AD patients we have 0 to 5 components. It can be observed that for control subjects 4 out of 5 have more than 2 alpha components. In contrast only 2 out of 5 AD patients have more than 2 alpha components. Also 2 out of five AD patients have no alpha components.

## 5.2.2 Component relations analysis

A very important parameter for the analysis which has to be taken under consideration is the kind of relations between the components of the different or same frequency bands.

### 5.2.2.1 One way directed coherence

A synoptic table having the one-way directed coherence of the components follows. It shows in detail for every subject, the amount of one-way directed relations they have. In this case we have a “driver and response” relationship between components-observations.

		Number of components								
Relations		d->d	d->t	d->a	t->d	t->t	t->a	a->d	a->t	a->a
Control subjects	Subject 1	0	0	1	1	0	0	2	3	1
	Subject 2	0	0	0	1	3	0	2	1	1
	Subject 3	1	0	0	3	1	0	0	0	0
	Subject 4	0	0	1	2	0	0	5	1	1
	Subject 5	1	0	1	1	0	0	0	0	2
AD	Subject 6	2	1	0	2	3	0	0	0	0
	Subject 7	0	1	0	1	6	1	0	0	0
	Subject 8	2	1	0	2	1	0	0	0	2
	Subject 9	0	1	0	1	0	4	0	1	2
	Subject 10	0	0	0	3	3	0	0	0	0
SUM	ALL	6	4	3	17	17	5	9	6	9
	CONTROL SUBJECTS	2	0	3	8	4	0	9	5	5
	AD	4	4	0	9	13	5	0	1	4

Table 5.2

According to the table above, follows the figure which shows for the control subjects and the AD patients, the number of components, having each one of the one-way relations described such as  $d \rightarrow d$ ,  $d \rightarrow t$ ,  $d \rightarrow a$ , etc.

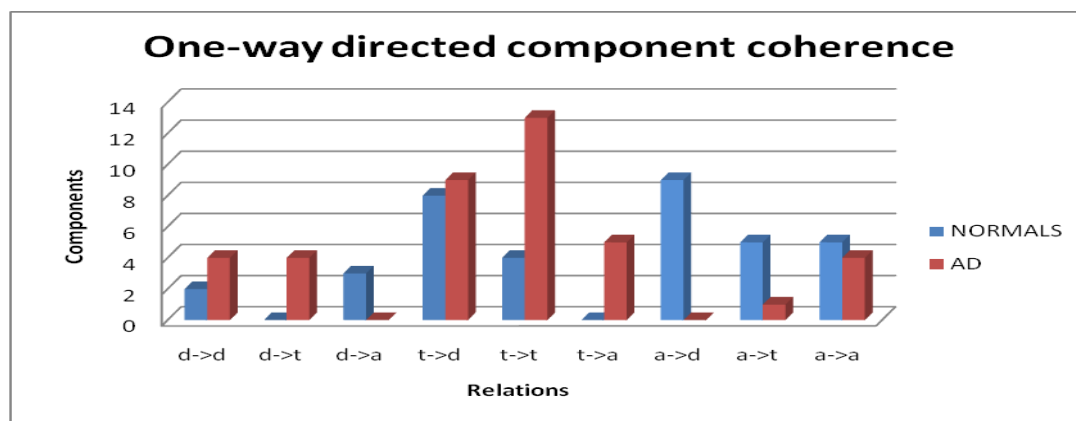


Fig. 5.15: Number of components having one-way directed coherence for each subject.



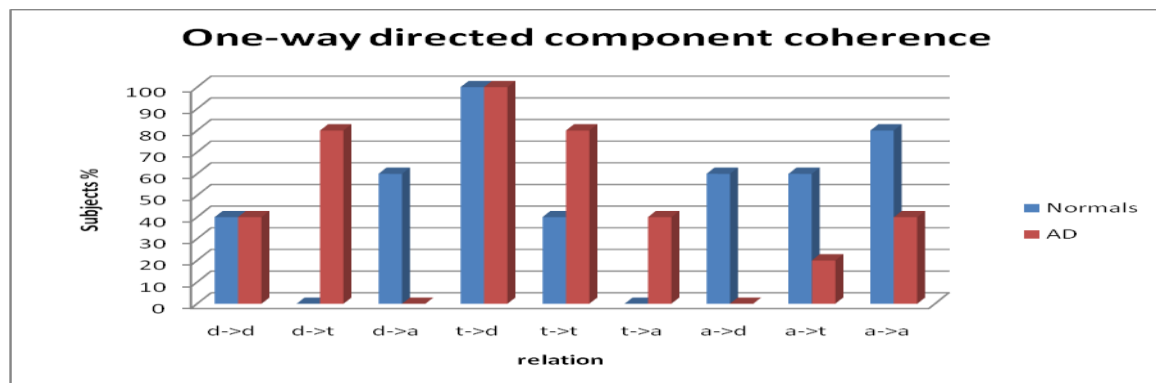
We can see by observing the figure and the table above that for the control subjects most of the one-way directed coherence relations are for theta leading delta components and for alpha leading delta components. There is also a considerable number of alpha components leading theta components and alpha components leading alpha components. For the AD patients we have a great deal of theta components leading delta components (just like the control subjects) as well as theta components leading theta components. There are also less cases of theta components leading alpha components, delta components leading theta components as well as alpha components leading alpha components.

For a more accurate analysis the table which follows shows what percent of the normal or AD people having at least one of the specific (previously described) directed coherence relations:

	Subjects %								
Relations	d->d	d->t	d->a	t->d	t->t	t->a	a->d	a->t	a->a
% subjects (all)	40	40	30	100	60	20	30	40	60
% Normal	40	0	60	100	40	0	60	60	80
% AD	40	80	0	100	80	40	0	20	40

**Table 5.3**

To have a better visualization result, the results of the table above are demonstrated in the following figure. It shows for the control subjects and the AD patients, the percentage of components, having at least one of the one-way relations described such as  $d \rightarrow d$ ,  $d \rightarrow t$ ,  $d \rightarrow a$ , etc.



**Fig. 5.16: The percentage of the subjects having at least one of the relations described.**

We can see that all of the subjects have at least one relation theta leading delta. More than 50% of the control subjects have relations such as delta leading alpha, alpha leading delta, alpha leading theta and alpha leading alpha. Also no normal subject has a relation of delta leading theta or theta leading alpha. On the other hand more than 50% of the AD patients have relations such as delta leading theta (no normal subject has this kind of relation), and theta leading theta. No AD patient is observed to have a relation of delta leading alpha or alpha leading delta something which for the most of normal people occurs.

### 5.2.2.2 Both-ways directed (partial-directed) coherence

A synoptic table having the both-ways directed coherence of the components follows. It shows in detail for every subject, the amount of both-ways directed relations they have. By both-ways directed relations we mean that we cannot say which is the “driver” and which is the “response” component, which means that the arrow shows both ways.

		Number of components					
Relations		d<->d	d<->t	d<->a	t<->t	t<->a	a<->a
Control subjects	Subject 1	0	1	2	0	0	1
	Subject 2	0	0	0	0	0	1
	Subject 3	0	1	0	0	1	0
	Subject 4	0	0	0	0	0	2
	Subject 5	0	1	2	0	1	0
AD	Subject 6	0	2	0	1	0	0
	Subject 7	0	0	0	1	0	0
	Subject 8	0	1	1	0	0	0
	Subject 9	0	0	0	1	2	1
	Subject 10	1	4	0	1	0	0
SUM	ALL	1	10	5	4	4	5
	CONTROL SUBJECTS	0	3	4	0	2	4
	AD	1	7	1	4	2	1

Table 5.4

According to the table above, follows the figure which shows for the control subjects and the AD patients, the number of components, having each one of the both-ways relations described such as  $d \leftrightarrow d$ ,  $d \leftrightarrow t$ ,  $d \leftrightarrow a$ , etc.

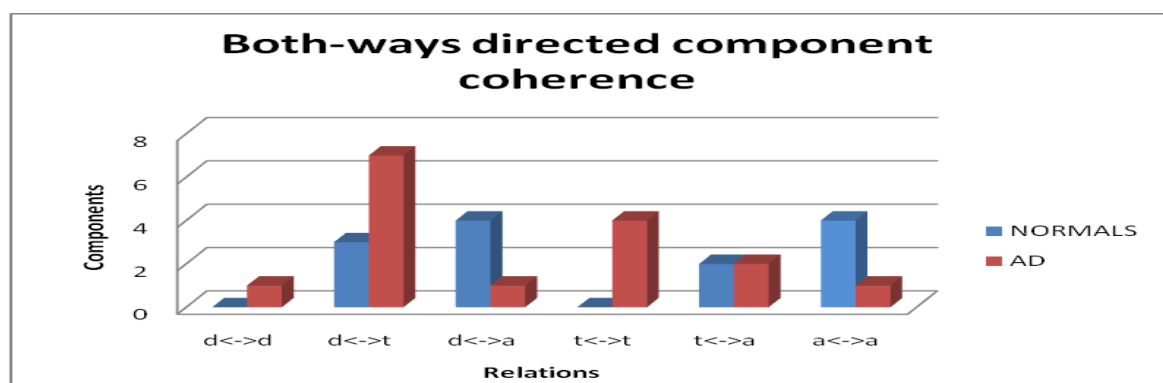


Fig. 5.17: Number of components having both-ways directed coherence for each subject.

We can see by observing the figure and the table above that for the control subjects most of the both-way directed coherence relations are between delta-alpha components and alpha components. No such relations are observed between delta components and theta components. For the AD patients we have a great deal of both-

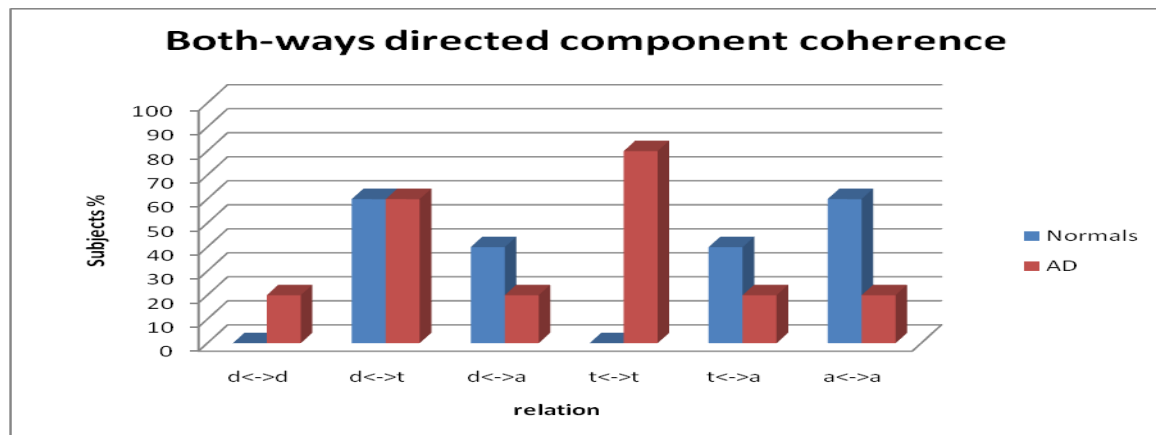
way directed coherence relations between delta-theta components as well as theta components.

For a more accurate analysis the table which follows shows what percent of the normal or AD people having at least one of the specific (previously described) both-ways directed coherence relations ( $d \leftrightarrow d$ ,  $d \leftrightarrow t$ ,  $d \leftrightarrow a$ ...):

	Subjects %					
Relations	$d \leftrightarrow d$	$d \leftrightarrow t$	$d \leftrightarrow a$	$t \leftrightarrow t$	$t \leftrightarrow a$	$a \leftrightarrow a$
% subjects (all)	10	60	30	40	30	40
% Normal	0	60	40	0	40	60
% AD	20	60	20	80	20	20

**Table 5.5**

To have a better visualization result, the results of the table above are demonstrated in the following figure. It shows for the control subjects and the AD patients, the percentage of components, having at least one of the both-ways relations described such as  $d \leftrightarrow d$ ,  $d \leftrightarrow t$ ,  $d \leftrightarrow a$ , etc.



**Fig. 5.18:** The percentage of the subjects having at least one of the relations described.

Observing the figure above, we can see that more than 50% of the normal and AD subjects have relations such between delta-theta. More than 50% of the control subjects also have both-ways directed coherence between alpha components and 40% have delta-alpha and theta-alpha relations. Also no normal subject has this kind of relation between delta components or between theta components. On the other hand more than 50% of the AD patients have relations between theta components. Only a few AD patients are observed to have the rest kinds of relations.

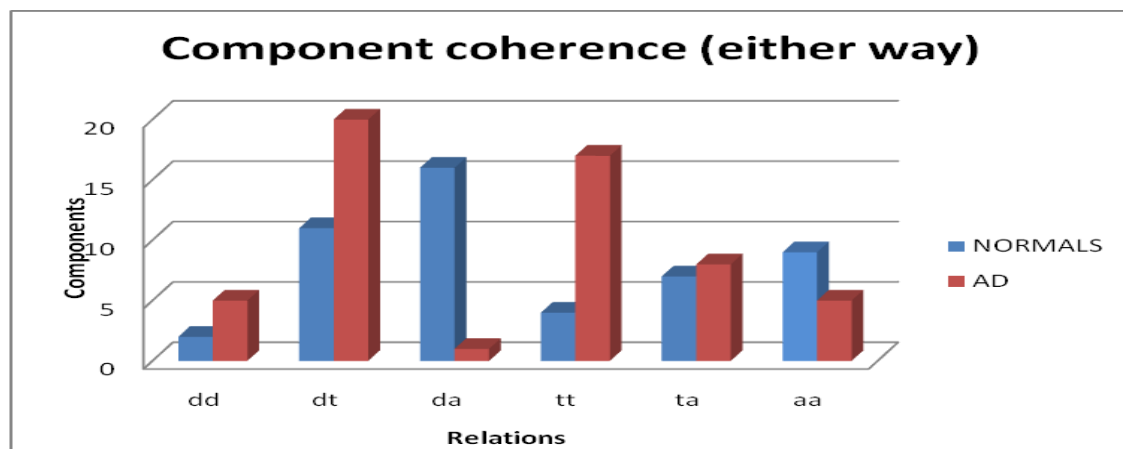
### 5.2.2.3 Either-way (one-way or both-ways) directed coherence

A synoptic table having the either-way directed coherence relations of the components follows. It shows in detail for every subject, the amount of either-way directed relations they have. In this case we do not care for the direction of the relation, this is why we name it either-way relation.

		Number of components					
Relations		dd	dt	da	tt	ta	aa
Control subjects	Subject 1	0	2	5	0	3	2
	Subject 2	0	1	2	3	1	2
	Subject 3	1	4	0	1	1	0
	Subject 4	0	2	6	0	1	3
	Subject 5	1	2	3	0	1	2
AD	Subject 6	2	5	0	4	0	0
	Subject 7	0	2	0	7	1	0
	Subject 8	2	4	1	1	0	2
	Subject 9	0	2	0	1	7	3
	Subject 10	1	7	0	4	0	0
SUM	ALL	7	31	17	21	15	14
	CONTROL SUBJECTS	2	11	16	4	7	9
	AD	5	20	1	17	8	5

**Table 5.6**

According to the table above, follows the figure which shows for the control subjects and the AD patients, the number of components, having each one of the relations described such as dd, dt, da, etc.



**Fig. 5.19: Number of components having either-way directed coherence for each subject.**

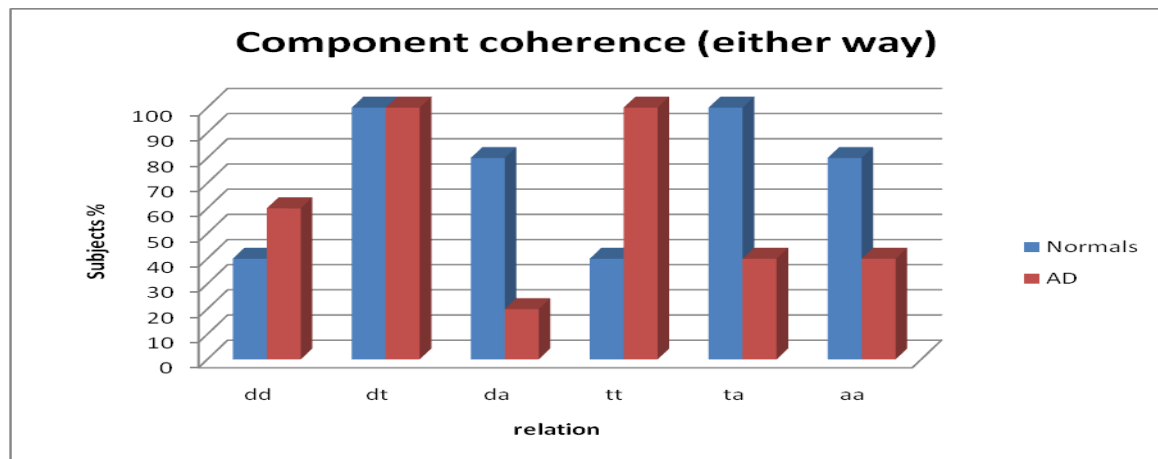
We can see by observing the figure and the table above that for the control subjects most of the relations are between delta-alpha components and delta-theta components. For the AD patients we also have a great deal delta-theta relations (just like the control subjects) as well as theta relationships.

For a more accurate analysis the table which follows shows what percent of the normal or AD people having at least one of the specific (previously described) either-way directed coherence relations (dd, dt, da...):

	Subjects %					
Relations	dd	dt	da	tt	ta	aa
% subjects (all)	50	100	50	70	70	60
% Normal	40	100	80	40	100	80
% AD	60	100	20	100	40	40

**Table 5.7**

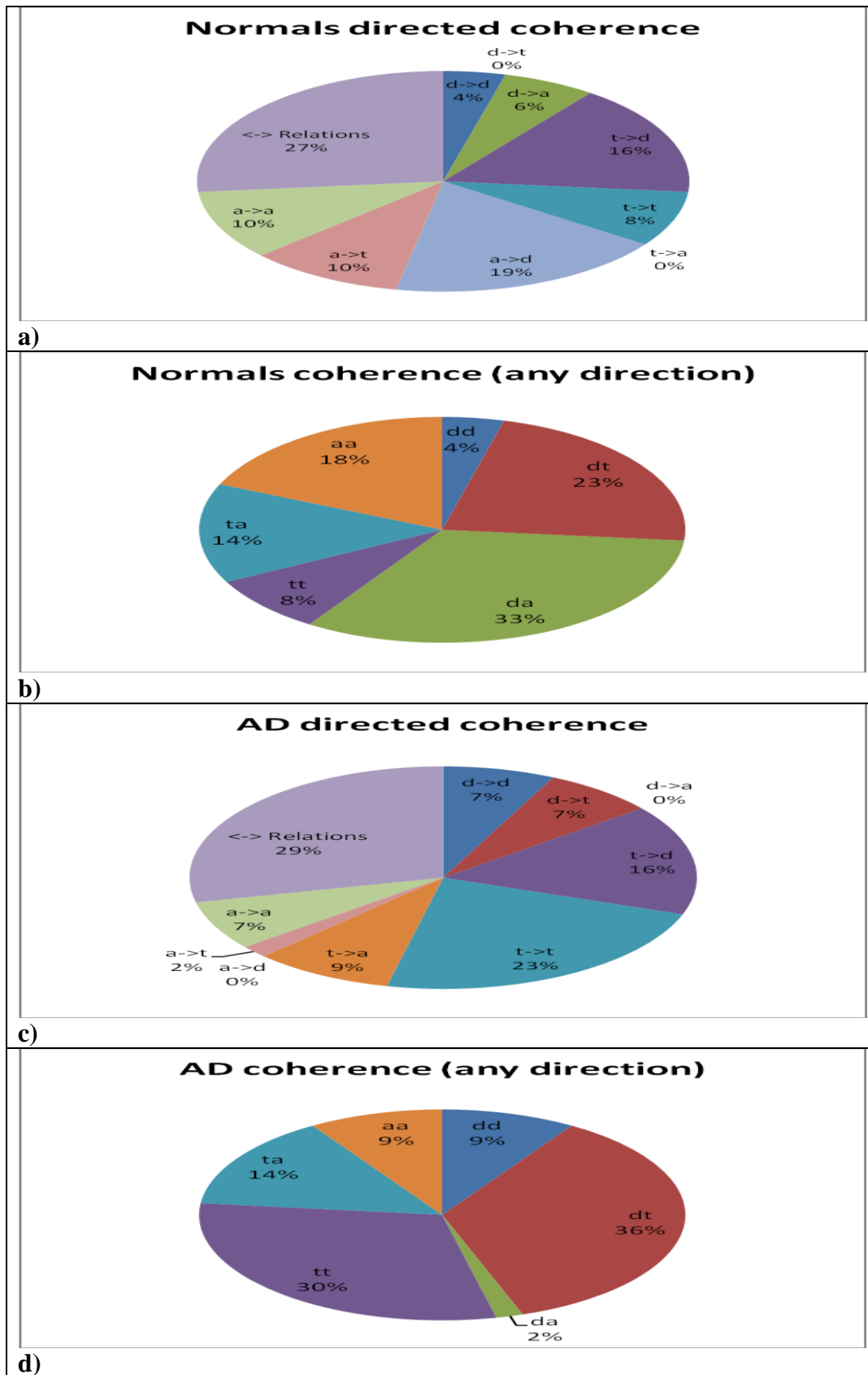
To have a better visualization result, the results of the table above are demonstrated in the following figure. It shows for the control subjects and the AD patients, the percentage of components, having at least one of the either-way relations described.



**Fig. 5.20: The percentage of the subjects having at least one of the relations described.**

We can see from the figure above that all of the subjects have at least one relation between delta-theta components. Also all the control subjects have at least one relation between theta-alpha components and all the AD patients have at least one relation between theta components. Also more than 50% of the control subjects have relations between delta-alpha components and alpha components. On the other hand more than 50% of the AD patients have relations between delta components.

According to the tables presented above, we can see what percentage we have for each relation compared to all the other relations. More specifically, we can see, for example, the percentage of  $d \rightarrow a$  relations compared to all the other relations such as  $t \rightarrow d$ ,  $t \rightarrow t$ , etc. This is something we do for every relation, respectively and the results are presented in figure 5.21.



**Fig. 5.21: Percentage of each relation compared to all the others.**

In figure 5.21 we have a visualization of the percentage of each relation between the same or different frequency bands compared to all the others. In figures 5.21a,c we have the percentage of relations (having direction) for control and AD subjects correspondingly and in figures 5.21b,d we have the percentage of relations (no direction information) for control and AD subjects correspondingly. By observing the figures, we can see that most of the relations for the control subjects are between delta-alpha (33%) and delta-theta (23%) components. Most of delta-alpha relations are alpha leading delta components and most of delta-theta relations are theta leading delta components. In AD patients' case, most of the relations are between delta-theta components (36%) and between theta components (30%). Most of delta-theta relations are theta leading delta (like in the control subjects' case) and most of the theta relations are one-way directed (one theta component leading another).

### 5.3 Discussion of Results

Event related oscillatory activity represents different aspects of information processing. Oscillations at delta band are related to focused attention, detection, recognition and decision making processes. Oscillations in theta band are correlated with attention, memory load, task difficulty and recognition of previous stimuli. Theta band oscillations also correlate with alpha-band power suppression or enhancement. Alpha-band oscillatory responses increase with simple memory tasks and decrease with demanding cognitive tasks.

During an oddball experiment, delta, theta and alpha bands are activated either in sequence or with overlap in time and are mapped in different brain locations reflecting the different stages of brain processing of the receiving stimuli. Delta and theta bands are activated in sequence and correlate with P300 ERP amplitude. Theta generally precedes delta in the P300 response and is more anterior in topography, while delta has rather posterior response [118],[119]. Theta activity is selectively enhanced during novelty stimulus presentations, linking it to the orienting processes associated with novelty processing. Delta response is related to target P300 amplitude and cognitive processing. In general an anterior theta response, related to orientation is followed by a posterior delta response which is related to cognitive processing.

Delta and theta responses present strong phase locking relative to the event, reflecting brain activity directly linked to event processing. With respect to the alpha-band response, however, recent reports have demonstrated that both phase-locked (evoked) and non-phase-locked (induced) alpha oscillations are functionally relevant to the oddball task processing and P300 response [112]. Furthermore, both fast and slow alpha have been related to specific aspects of cognitive processing indexed by P300 [30]. Several studies have also demonstrated that encoding typically induces alpha ERS activity, whereas retrieval elicits alpha ERD [106], [107]. Through a more detailed consideration of frequency bands in tasks of high mental load, a long lasting synchronization of the theta band is observed, along with a desynchronization of the upper alpha band (10-12 Hz) [112]. Similar results are reported in other studies, where theta and alpha synchronization is observed during the encoding of new items in short-term memory [21]. In addition, however, the response of the lower alpha band (8-10 Hz) in anterior sites is characterized by an ERS activation, which is attributed to a likely excess of attentional capacities during such mentally demanding tasks [112].

The evoked alpha activity is mostly identified at the lower range of the alpha band (slow alpha) and synchronizes during the first 100-200ms post-stimulus, inhibiting the semantic network, in order to facilitate the attention tasks. The evoked part of alpha band has a transient phase locking that can be observed in frontal (occipital) sites. It typically coincides with the exogenous ERP components and has been related with unspecific attention processes. In contrast, induced alpha activity has a damped character at higher frequencies of the alpha band, which is described by event related desynchronization (ERD) and is widespread in posterior (parietal) electrodes. The semantic network activates 400-500ms after the stimulus, when the maximum alpha desynchronization occurs. Alpha ERD reaches its maximum later than the P300 response and demonstrates increased cognitive processing during the oddball task [120]. Induced activity has been related to cognitive processing and episodic memory and increases with cognitive load and stimulus significance, while evoked alpha activity acts as an inhibitor of irrelevant processes in order to facilitate other, relevant, brain processes. The simultaneous existence of ERS and ERD in



distinct scalp areas is explained by accepting that ERD reflects functionally activated cortical regions, and ERS manifests temporary deactivation in other cortical fields. In agreement with these considerations, the model proposed by [J. Polich “Updating P300: An integrative theory of P3a and P3b,” *Clin. Neurophys.*, Vol. 118, pp. 2128–2148, 2007] for the generation of P300 waveforms comprises an early process (P3a) localized in frontal working memory related to attention and a late stimulus-related process (P3b) driven by attention that relates to memory processing. Even though the relation of P3a to alpha ERD has not been verified, it is clear that P3b related to the late alpha frequency of the EEG.

In summary, delta and theta bands present phase-locked activity contributing directly to the ERP waveform, while alpha band contains both evoked and induced aspects. In general, the frontal increase in phase-locked slow alpha activity to targets is accompanied by a parietal suppression of non-phase locked fast alpha activity. More specifically, desynchronization of upper alpha band is related to semantic processing, whereas synchronization of the lower alpha reflects inhibition. The topographic properties of this simultaneous appearance of evoked synchronization and induced desynchronization are highly dependent to the nature of stimulus and not in task in general. Despite the wide use of event related power measures, it is becoming widely accepted that both phase and non-phase locked oscillations are contributing to attention and mental tasks [112]. We should note that the traditional application of the ERD/ERS measure eliminates the phase-locked power and, as such, it does not take into account the evoked part of alpha activity. According to phase-related considerations, the early alpha synchronization in the lower alpha band is attributed to phase locked oscillations, whereas the later desynchronization in the higher alpha band is attributed to non-phase locked activity [112]. Our study forms a first attempt to address such phase locking issues related to event-related responses. Accordingly, the focus of frequency alterations is on the phase or non-phase locked nature of oscillations that can lead to synchronization and desynchronization of oscillations, rather than the ERD/ERS change in non-phase locked power. Even though we do not directly measure power changes, we consider phase synchronization that further contributes to ERS and phase desynchronization that is possibly reflected in ERD, so that a close agreement of these studies with our results can be established.

### 5.3.1 Control Subjects

It is observed that the PIC measure expressed the existing phase-locked activity differently than the average energy spectrum, indicating an influence of the latter on factors irrelevant to phase consistency throughout the trials. The synchronization maps of the subjects are organized by means of the directed influence of the bands Delta, Theta and Alpha. Besides the strong synchronization among bands, these figures elucidate the order of functional activity in terms of the driver system. The oddball ERP response is mainly characterized by a positive peak 300ms after the stimulus, also known as P300, appearing in the target stimuli. P300 response forms a processing sequence that varies in time, topography and frequency. Studies on the topography of P300 response reveal both earlier anterior and later posterior contributions to P300 responses. Target stimuli involve parietal activations that vary with the degree of cognitive engagement. Anterior contributions were found in unexpected and novel stimuli; in general, infrequent task irrelevant stimuli produce earlier strong anterior activations [85]. A number of investigators have also evaluated ERP activity to novel and target stimuli using time–frequency methods, and findings

suggest that the anterior to posterior processing sequence also varies in frequency. Specifically, the anterior activation is generally higher in frequency (e.g. theta) relative to the posterior activation (e.g. delta). Both theta and delta activities have been related to oddball target responses [103]. Delta activity (from 0 to 3Hz), has been most directly related to target P300 amplitude and the cognition as related to memory processing.

Theta activity (from 3 to 7Hz) has also been strongly implicated in oddball processing, related to attention, short-term memory and encoding of new information. Theta generally precedes delta in the P300 response, and is more anterior in topography, while delta is later and more posterior. Interestingly, theta is selectively enhanced during novelty stimulus presentations, linking it to the orienting processes associated with novelty processing [85]. Thus, the processing sequence involves an anterior theta response first, more closely tied to orienting, and then a posterior delta response more closely tied to cognitive processing. Sensory stimulation also induces phasic alpha ERD. In accordance to the neuronal inhibition hypothesis of stimulus processing, the P300 waveform may originate from a reduction and desynchronization of fast (alpha rhythm) non-phase locked oscillations, even though it is composed of phase locked delta and theta synchronized oscillations.

Alpha ERD is induced by task requiring cognitive processing with attention and memory components. Task induced alpha rhythms are associated with the above functional meaning, with an early slow (8-10Hz) alpha being associated to attention and late fast (10-12Hz) alpha linked to memory. The model proposed by [85] for the generation of P300 waveforms comprises an early process (P3a) localized in frontal working memory related to attention and a late stimulus-related process (P3b) driven by attention that relates to memory processing. Even though the relation of P3a to alpha ERD has not been verified, it is clear that P3b related to the late alpha frequency of the EEG.

The functional role of frequency components is also verified in this work. The independent components identified relate to the alpha, theta and delta bands. The delta components (related to cognitive processing) strongly relate to the alpha components. In both cases, we can also observe two types of alpha components, an early one at lower alpha band and a late one at the fast alpha range, with the early one driving the late alpha. The early alpha associated with attention drives a delta component related to cognition, which in turn drives the late alpha component that also related to memory operations; this relationship is marked with orange arrows. This tendency supports the clear relationship of P3b (mostly activated in delta band) to the late alpha activity [85]. The theta components (related to episodic short-term memory and attention) are also related to delta activity, with a tendency of theta to drive delta components. Instead, there is only weak relation between alpha and theta verifying the unclear relation of P3a (mostly associated to theta content) with alpha activity.

The dynamic synchronization of components, further support the functional relationships among band activities. Consistently, early alpha components lead some theta components and, in turn, theta lead delta components. There are also two-way interactions between early alpha and delta pairs, as well as theta and delta pairs of components. Furthermore, there is a consistent tendency of early alpha and delta components to lead a late alpha component. This pattern is in close agreement with the proposed involvement of early alpha and theta bands in alertness and attention, followed by the subsequent stimulation of delta band and late alpha activations for the completion of cognitive tasks.

### 5.3.2 AD Patients

In terms of pathological aspects, Alzheimer's disease reduces the ability to perform functions related to memory and complex attention, which also influences the activity at various bands of the EEG of AD subjects compared to controls. The P300 ERP paradigm is related directly to mental efficiency and reflects brain processes that demand attentional allocation and fast memory processing. Furthermore, the P300 response has been primarily located in the temporo-parietal cortex, which is the area most severely affected by Alzheimer's disease. The amplitude of delta and theta-band response increases during an oddball experiment, in response to basic information processing mechanisms of attention allocation and immediate memory. In Alzheimer's disease, however, memory and complex attention functions are highly affected, resulting in reduced delta and theta activity. Delta and theta phase-locking is also reduced in Alzheimer's patients compared to control [121], [122]. Furthermore, with the use of drug treatment, the activity of the theta band and its phase-locking increase and can become comparable to control subjects. This is a crisp indication that theta and delta responses originate from different brain processes.

With respect to the alpha band, there are relevant indications that it is weaker or completely absent in AD subjects. In essence, AD responses may attempt to initiate the alpha ERS process in order to inhibit other processes to the benefit of attention, but completely lose their dynamic coupling afterwards (absence of alpha activity), possibly not being able to compensate with the increased difficulty of the task at hand [105]. Thus, many studies have generally related AD with alpha-spectral changes of the EEG, which include a significant decrease in upper-alpha to beta power, followed by a decrease in the entire alpha band activity [108], [109], [110]. Significant group differences between controls and ADs have been reported in [105], observed in the 7-17 Hz frequencies and localized in the areas of frontal, central and left temporal electrodes. The alpha/beta band decrease in reactivity has also been shown through spectral [117]. Furthermore, the increase in relative theta power has been used to predict MCI patients that will progress to AD [116].

The experiment considered in this work focuses on the retrieval of events with the presentation of an acoustic probe. Processes related to working memory, especially during retrieval, have been found to weaken with normal aging, but in general they follow the same patterns of activation [104]. Nevertheless, the activation pattern is expected to change with the progression of AD, as described above. Indeed, in our experiments theta band in the AD cases is reflected in many more independent components than in controls. The AD patients considered were all under drug treatment, thus explaining the increased theta activity reported in all cases. Furthermore, the late (upper) alpha activity is nonexistent in ADs. Only the lower alpha activity is observed as an early activation in some AD subjects, but it appears as phase locked (synchronized) activity rather than non-phase locked oscillations contributing to ERD. This pattern of activation could be attributed to the increased effort of ADs for alertness and attention that relates more to the encoding than to the retrieval states [105]. Furthermore, the similar activation pattern shared by controls and ADs in the delta/theta bands (3-7 Hz) during retrieval is also reported in [Karrasch 2006, Hogan 2003], even though theta activity has been reported to increase in ADs during spontaneous EEG recording [108]. Our study reflects a more clear representation of the theta band in ADs, without however methodologically being able to indicate an increase in the total power of this band.

The lack of alpha-band activity during the retrieval phase in AD has been recently reported in [105] indicating that the pathology possibly affects alpha/beta desynchronization during short-term memory processing. Our results indicate a similar absence in ADs of the upper alpha band, which is mainly modulated by stimulus-induced effects and semantic memory processing. Nevertheless, the early activation of the lower alpha band, which relates to attention, indicates an ERS phase-locked pattern as opposed to the non phase-locked activation in controls. These findings further elucidate the changes in alpha-band activities, indicating that AD is related to a synchronized effort towards alertness, while it completely eliminates alpha activity related to working memory in retrieval. This performance is in close correspondence with the ERS increase of theta and lower alpha and the ERD increase in upper alpha bands during the experimental condition of demanding mental tasks, where the attentional capacities are exceeded and “alternative” cognitive strategies are utilized [112]. A similar interpretation of our results indicates that the attentional resources are exceeded in ADs, which is reflected by the alpha ERS due to phase locked early alpha amplitude increase [112]. Indeed, the alpha-related components in ADs reflected in the figures presented if exist, reflect only early alpha activity of mostly phase locked nature.

In the case of AD patients there exists an increased number of theta components, with less delta components. We can observe a close connectivity of theta components and a two-way interaction between the delta and theta components, with the theta band mostly driving the delta band. The alpha activity does not appear in all AD subjects and if it does it shows only early alpha (lower alpha) oscillations. The alpha band oscillations demonstrate a large phase locked portions, at least appearing in the power of the average trial signal, as opposed to the control subjects that typically demonstrate non-phase locked oscillations in alpha band. Furthermore, the functional synchronization of the alpha band with lower frequency bands is not clear. In some cases alpha is driven by delta, whereas in other cases alpha is driven by theta components.

## CHAPTER 6: Conclusion and Future Work

Our study was able to identify and characterize the intertrial coherence of independent components involved in the multiple trials of the auditory oddball experiment. Furthermore, it enabled the efficient visualization of established brain networks in three frequency bands, by means of the PDC synchrony measure. This study considered a population of five control and five AD subjects of which four representative activation networks were presented. Denoting the limitations of the study, which should be extended to a larger population sample, the initial results presented indicate that the proposed synchronization analysis framework is able to reflect not only the brain network topology during a certain mental task (like the working memory), but also the directional coupling between related brain regions.

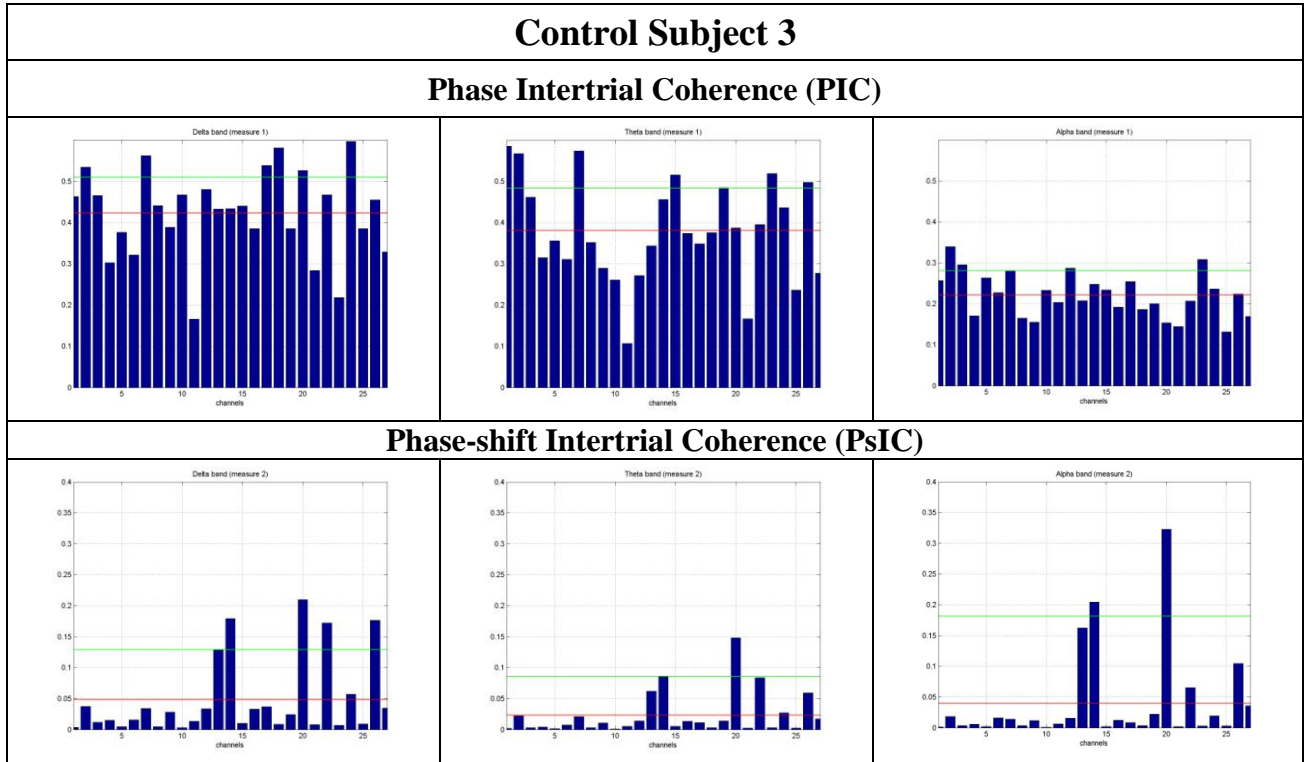
It has to be mentioned that this work can have even better results with further analysis. Some ideas which would help in the progress of this work are discussed here.

- **Larger population:** It is important for further statistical analysis to be accurate, that these measures and methods are applied on more subjects so as to make safer conclusions.
- **Stricter subject clustering:** We could use subjects for which we know more information of the specific age, sex etc. Especially in AD case it is important to have a common base because AD alters depending on different parameters and has different stages.
- **Better component filtering:** Some of the components we use as representatives for the different frequency bands have mixed activation. We could use filtering methods so as to clear the irrelevant activation.
- **TF-representation:** We could try different methods of TF transforms besides Morlet wavelets so as to compare and contrast the different representations.
- **PCA:** Another analysis which could alternatively be tried is PCA on EEG concatenated trials on the TF domain.

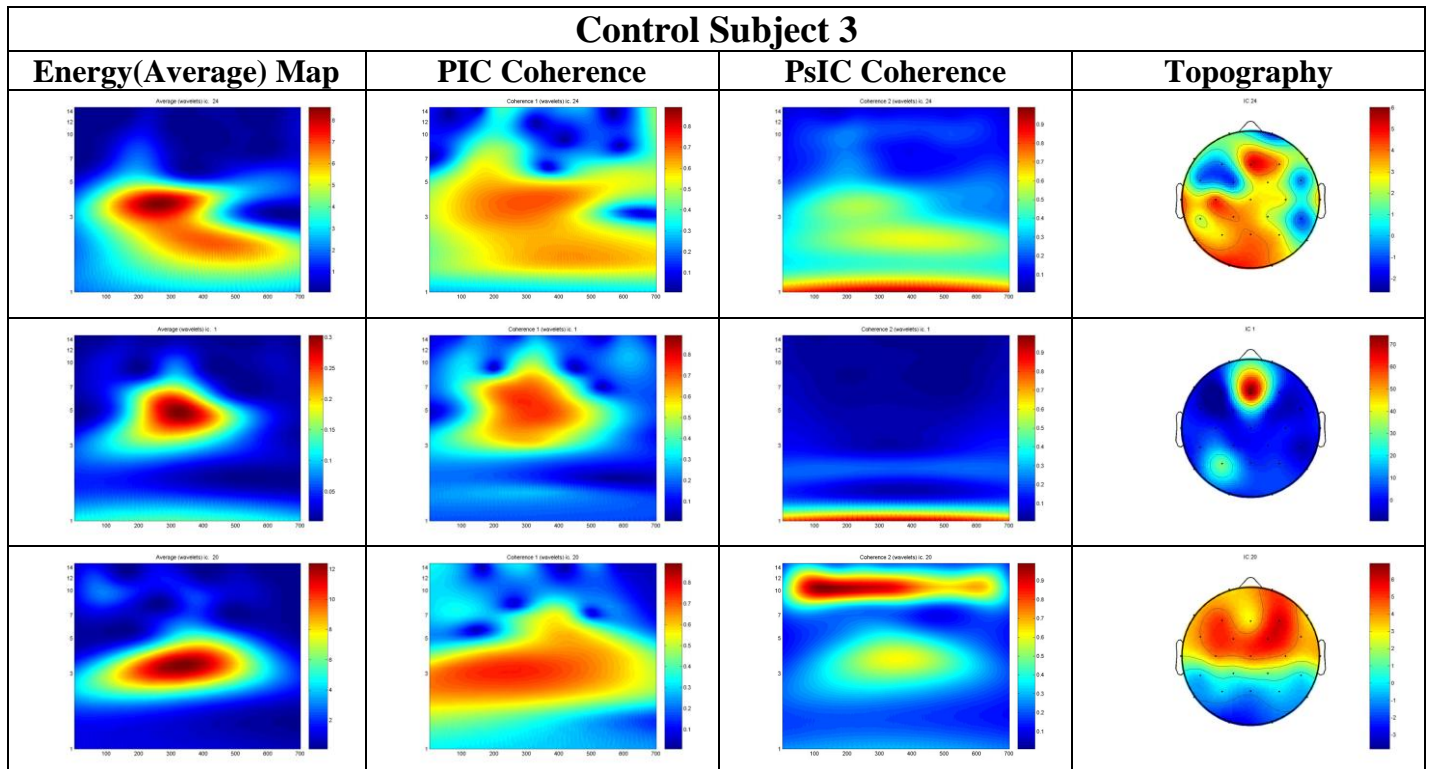
There are many ways in which our analysis could expand and progress, though it has to be mentioned that our method and the results which come up have novelties, as most of other researchers used to only examine the components' content without trying to find out the relations between the components.

# Appendix

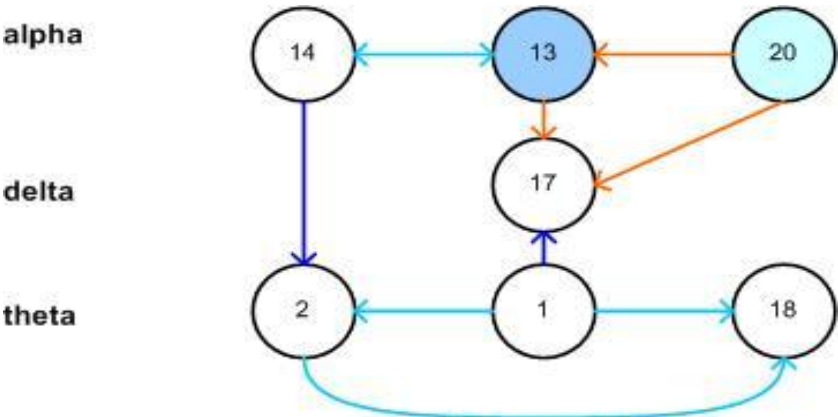
Here we will give some more examples on the subjects we did not include in the main analysis.



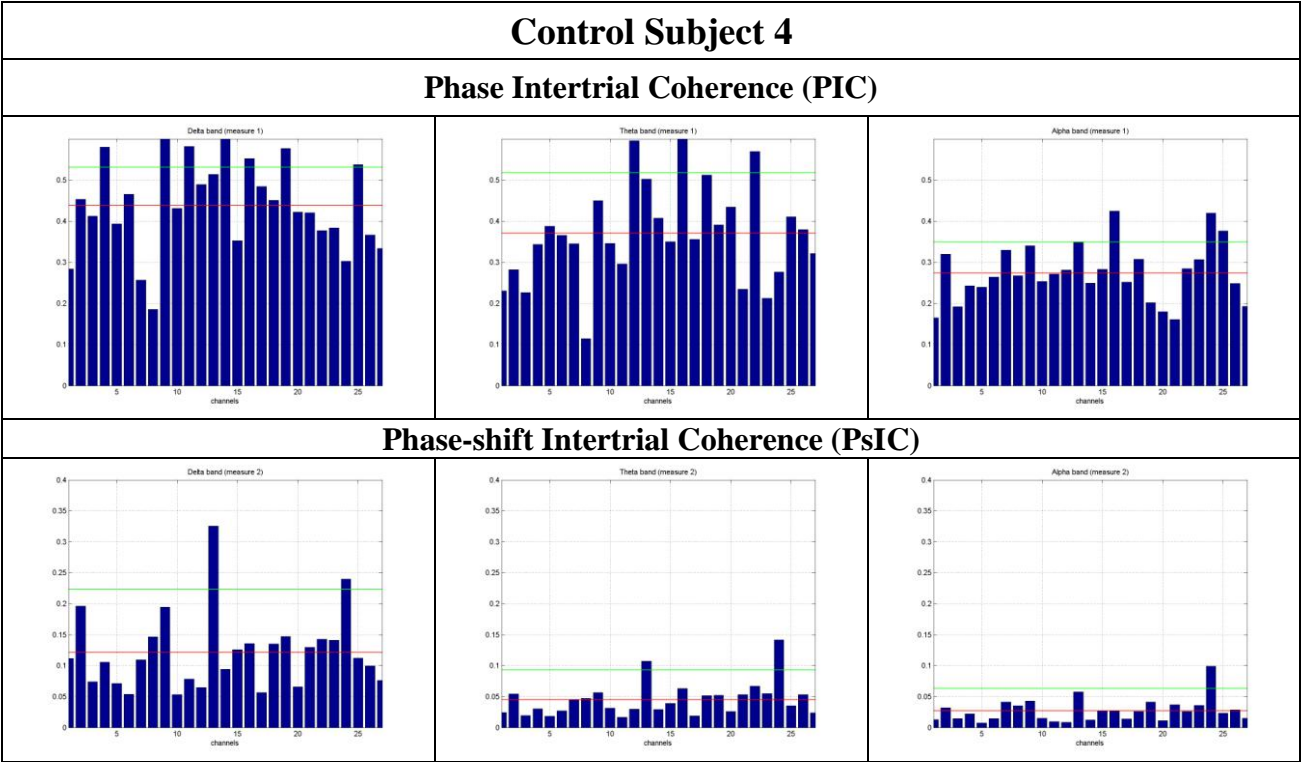
**Fig. 1: Components bar diagrams for delta, theta and alpha band.**



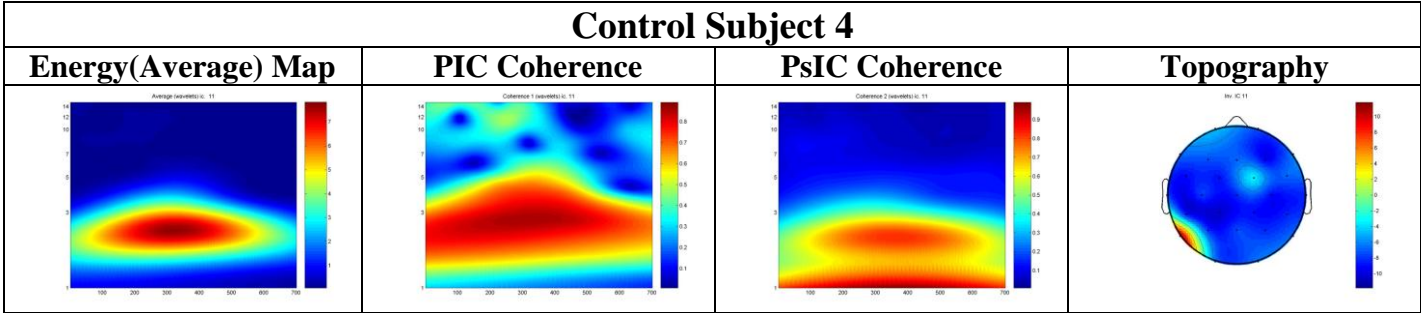
**Fig. 2: The maps of three intertrial coherence measures and their scalp topography. First row (cmp. 24), second row (cmp. 1) and last row (cmp. 20).**



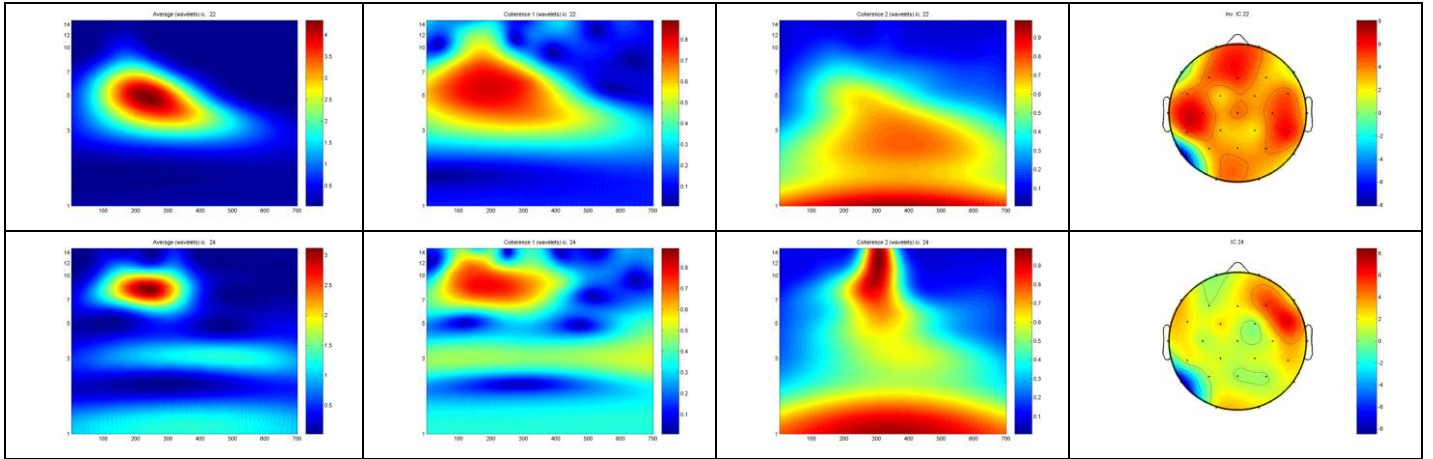
**Fig. 3: The PDC result graph which shows the relations between the chosen components.**



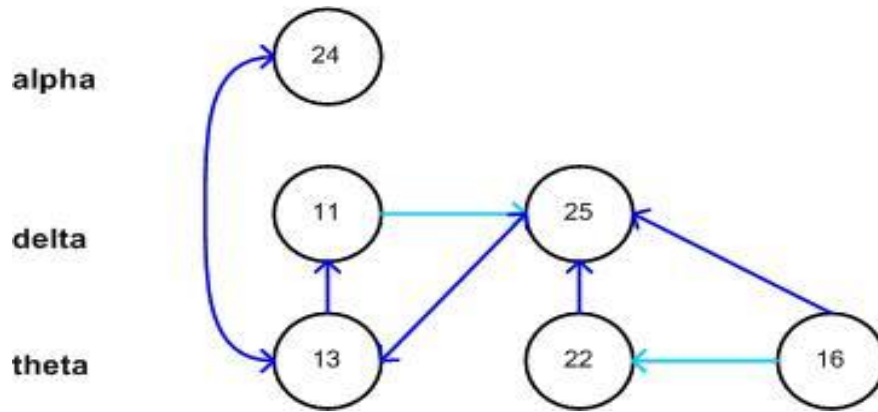
**Fig. 4: Components bar diagrams for delta, theta and alpha band.**



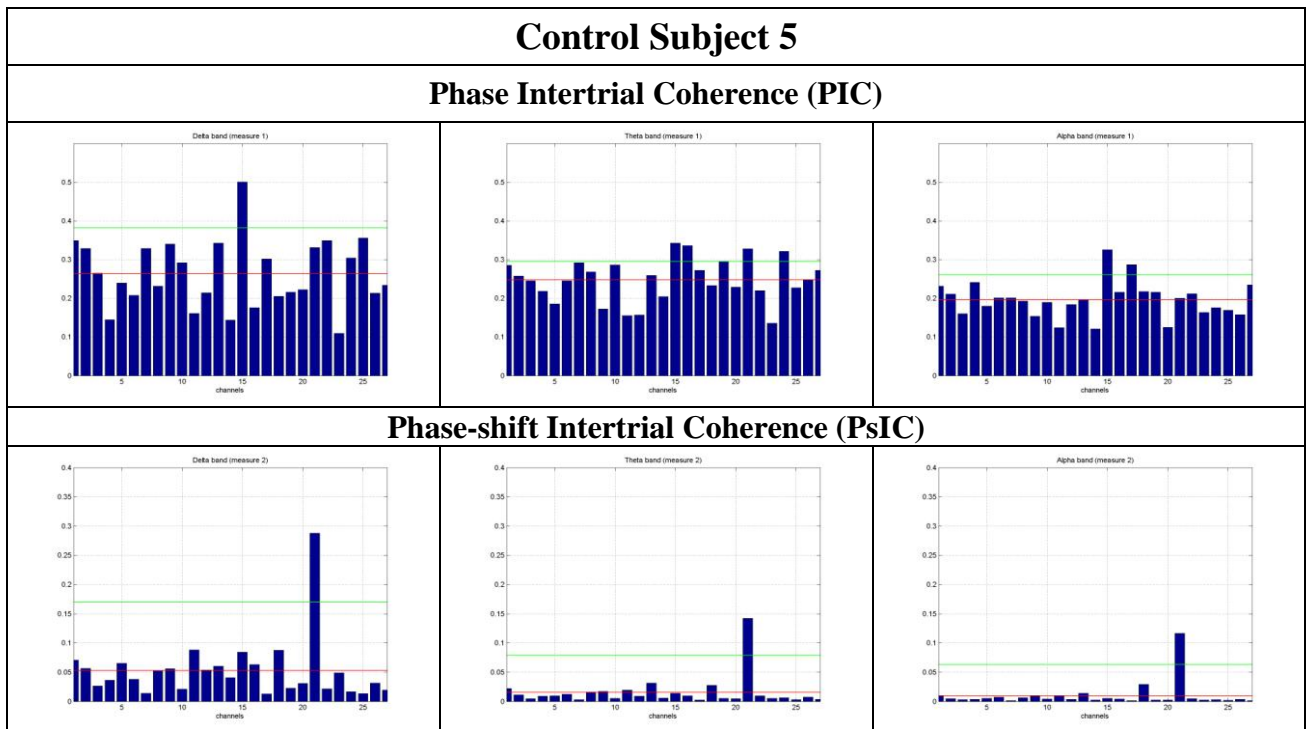




**Fig. 5:** The maps of three intertrial coherence measures and their scalp topography. First row (cmp. 11), second row (cmp. 22) and last row (cmp. 24).



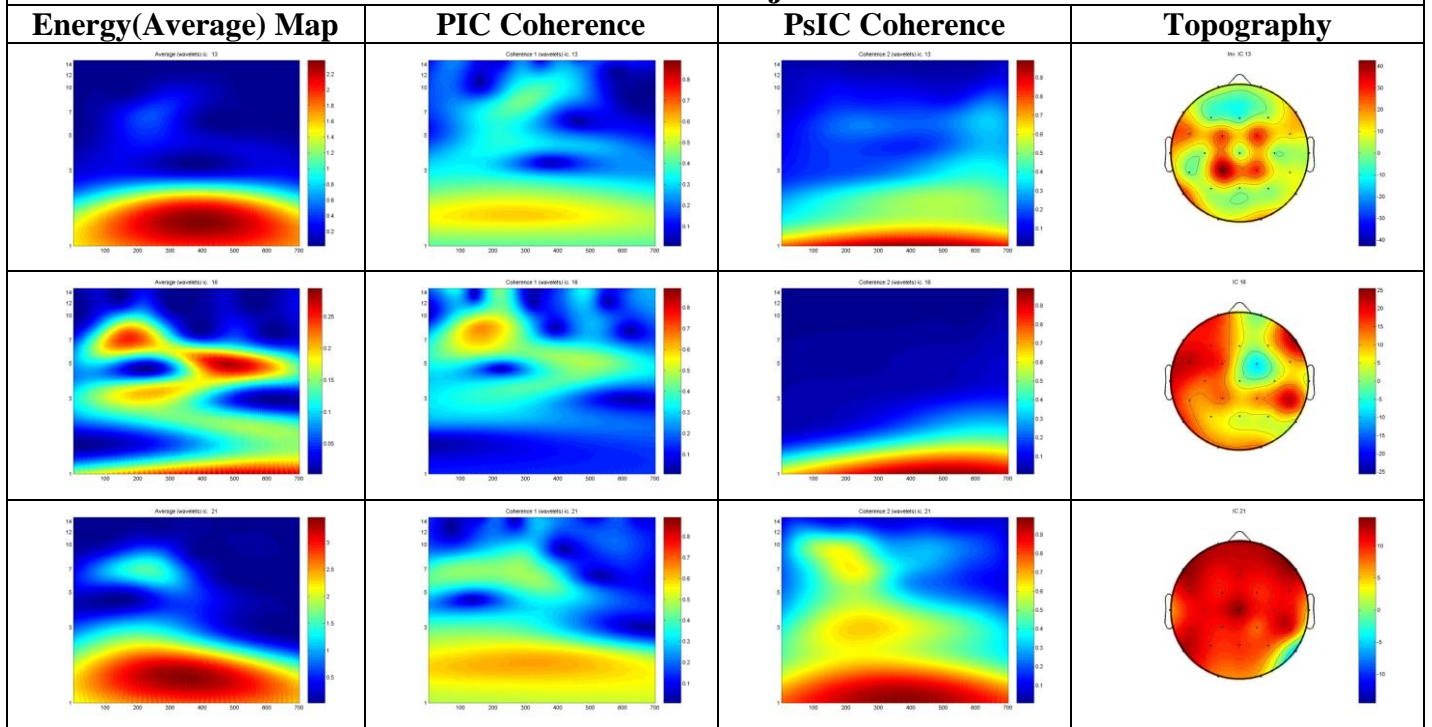
**Fig. 6:** The PDC result graph which shows the relations between the chosen components.



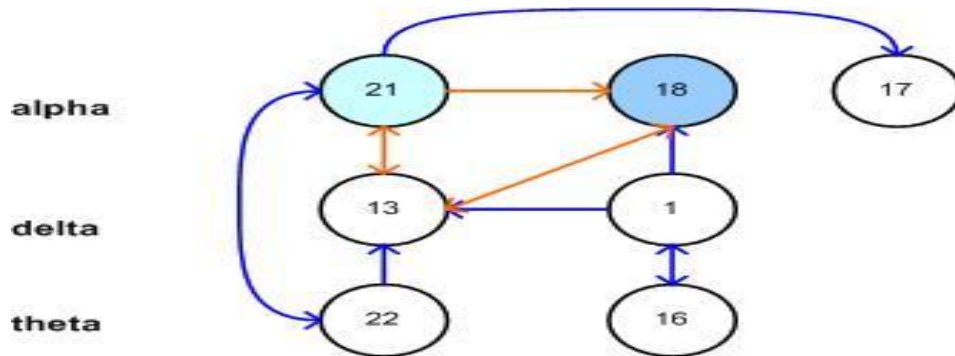
**Fig. 7:** Components bar diagrams for delta, theta and alpha band.



### Control Subject 4



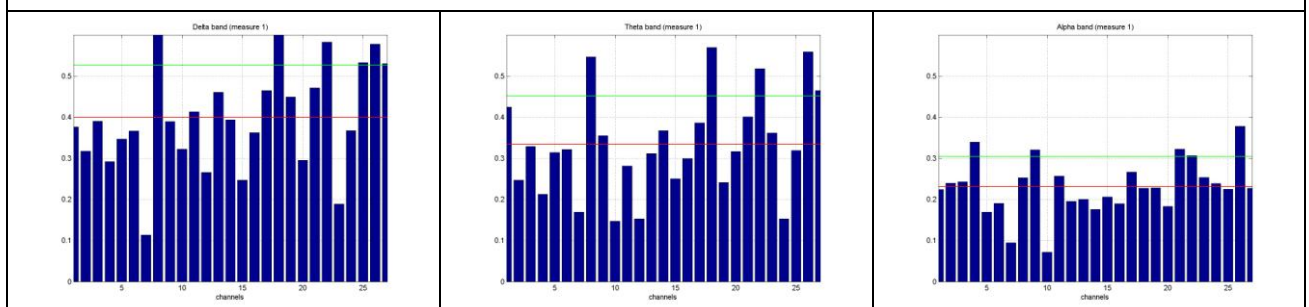
**Fig. 8:** The maps of three intertrial coherence measures and their scalp topography. First row (cmp. 13), second row (cmp. 16) and last row (cmp. 21).

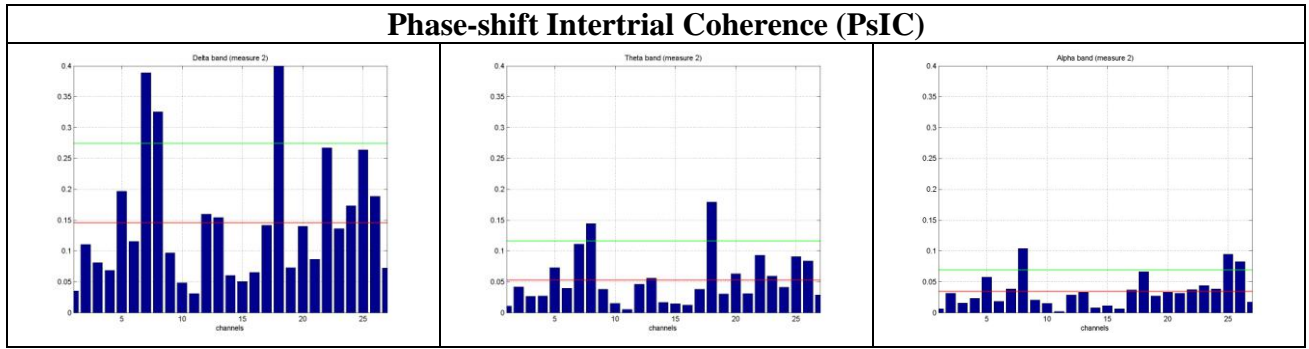


**Fig. 9:** The PDC result graph which shows the relations between the chosen components.

### AD Subject 3

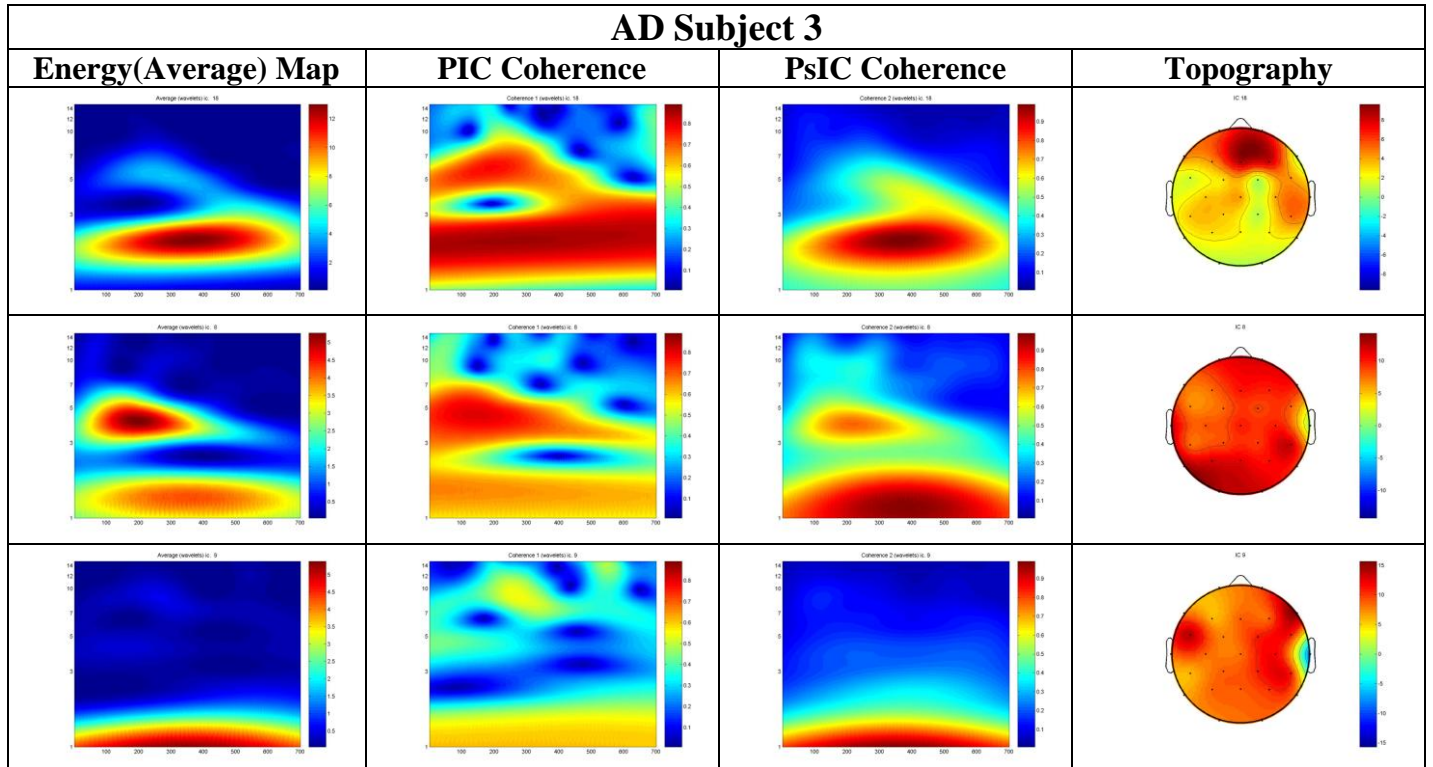
#### Phase Intertrial Coherence (PIC)



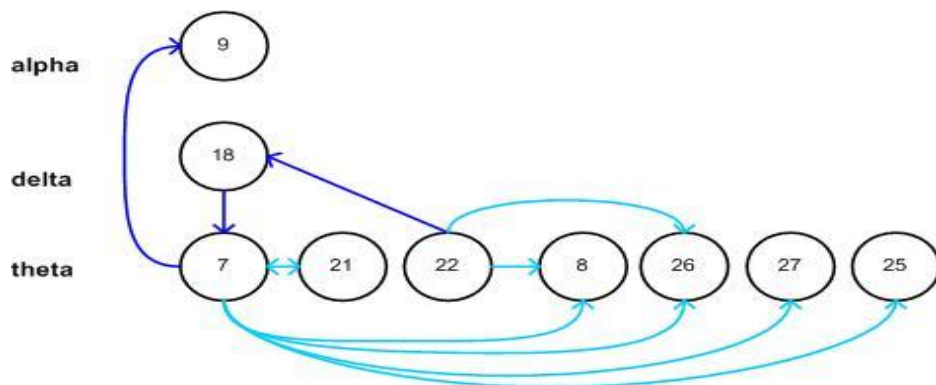


**Fig. 10: Components bar diagrams for delta, theta and alpha band.**

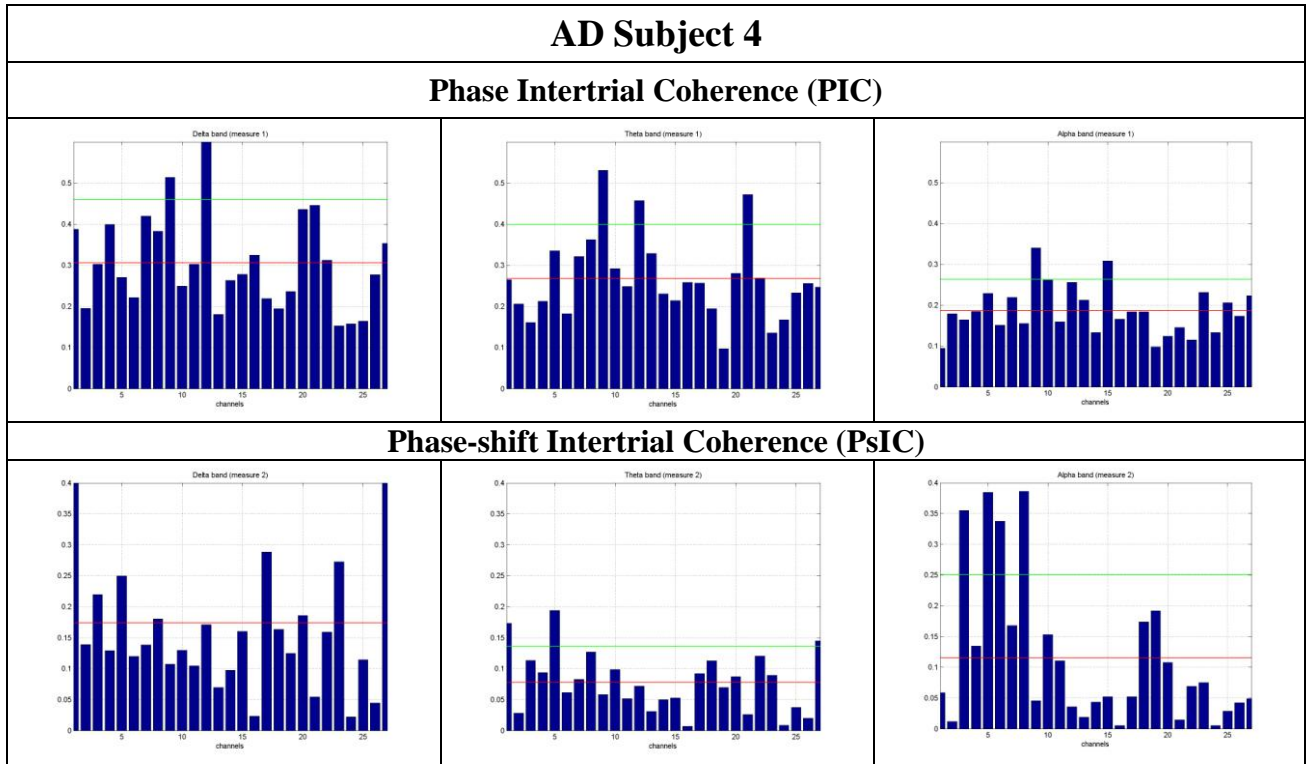
### AD Subject 3



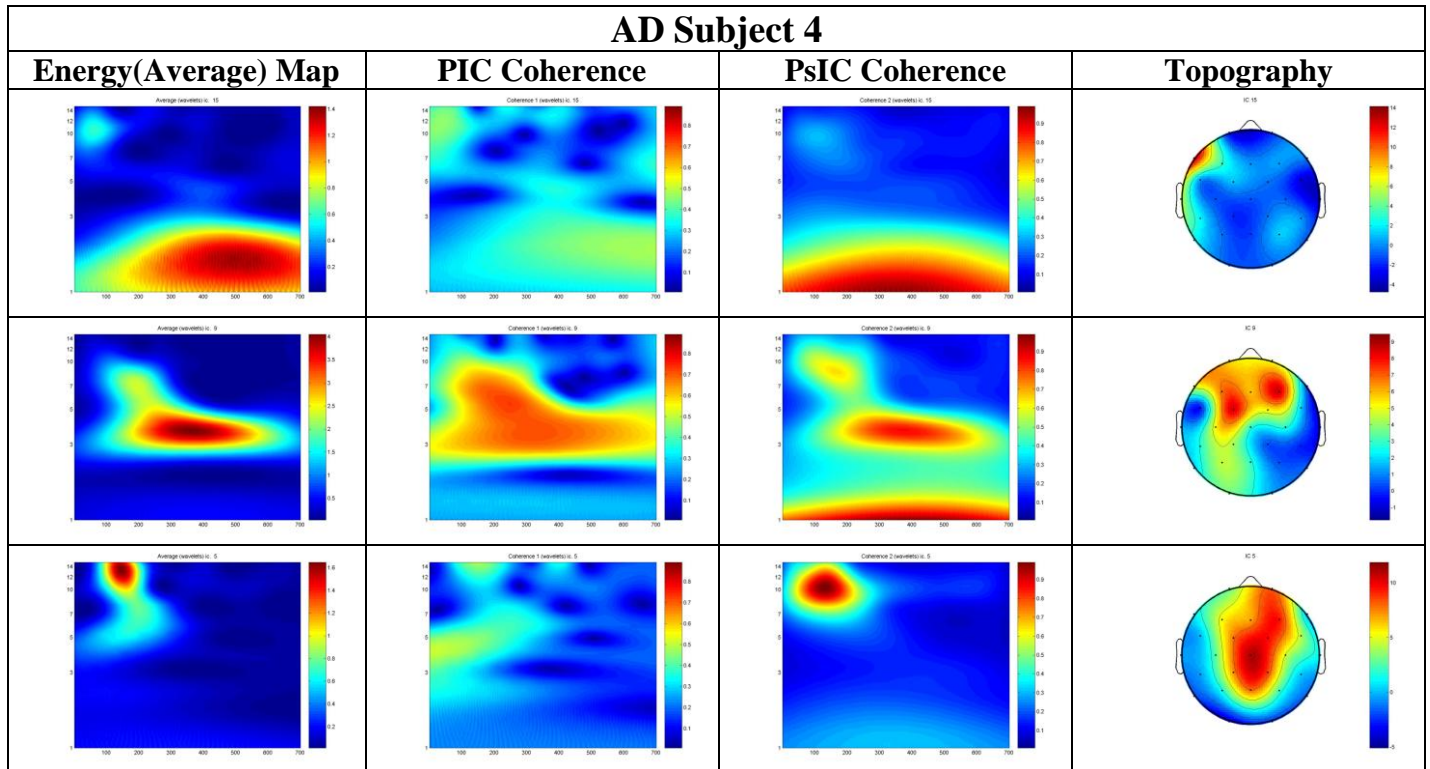
**Fig. 11: The maps of three intertrial coherence measures and their scalp topography. First row (cmp. 8) and second row (cmp. 17) .**



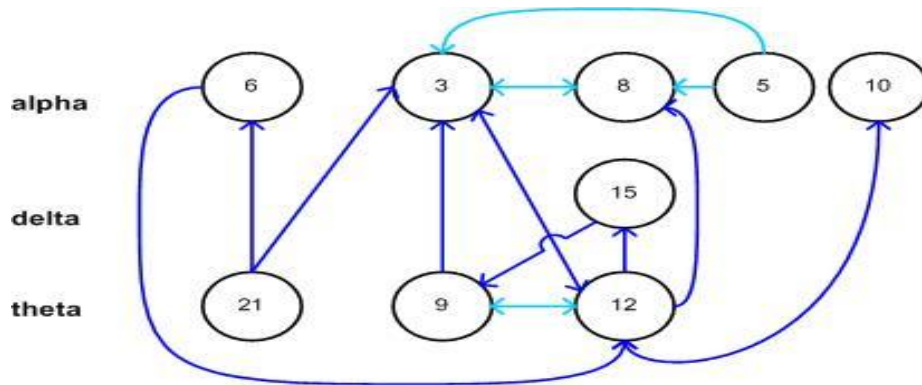
**Fig. 12: The PDC result graph which shows the relations between the chosen components.**



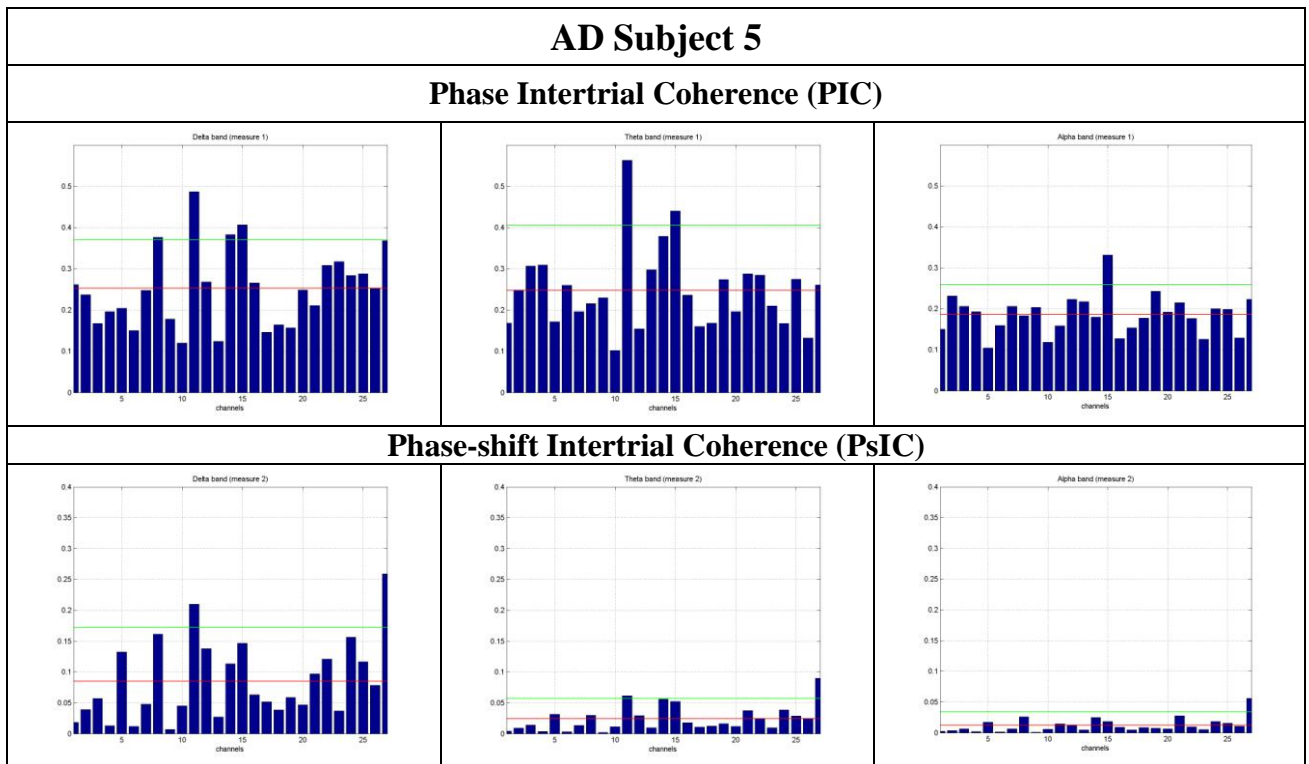
**Fig. 13: Components bar diagrams for delta, theta and alpha band.**



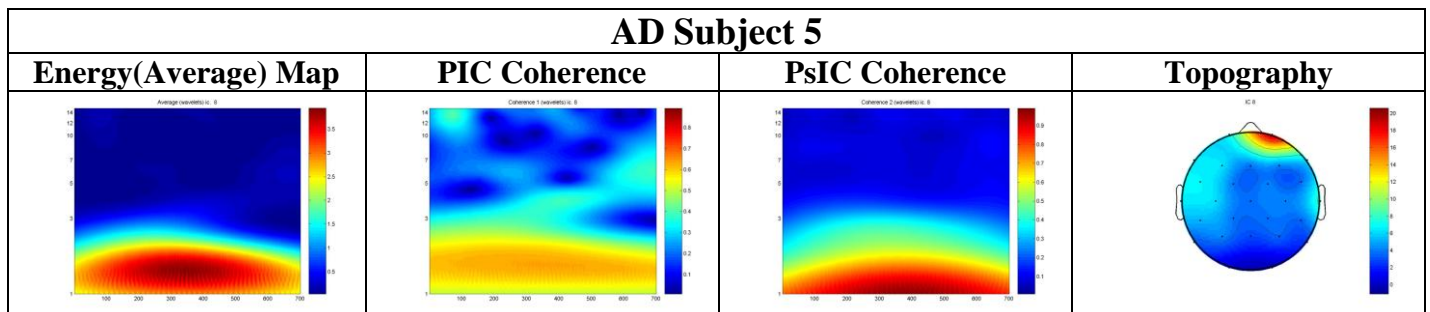
**Fig. 14: The maps of three intertrial coherence measures and their scalp topography. First row (cmp. 15), second row (cmp. 9), and last row (cmp. 21). .**



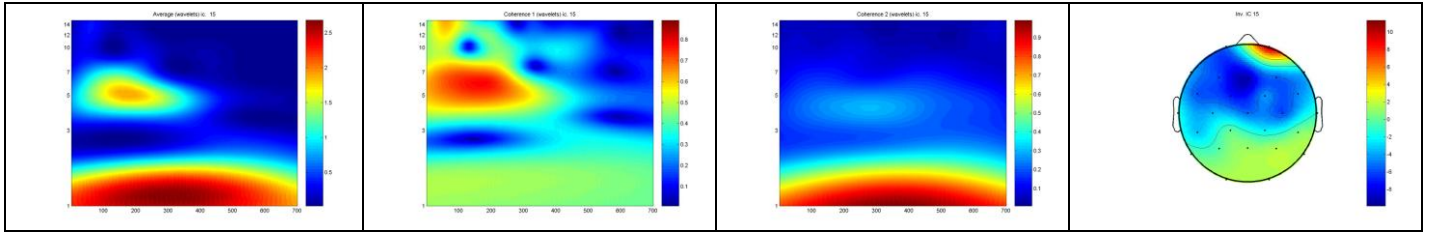
**Fig. 15:** The PDC result graph which shows the relations between the chosen components.



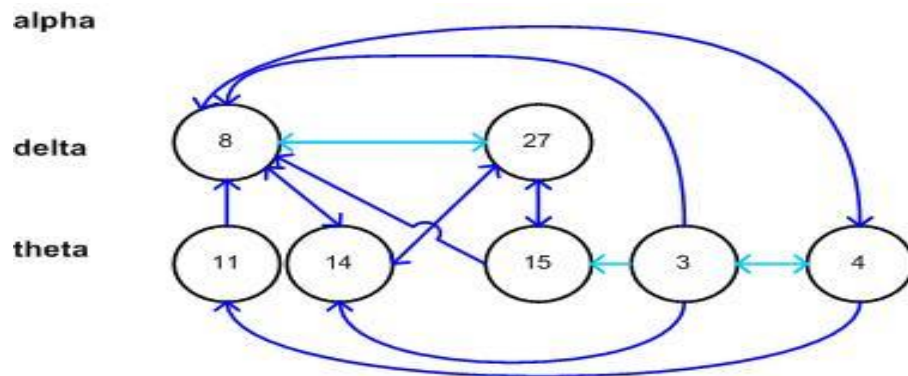
**Fig. 16:** Components bar diagrams for delta, theta and alpha band.







**Fig. 17: The maps of three intertrial coherence measures and their scalp topography. First row (cmp. 8) and second row (cmp. 15) .**



**Fig. 18: The PDC result graph which shows the relations between the chosen components.**

# References

- [1] Andrews, (2001), Neuropsychology, Psychology Press.
- [2] <http://www.brainanatomy.net/>
- [3] [http://en.wikipedia.org/wiki/Frontal\\_lobe](http://en.wikipedia.org/wiki/Frontal_lobe)
- [4] Kendra Van Wagner, The Anatomy of the Brain, About.com
- [5] Blakemore & Frith, (2005), The Learning Brain. Blackwell Publishing.
- [6] [http://en.wikipedia.org/wiki/Parietal\\_lobe](http://en.wikipedia.org/wiki/Parietal_lobe)
- [7] [http://en.wikipedia.org/wiki/Occipital\\_lobe](http://en.wikipedia.org/wiki/Occipital_lobe)
- [8] SparkNotes, Brain Anatomy: Parietal and Occipital Lobes.
- [9] [http://en.wikipedia.org/wiki/Temporal\\_lobe](http://en.wikipedia.org/wiki/Temporal_lobe)
- [10] Fine EJ, Ionita CC, Lohr L,(2002), The history of the development of the cerebellar examination, Semin Neurol 22 (4): 375–84.
- [11] Rapp, Brenda, (2001), The Handbook of Cognitive Neuropsychology: What Deficits Reveal about the Human Mind. Psychology Press. pp. 481.
- [12] <http://en.wikipedia.org/wiki/Pons>
- [13] Saunders Comprehensive Veterinary Dictionary, 3 ed. © 2007 Elsevier, Inc.
- [14] Swartz E. Barbara, Goldensohn Eli S. , (1998).,Timeline of the history of EEG and associated fields, Electroencephalography and clinical Neurophysiology 106, pp. 173–176
- [15] <http://www.ohsu.edu/health/health-topics>
- [16] <http://www.brainmaster.com>
- [17] Niedermeyer Ernst, Fernando Lopes da Silva, (2004), Electroencephalography: Basic Principles, Clinical Applications, and Related Fields - Page 140, Lippincott Williams & Wilkins.
- [18] <http://en.wikipedia.org/wiki/Eeg>
- [19] Benbadis R. Selim, Rielo Diego, (2008), EEG Artifacts, <http://emedicine.medscape.com/>
- [20] Krause C., Boman P., Sillanmäki L. , Varho T., Holopainen I., (2008), Brain oscillatory EEG event-related desynchronization (ERD) and -synchronization (ERS) responses during an auditory memory task are altered in children with epilepsy. Seizure 17, pp. 1-10.
- [21] Klimesch W, Dopplemayr M, Russegger H, Pachinger T, (1996), Theta band power in the human scalp EEG and the encoding of new information. NeuroReport 7: pp. 1235-1240.

- [22] Klimesch W,( 1999), EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis, *Brain Research Reviews* 29, pp. 169–195.
- [23] Krause C., Boman P., Sillanmäki L. , Varho T., Holopainen I., (2008), Brain oscillatory EEG event-related desynchronization (ERD) and -synchronization (ERS) responses during an auditory memory task are altered in children with epilepsy. *Seizure* 17, pp. 1-10.
- [24] Cahn BR, & Polich J. (2006), Meditation states and traits: EEG, ERP, and neuroimaging studies. *Psychological Bulletin*. 132 (2), pp. 180-211.
- [25] Klimesch W., Doppelmayr M., Pachinger T., Russegger H., (1997), Event-related desynchronization in the alpha band and the processing of semantic information. *Cognit. Brain Res.*, 6, pp. 83–94.
- [26] W. Klimesch, M. Doppelmayr, T. Pachinger and H. Russegger, (1997), Event-related desynchronization in the alpha band and the processing of semantic information. *Cognit. Brain Res.* **6**, pp. 83–94.
- [27] W. Klimesch, (1996), Memory processes, brain oscillations and EEG synchronization. *Int. J. Psychophysiol.* **24**, pp. 61–100.
- [28] W. Klimesch, M. Doppelmayr, H. Schimke and B. Ripper, (1997), Theta synchronization in a memory task. *Psychophysiology* **34**, pp. 169–176.
- [29] Klimesch W., Doppelmayr M., Russegger H., Pachinger T., Schwaiger J., (1998). Induced alpha band power changes in the human EEG and attention. *Neuroscience Letters* 244, pp. 73–76.
- [30] Yordanova,J. Kolev V., Polich J., (2001), P300 and alpha event-related desynchronization (ERD), *Psychophysiology*, pp. 143–152.
- [31] Basar, E., Hari, R., Lopes da Silva, F. H., & Schürmann, M. (1997), Brain alpha activity—New aspects and functional correlates, @Special issue 1–3#. *International Journal of Psychophysiology*, 26.
- [32] Krause, C., Lang, H., Laine, M., Helle, S., Kuusisto, M., & Pörn, B., (1994), Event-related desynchronization in the auditory stimulus modality., *Brain Topography*, 7, pp. 107–112.
- [33] Fitzgibbon S., Pope K., Mackenzie L., Clark C., Willoughby J., (2004), Cognitive tasks augment gamma EEG power., *Clinical Neurophysiology*, Vol. 115, pp. 1802 – 1809.
- [34] Willoughby J. O., Fitzgibbon S. P., Pope K. J., Mackenzie L., Medvedev A. V., Clark C. R., Davey M. P., Wilcox R. A., (2003), Persistent abnormality detected in the non-ictal electroencephalogram in primary generalised epilepsy. *Journal of Neurology Neurosurgery and Psychiatry*, pp. 51-55.
- [35] Brovelli A., Daprati E., Naranjo R.H., Muzur A., Longo R., Battaglini P.P., (2000), A method for the discrimination between induced and evoked gamma band activity. *Advances in Medical Signal and Information Processing*, 2000, pp. 226-230.

- [36] Tallon-Baudry C., Bertrand O., (1999), Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences* , Vol. 3, No. 4, pp. 151-162.
- [37] Muller M. M., Junghofer M., Elbert T., Rochstroh B., (1997), Visually induced gamma-band responses to coherent and incoherent motion: a replication study. *Neuroreport* 8, pp. 2575-2579.
- [38] Kaiser J., Lutzenberger W., (2003), Induced gamma-band activity and human brain function. *The Neuroscientist*, Vol. 9, No. 6, pp. 475-484.
- [39] Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J, (1996), Stimulus specificity of phase-locked and non-phase-locked 40 hz visual responses in human. *J Neurosci* 16: pp. 4240-4249.
- [40] Berchtold NC, Cotman CW, (1998), "Evolution in the conceptualization of dementia and Alzheimer's disease: Greco-Roman period to the 1960s". *Neurobiol. Aging* 19 (3): pp. 173–89.
- [41] Tiraboschi P, Hansen LA, Thal LJ, Corey-Bloom J., (2004), The importance of neuritic plaques and tangles to the development and evolution of AD. *Neurology* 62 (11): pp. 1984–9.
- [42] <http://www.alzheimers-research.org.uk/info/dementia/>
- [43] Alzheimer's association, Basics of Alzheimer's disease.
- [44] Anne Brown Rodgers, Alzheimer's disease: Unraveling the mystery, National Institute on Aging, National Institute on Health
- [45] [http://www.alz.org/alzheimers\\_disease\\_what\\_is\\_alzheimers.asp](http://www.alz.org/alzheimers_disease_what_is_alzheimers.asp)
- [46] <http://www.alzheimer.ca/english/disease/progression-intro.htm>
- [47] [http://www.alz.org/alzheimers\\_disease\\_stages\\_of\\_alzheimers.asp](http://www.alz.org/alzheimers_disease_stages_of_alzheimers.asp)
- [48] Carrie Hill, (2008), What causes Alzheimer disease?, About.com, reviewed by the Medical Review Board.
- [49] Carrie Hill, (2009), Diagnosis of Alzheimer Disease, About.com, reviewed by the Medical Review Board.
- [50] Aapo Hyvärinen, (1999), Survey on independent component analysis, *Neural Computing Surveys*, Vol. 2, pp. 94-128.
- [51] H. H. Harman. , (1967), *Modern Factor Analysis*. University of Chicago Press, 2nd edition.
- [52] M. Kendall., (1975), *Multivariate Analysis*. Charles Griffin&Co.
- [53] I.T. Jolliffe. (1986), *Principal Component Analysis*. Springer-Verlag.
- [54] P. Comon., (1994), Independent component analysis - a new concept? *Signal Processing*, 36: pp. 287-314.



- [55] C. Jutten and J. Herault. (1991), Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24: pp. 1-10.
- [56] D.-T. Pham, P. Garrat, and C. Jutten. (1992), Separation of a mixture of independent sources through a maximum likelihood approach. In *Proc. EUSIPCO*, pp. 771-774.
- [57] J.-F. Cardoso. (1997), Infomax and maximum likelihood for source separation. *IEEE Letters on Signal Processing*, 4: pp. 112-114.
- [58] A.J. Bell and T.J. Sejnowski. , (1995), An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7: pp. 1129-1159.
- [59] [http://en.wikipedia.org/wiki/Time-frequency\\_representation](http://en.wikipedia.org/wiki/Time-frequency_representation)
- [60] W. Christopher Lang and Kyle Forinash, (1998), Time-frequency analysis with the continuous wavelet transform, American Association of Physics Teachers.
- [61] Valens C., (1999), A really friendly guide to wavelets.
- [62] <http://en.wikipedia.org/wiki/Wavelets>
- [63] Graps A., (1995), An Introduction to wavelets, vol. 2, num. 2, published by the IEEE Computer Society.
- [64] [http://en.wikipedia.org/wiki/Evoked\\_potentials](http://en.wikipedia.org/wiki/Evoked_potentials)
- [65] David R. Stapells, (2004-2005), <http://www.audiospeech.ubc.ca/haplab/aep.htm>
- [66] Κουτσούρης Δ., Παυλόπουλος Σ., Πρέντζα Α., (2004), Εισαγωγή στην βιοιατρική τεχνολογία και ανάλυση ιατρικών σημάτων, Εκδόσεις Τζιολα.
- [67] Σακκαλής Β., Ηλεκτοεγκεφαλογράφημα και εφαρμογές. Πανεπιστημιακές σημειώσεις, Πολυτεχνείο Κρήτης.
- [68] Warnke A, Remschmidt H, Hennighausen K. (1994), Verbal information processing in dyslexia--data from a follow-up experiment of neuro-psychological aspects and EEG. *Acta Paedopsychiatr.*;56(3): pp. 203-8.
- [69] Pause BM, Sojka B, Krauel K, Ferstl R., (1996), The nature of the late positive complex within the olfactory event-related potential (OERP). *Psychophysiology*. 33(4): pp. 376-84.
- [70] Greffrath W, Baumgärtner U, Treede RD., (2007), Peripheral and central components of habituation of heat pain perception and evoked potentials in humans. *Pain*. 132(3): pp. 301-11.
- [71] Quant S, Maki BE, McIlroy WE., (2005), The association between later cortical potentials and later phases of postural reactions evoked by perturbations to upright stance. *Neurosci Lett*. 381(3): pp. 269-74.

- [72] Chan PY, Davenport PW., (2008), Respiratory-related evoked potential measures of respiratory sensory gating. *J Appl Physiol.* 105(4): pp. 1106-13.
- [73] Wang AL, Mouraux A, Liang M, Iannetti GD., (2008), enhancement of the N1 wave elicited by sensory stimuli presented at very short inter-stimulus intervals is a general feature across sensory systems. *PLoS ONE.* 3(12):e3929.
- [74] Spreng M., (1980), Influence of impulsive and fluctuating noise upon physiological excitations and short-time readaptation. *Scand Audiol Suppl.* (Suppl 12): pp. 299-306
- [75] Keidel WD, Spreng M., (1965), Neurophysiological evidence for the Stevens power function in man. *J Acoust Soc Am.* 38: pp. 191-5.
- [76] Davis H, Mast T, Yoshie N, Zerlin S., (1966), The slow response of the human cortex to auditory stimuli: recovery process. *Electroencephalogr Clin Neurophysiol.* Aug;21(2): pp. 105-13.
- [77] Butler RA., (1968), Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. *J Acoust Soc Am.* 44(4): pp. 945-50.
- [78] Pantev C, Hoke M, Lehnertz K, Lütkenhöner B, Anogianakis G, Wittkowski W., (1988), Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. *Electroencephalogr Clin Neurophysiol.* 69(2): pp. 160-70.
- [79] Nash AJ, Williams CS., (1982), Effects of preparatory set and task demands on auditory event-related potentials. *Biol Psychol.* 15(1-2): pp. 15-31.
- [80] Hillyard SA, Hink RF, Schwent VL, Picton TW., (1973), Electrical signs of selective attention in the human brain. *Science.* 182(108): pp. 177-80.
- [81] Schafer EW, Marcus MM., (1973), Self-stimulation alters human sensory brain responses. *Science.* 181(95): pp. 175-7.
- [82] Curio G, Neuloh G, Numminen J, Jousmäki V, Hari R., (2000), Speaking modifies voice-evoked activity in the human auditory cortex. *Hum Brain Mapp.* 9(4): pp. 183-91.
- [83] Sutton, S., Braren, M., Zubin, J., & John, E. R., (1965), Evoked-Potential Correlates of Stimulus Uncertainty. *Science*, 150(3700), pp. 1187-1188.
- [84] Polich J, Criado JR, (2006), Neuropsychology and neuropharmacology of P3a and P3b, *Intl J Psychophysiol* 60 (2): pp. 172-185.
- [85] Polich J, (2007), "Updating P300: An integrative theory of P3a and P3b" *Clin Neurophysiol.*
- [86] Hagoort, P., (2003), Interplay between Syntax and Semantics during Sentence Comprehension: ERP Effects of Combining Syntactic and Semantic Violations. *Journal of Cognitive Neuroscience* 15 (6): pp. 883-899.
- [87] Hagoort, P., (2003), How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *NeuroImage* 20: S20.

- [88] Zheng Ye, Yue-jia Luo, Angela D. Friederici, and Xiaolin Zhou, (2006). "Semantic and syntactic processing in Chinese sentence comprehension: Evidence from event-related potentials." *Brain Research* 1071, pp. 186-196.
- [89] Hagoort, Peter; Brown, Colin; Osterhout, Lee, (1999), "The neurocognition of syntactic processing". *The Neurocognition of Language*. New York: Oxford University Press.
- [90] B. Barlow and J.D. Mollon, (1982), *The Senses*, Cambridge University Press.
- [91] [http://en.wikibooks.org/wiki/Consciousness\\_Studies/Print\\_version](http://en.wikibooks.org/wiki/Consciousness_Studies/Print_version)
- [92] <http://brainlang.georgetown.edu/erplab.htm>
- [93] J. Kalcher, a and G. Pfurtscheller, (1995), Discrimination between phase-locked and non-phase-locked event-related EEG activity, *Electroencephalography and Clinical Neurophysiology*, Volume 94, Issue 5, pp. 381-384.
- [94] Rolf G. Winter; Aephraim M. Steinberg. *Coherence*. AccessScience@McGraw-Hill.
- [95] [http://en.wikipedia.org/wiki/Coherence\\_\(physics\)](http://en.wikipedia.org/wiki/Coherence_(physics))
- [96] L. Baccala, K. Sameshima, and D. Y. Takahashi, (2007), Generalized partial directed coherence, *Digital Signal Processing*, 15th Intern. Conf., pp. 163-166.
- [97] Yasumasa Takahashi D., Baccala L. A., Sameshima K., (2007), Connectivity Inference between Neural Structures via Partial Directed Coherence, *Journal of Applied Statistics*, Vol. 34, No. 10, pp. 1255–1269.
- [98] Makeig, M. Westerfield, T-P. Jung, S. Enghoff, J. Townsend, E. Courchesne, T.J. Sejnowski, , (2002), "Dynamic Brain Sources of Visual Evoked Responses," *Science*, Vol. 295, pp. 690-694.
- [99] T.P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, T.J. Sejnowski, , (2001), "Analysis and Visualization of Single-Trial Event-Related Potentials," *Human Brain Mapping*, Vol. 14, pp. 166-185.
- [100] T.P. Jung, S. Makeig, M.J. McKeown, A.J. Bell, T.W. Lee, T.J. Sejnowski, , (2001), "Imaging Brain Dynamics Using Independent Component Analysis," *Proc. of the IEEE*, Vol. 89, No. 7, pp. 1107-1122.
- [101] V. Wyart and C. Tallon-Baudry, (2008), "Neural Dissociation between Visual Awareness and Spatial Attention," *The Journal of Neuroscience*, Vol. 28, No. 10, pp. 2667-2679.
- [102] E.M. Bernat, S.M. Malone, W.J. Williams, C.J. Patrick, W.G. Iacono, , (2007), "Decomposing delta, theta, and alpha time-frequency ERP activity from a visual oddball task using PCA," *Psychophysiology*, pp. 62-74.
- [103] C. Başar-Eroglu, T. Demiralp, M Schürmann, E Başar, (2001), "Topological distribution of odd-ball "P-300" responses," *Int. Journal of Psychophysiology*, Vol. 39, pp. 213-220.

- [104] Karrasch, M., Laine, M., Rapinoja, P., Krause, C.M., (2004), Effects of normal aging on event-related desynchronization/synchronization during a memory task. *Neuroscience Letters* 366, pp. 18–23.
- [105] M. Karrasch, M. Laine, J. O. Rinne, P. Rapinoja, E. Sinerva, C. M. Krause, (2006), “Brain oscillatory responses to an auditory-verbal working memory task in mild cognitive impairment and Alzheimer’s disease”, *International Journal of Psychophysiology* 59, pp. 168 – 178.
- [106] Karrasch, M., Laine, M., Rapinoja, P., Krause, C.M., (2004), Effects of normal aging on event-related desynchronization/synchronization during a memory task. *Neuroscience Letters* 366, pp. 18–23.
- [107] Krause, C.M., Salminen, P.-A., Sillanmaki, L., Holopainen, I.E., (2001), Event-related desynchronization and synchronization during a memory task in children. *Clinical Neurophysiology* 112, pp. 2233– 2240.
- [108] Bennys, K., Rondouin, G., Vergnes, C., Touchon, J., (2001), Diagnostic value of quantitative EEG in Alzheimer’s disease. *Neurophysiologie Clinique* 31, pp. 153–160.
- [109] Jeong, J., (2004), EEG dynamics in patients with Alzheimer’s disease. *Clinical Neurophysiology* 115, pp. 1490– 1505.
- [110] Kowalski, J.W., Gawel, M., Pfeffer, A., Barcikowska, M., (2001), The diagnostic value of EEG in Alzheimer disease. Correlation with the severity of mental impairment. *Journal of Clinical Neurophysiology* 18, pp. 570– 575.
- [111] Hogan, M.J., Swanwick, G.R.J., Kaiser, J., Rowan, M., Lawlor, B., (2003), Memory-related EEG power and coherence reductions in mild Alzheimer’s disease. *International Journal of Psychophysiology* 49, pp. 147– 163.
- [112] Krause, C.M., Sillanmaki, L., Koivisto, M., Saarela, C., Haggqvist, A., Laine, M., Hamalainen, H., (2000), The effects of memory load on event-related EEG desynchronization and synchronization. *Clinical Neurophysiology* 111, pp. 2071– 2078.
- [113] Klimesch W, Doppelmayr M, RoÈhm D, PoÈllhuber D, Stadler W., (2000), Simultaneous desynchronization and synchronization of different alpha responses in the human electroencephalograph: a neglected paradox? *Neurosci Lett* ;284:97±100.
- [114] Krause, C.M., Lang, A.H., Laine, M., Kuusisto, M., Po”rn, B., (1996), Event related EEG desynchronization and synchronization during an auditory M. Karrasch et al. / *International Journal of Psychophysiology* 59 (2006) 168–178 177 memory task. *Electroencephalography and Clinical Neurophysiology* 98, pp. 319–326.
- [115] Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., (1996). Theta band power in the human scalp EEG and the encoding of new information. *NeuroReport* 7, pp. 1235– 1240.

- [116] Jelic, V., Johansson, S.-E., Almkvist, O., Shigeta, M., Julin, P., Nordberg, A., et al., (2000), Quantitative electroencephalography in mild cognitive impairment: longitudinal changes and possible prediction of Alzheimer's disease. *Neurobiology of Aging* 21, pp. 533–540.
- [117] Berendse, H.W., Verbunt, J.P.A., Scheltens, P., van Dijk, B.W., Jonkman, E.J., (2000), Magnetoencephalographic analysis of cortical activity in Alzheimer's disease: a pilot study. *Clinical Neurophysiology* 111, pp. 604–612.
- [118] Demiralp T, Ademoglu A, Comerchero M, Polich J., (2001), Wavelet analysis of P3a and P3b. *Brain Topogr* 13, pp. 251-67.
- [119] Porjesz B, Rangaswamy M, Kamarajan C, Jones KA, Padmanabhapillai A, Begleiter H., (2005), The utility of neurophysiological markers in the study of alcoholism. *Clin Neurophysiol*, 116, pp. 993-1018.
- [120] Yordanova, J., Kolev, V. (1998). Event-related alpha oscillations are functionally associated with P300 during information processing. *NeuroReport*, 9: pp. 3159-3164. Copyright © 1998 Lippincott Williams & Wilkins.
- [121] G.G. Yener, B. Güntekin, A. Öñiz, E. Başar., (2007), Increased frontal phase-locking of event-related theta oscillations in Alzheimer patients treated with cholinesterase inhibitors. *International Journal of Psychophysiology* 64, pp. 46 – 52.
- [122] G. Yener, B. Guntekinb and E. Basar., (2008), Event-related delta oscillatory responses of Alzheimer patients. *European Journal of Neurology*, 15: pp. 540–547.
- [123] Miwakeichi F., Martinez-Montes E., Valdes-Soza A. P., Nishiyama N., Mizuhara H., Yamaguchi Y., (2004). Decomposing EEG data into space-time-frequency components using Parallel Factor Analysis. *NeuroImage* 22, pp. 1035-1045.
- [124] Wang T., Deng J., He B., (2004). Classifying EEG-based motor imagery tasks by means of time-frequency synthesized spatial patterns. *Clinical Neurophysiology* 115, pp. 2744-2753.
- [125] Melissant C., Ypma A., Frietman EE., Stam CJ., (2005). A Method for Detection of Alzheimer's disease using ICA-Enhanced EEG Measurements. *Artificial Intelligence in Medicine*, 33, pp. 209-222.
- [126] Lee PL., Wu YT., Chen LF., Chen YS., Cheng CM., Yeh TC., Ho LT., Chang MS., Hsieh JC., (2003). ICA-based spatiotemporal approach for single-trial analysis of postmovement MEG beta synchronization. *NeuroImage* 20, pp. 2010-2030.
- [127] Lehman D., Faber P.L., Gianotti L.R.R., Kochi K., Pascual-Marqui R.D., (2006). Coherence and phase locking in the scalp EEG and between LORETA model sources, and microstates as putative mechanisms of brain tempo-

- spatial functional organization. *Journal of Physiology-Paris*, Vol. 99, pp. 29-36.
- [128] Delorme A., Makeig S., Fabre-Thorpe M., Sejnowski T., (2002). From Single-trial EEG to Brain Area Dynamics. *Neurocomputing* 44-46, pp. 1057-1064.
- [129] Meinecke F. C., Ziehe A., Kurths J., Müller K. R., (2005). Measuring Phase Synchronization of Superimposed Signals. *The American Physical Society*, pp. 1-4.
- [130] Allefeld C., Frisch S., Schlesewsky M., (2005). Detection of Early Cognitive Processing by Event-Related Phase Synchronization Analysis. *Neuroreport*, 16(1), pp. 13-16.
- [131] Sakkalis V., Tsiaras V., Michalopoulos K., Zervakis M., (2008). Assessment of neural dynamic coupling and causal interactions between independent EEG components from cognitive tasks using linear and nonlinear methods. *EMBS*, pp. 1-4.
- [132] L. A. Baccala, K. Sameshima, (2001), Partial directed coherence: a new concept in neural structure determination, *Biological Cybernetics* vol. 84, No 6 , pp. 463-474.
- [133] Makeig S., Delorme A., Westerfield M., Jung TP., Townsend J., Courchesne E., Sejnowski T. J., (2004). Electroencephalographic Brain Dynamics following manually responded visual targets. *Plos Biology*, Vol. 2, pp. 747-762.
- [134] Spironelli C., Angrilli A., (2009), EEG delta band as a marker of brain damage in aphasic patients after recovery of language, *Neuropsychologia* 47, pp. 988–994.
- [135] [http://en.wikipedia.org/wiki/Factor\\_analysis](http://en.wikipedia.org/wiki/Factor_analysis)
- [136] J. H. Friedman and J. W. Tukey. , (1974), A projection pursuit algorithm for exploratory data analysis. *IEEE Trans. of Computers*, c-23(9):pp. 881-890.
- [137] P.J. Huber., (1985), Projection pursuit. *The Annals of Statistics*, 13(2): pp. 435-475.
- [138] M.C. Jones and R. Sibson., (1987), What is projection pursuit ? *J. of the Royal Statistical Society, ser. A*, 150: pp. 1-36.
- [139] J. Schmidhuber, (1996), M. Eldracher, and B. Foltin. Semilinear predictability minimization produces well-known feature detectors. *Neural Computation*, 8: pp.773-786.
- [140] H. B. Barlow. , (1972), Single units and sensation: A neuron doctrine for perceptual psychology? *Perception*, 1: pp. 371-394.
- [141] A.J. Bell and T.J. Sejnowski. , (1997), The 'independent components' of natural scenes are edge filters. *Vision Research*, 37: pp. 3327-3338.

- [142] Pfurtscheller G., Lopes de la Silva F.H., (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology* 110, pp. 1842-1857.

# Acknowledgments

At this point, I would like to thank everyone who helped me so as to conclude this work.

First of all, I would like to thank Professor Michalis Zervakis for his continue help and the precious guidelines he gave me. I would also like to thank him for his trust in the demanding field of research, where we managed to publish two contribution papers.

Two colleagues I would also like to thank are Kostas Michalopoulos and Eyaggelos Sakkalis, with whom we made a very good team so as to complete this work.

Finally, I would like to thank my family, who always stand by me.