



**TECHNICAL UNIVERSITY OF CRETE
DEPARTMENT OF ELECTRONIC AND
COMPUTER ENGINEERING
JUNE 2012**

STABILITY AND ANALYSIS OF INDEPENDENT COMPONENTS OF EEG.

MICHALOPOULOS KONSTANTINOS

SUPERVISOR: Prof. Zervakis Michalis

Acknowledgements

I would like to thank my advisor, Michalis Zervakis for his support, guidance and patience all this time. I would also like to thank Vassiliki Iordanidou for the great cooperation and friendly spirit. Special thanks to Vangelis Sakkalis for his help, advice and for being a great person. Finally I would like to thank Alexia for all the patience and support that she has displayed all this time!

ABSTRACT

Over the past years EEG analysis has gained focus as a low cost method to obtain insight in the complex brain processes and help us better understand the way brain works. In parallel it has been used in order to study and identify different brain pathologies, especially Alzheimer and epilepsy.

In this work we are focusing in the analysis of a specific type of EEG recordings, Event Related Potentials (ERPs). Event related potentials are essentially an EEG recording which contains different executions of the same experiment. Recent studies revealed a variety of activations and oscillations, from phase resetting of ongoing EEG activity in the different bands to phase-locked evoked and non-phase-locked induced oscillations. In this work we present methods and measures that can distinguish different patterns of activations and characterize the activity that takes place during the experiment.

Independent component analysis (ICA) is a popular method in the field of EEG signal processing. Independent Component Analysis(ICA) has been successfully applied 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. A major problem in application of ICA is that the stability of the estimated independent components is not known. Since the ICA techniques are stochastic, their results may be somewhat different in different runs of the algorithm due to different initial conditions. Other reasons are that the data does not always follow the ICA model exactly and in many cases the small number of training samples available is not adequate for the algorithm to run. We presented results in simulated and real EEG data that demonstrate the effect of sample size and initial weights on the stability of ICA results. We explored a methodology to evaluate the stability of the independent components before used in analysis and developed an analysis framework that takes under consideration the stable components.

We applied the proposed technique in two datasets with subjects suffering from pathologies related to aging. The first dataset consisted from control and Alzheimer subjects performing an auditory oddball experiment. The second dataset consisted from control and progressive MCI subjects performing a visual memory test. The results of this study indicate that the proposed component analysis and framework is able to depict the synchronized activations during a certain mental task (like the working memory). As such, it can efficiently reveal and quantify group (as well as individualized) differences in pathology (AD/ MCI) populations.

ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια η ανάλυση του Ηλεκτροεγκεφαλογραφήματος(ΗΕΓ) κερδίζει της προσοχής σαν μία φτηνή μέθοδο που μπορεί να μας δώσει πληροφορίες για τις περίπλοκες εγκεφαλικές διεργασίες και να μας βοηθήσει να κατανοήσουμε τον τρόπο λειτουργίας του εγκεφάλου. Παράλληλα χρησιμεύει στην μελέτη και ταυτοποίηση διαφόρων παθολογιών του εγκεφάλου, όπως η επιληψία και η νόσος Alzheimer.

Σε αυτή την εργασία επικεντρωνόμαστε στη μελέτη ενός ειδικού τύπου καταγραφής του ΗΕΓ που ονομάζεται Προκλητό Δυναμικό. Τα προκλητά δυναμικά αποτελούν την καταγραφή του ΗΕΓ κατά την διάρκεια επαναλαμβανόμενων δοκιμών μίας συγκεκριμένης εργασίας απο τον εξεταζόμενο. Πρόσφατες μελέτες έδειξαν ότι τα Προκλητά Δυναμικά εμπεριέχουν μία ποικιλία απο δραστηριοποιήσεις του εγκεφάλου που ποικίλλουν, όπως συγχρονισμός της φάσης ή της ενέργειας του σήματος. Παρουσιάζουμε μεθόδους και μέτρα με τα οποία μπορούμε να διαχωρίσουμε και να αναγνωρίσουμε τους διαφορετικούς τύπους ενεργοποίησης που λαμβάνουν χώρα κατά την διάρκεια του πειράματος.

Η ανάλυση ανεξάρτητων συνιστωσών (ICA) είναι μία ιδιαίτερα δημοφιλής μέθοδος για την ανάλυση του ΗΕΓ. Έχει με επιτυχία εφαρμοστεί σε προκλητά δυναμικά αναλύοντας το σήμα σε χωρικά σταθερές και χρονικά ανεξαρτητες συνιστώσες με διαφορετικές κατανομές στα ηλεκτρόδια που μπορούν να αποδοθούν σε συγκεκριμένες εγκεφαλικές δραστηριότητες. Ένα σοβαρό πρόβλημα στην εφαρμογή του ICA είναι ότι η ευστάθεια του αποτελέσματος δεν είναι εγγυημένη. Εφόσον η διαδικασία ICA είναι στοχαστική τα αποτελέσματα μπορεί να είναι σχετικά διαφορετικά σε διαφορετικές εκτελέσεις του αλγορίθμου στα ίδια δεδομένα λόγω διαφορετικών αρχικών συνθηκών. Άλλοι λόγοι είναι το γεγονός πως τα δεδομένα ΗΕΓ δεν ακολουθούν πιστά το μοντέλο ICA και πολλές φορές τα δεδομένα δεν είναι αρκετά ώστε ο αλγόριθμος να συγκλίνει σε μία σταθερή λύση. Παρουσιάζουμε αποτελέσματα σε προσομοιωμένα δεδομένα και πραγματικά δεδομένα ΗΕΓ που δείχνουν την επίδραση του μεγέθους του δείγματος και των αρχικών συνθηκών στο παραγόμενο αποτέλεσμα. Τέλος εφαρμόζουμε μία μεθοδολογία για την εκτίμηση της ευστάθειας των συνιστωσών πριν τις χρησιμοποιήσουμε στην ανάλυση και προτείνουμε μία συνολική μεθοδολογία που παίρνει υπόψη της την ευστάθεια των συνιστωσών

Εφαρμόζουμε την μεθοδολογία σε δύο ομάδες δεδομένων που αποτελούνται απο καταγραφές ασθενών με παθολογίεςσχετιζόμενες με την γήρανση. Η πρώτη ομάδα αποτελείτε απο καταγραφές υγιών ατόμων και ατόμων που πάσχουν απο τη νόσο Alzheimer ενώ εκτελούν ένα ακουστικό πείραμα. Η δεύτερη ομάδα αποτελείτε απο υγιείς και άτομα που έχουν διαγνωστεί με προχωρημένη άνοια ενώ εκτελούν ένα οπτικό πείραμα μνήμης. Τα αποτελέσματα δείχνουν ότι η προτεινόμενη μέθοδος

ανάλυσης είναι ικανή να αποδώσει της συγχρονισμένες ενεργοποιήσεις κατά τη διάρκεια του πειράματος. Με αυτό τον τρόπο μπορέσαμε να εκτιμήσουμε και να ποσοτικοποιήσουμε διαφορές ανάμεσα στις ενεργοποιήσεις των παθολογιών.

Table of Contents

Chapter 1.	Introduction	10
Chapter 2.	Electroencephalography and Event Related Potentials	14
2.1	Electroencephalography	14
2.2	Brain Activity patterns.....	15
2.3	Event Related Potentials	18
2.4	Alzheimer’s Disease.....	20
2.4.2	Alzheimer Disease Symptoms and stages.....	22
2.4.3	Alzheimer’s effects on Event Related Activations.	23
Chapter 3.	Methods and Measures of ERP analysis.	27
3.1	Introduction.....	27
3.2	Methods	27
3.2.1	Independent Component Analysis on EEG data	27
3.2.2	Time-Frequency Transforms	28
3.2.3	The Short-Time Fourier Transform	29
3.2.4	Wavelet Transform	30
3.2.5	Wigner-Ville distribution.....	30
3.2.6	Application and comparison of time-frequency transforms.	32
3.2.7	Decomposition of Time-Frequency Data	32
3.3	Measures of evoked and induced activity	35
3.3.1	Event Related Synchronization and Desynchronization.....	37
3.3.2	Phase Intertrial Coherence (PIC).....	38
3.3.3	Phase-shift Intertrial coherence (PsIC)	38
3.3.4	Intertrial Power Entropy	39
3.3.5	Induced and Global Energy	40
3.3.6	Application of ERP activity measures.	40
3.4	Discussion.....	41
Chapter 4.	Independent component analysis application in EEG data.	44
4.1	Independent Component Analysis.....	44
4.2	ICA model.	44
4.3	Statistical Independence	45
4.4	Estimating Independence.....	46

4.5	ICA preprocessing.....	47
4.6	Infomax Algorithm	49
4.7	Application to EEG data.....	51
4.7.1	Selecting interesting Independent Components.....	53
4.8	Discussion.....	58
Chapter 5.	Stability Estimation methods.....	59
5.1	Introduction.....	59
5.2	Definition of Stability.....	59
5.3	Approaches to stability estimation	60
5.4	Bootstrapping.....	61
5.4.1	Using Bootstrap for estimating ICA stability.....	63
5.4.2	Grouping Estimates.....	64
5.4.3	Clustering the estimates	64
5.4.4	Visualization by nonlinear 2D projection.....	65
5.4.5	Experimental Results	66
5.4.6	Advantages and disadvantages of the bootstrap method.	72
5.5	Split-Half approach.....	73
5.5.1	Splitting the data.....	74
5.5.2	Pairing the Independent Components	74
5.5.3	Significantly similar Independent Components.....	75
5.5.4	Advantages and disadvantages of split-half method.	77
Chapter 6.	Proposed Methodology.	79
6.1	Introduction.....	79
6.2	Experimental Setup of the Acoustic Experiment	80
6.3	Evaluation of components of the 1 st experiment	81
6.3.1	Control Subjects.....	82
6.3.2	Alzheimer subjects	87
6.3.3	Discussion of activation patterns.....	90
6.3.4	Evaluation of Unstable Components.	91
6.4	Application on a working memory experiment	92
6.4.1	Dataset description	92
6.4.2	Procedure.....	93

6.5	Evaluation of Components of 2 nd Experiment	93
6.5.1	Control Subjects	93
6.5.2	PMCI Subjects	95
6.5.3	Discussion of activation patterns.....	96
6.6	Comparison between the two datasets.....	97
6.7	Discussion.....	104
Chapter 7.	Conclusions	106
7.1	Future Work	107

Chapter 1. Introduction

Event related or event locked activity induced by an external or internal stimulus involves both phase locked and non-phase locked rhythmic oscillations. Phase-locked (evoked) activity can be observed at different frequency bands as a result of a single or multiple brain processes, arising at different brain locations. Recent studies have also revealed responses that are non-phase locked to the event occurrence (induced), which vary with stimulus and interact with the ERP. The origins of brain sources relate to multiple task conditions and many stimulus types that define distinct-topography brain functions, some operating independently and some being coupled [1].

The ERP represents an important and extensively studied brain response. It has been proven extremely useful in clinical and physiological research. There is a rich literature about the functional meaning of the different peaks of ERP (such as the P1, N1 and P3), which are thought to reflect different aspects of information processing in the brain [2]. This classical point of view states that ERPs are generated by fixed latency, phase-locked responses [2]. Its underlying assumption implies that the interesting ERP response is evoked by the task and can be detected by averaging the recorded signals over trials, which increases the signal-to-noise ratio (SNR) in the average signal [3].

Induced activities are expressed through the increase or decrease of energy in a specific band post-stimulus, denoted as event related synchronization (ERS) or desynchronization (ERD), respectively [4]. Non phase-locked oscillations have been associated with a variety of different functions related to perception and different types of cognitive processes [5]. Extensive findings correlate alpha energy and alpha phase, on stimulus onset with the ERP amplitude, indicating that the ERP and EEG oscillations interact and relate to each other [6]. Furthermore, one part of ERP has been found to associate with phase resetting of ongoing EEG activities [7], which reorganizes the phase at a particular frequency to produce the ERP without implying the generation of a new response. Evoked and induced oscillations may be considered as coupled processes progressing in time, with different spatial localization of origin and partially overlapping frequency content [8].

Many studies suggest 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 [9] to phase-locked evoked and non-phase-locked induced oscillations especially in delta, theta and gamma bands [3], [8]. Their origins relate to multiple task conditions and many stimulus types engaged during the event presentation and execution of its consequent actions [10], which define distinct brain functions some operating independently and some being coupled [1]. Because of their separate neurophysiologic origins, phase-locked (evoked) and non phase-

locked (induced) responses are of different nature [8] and have different functional roles [11], even though they may correlate with similar cognitive events [11], [12]. Furthermore the induced activity has also been characterized as event-related spectral perturbations (ERSP)[13]. ERSP measures have been considered in relation to quantified changes (increase or decrease) in power of specific frequency bands relative to mean prestimulus power, which is also termed as event related desynchronization/ synchronization (ERD/ERS) [4].

A variety of methods to characterize the nature of EEG activity in terms of their major time/frequency activity and topographic origin has been employed. Phase-relevant techniques as intertrial coherence (ITC) have been used as measures to characterize the phase consistency of the detailed TF content throughout trials [14]. In our work we propose a similar measure for quantifying phase-locked activity in ERP trials called phase intertrial coherence (PIC). Alternatively, the techniques based on average power relative to prestimulus reveal non-phase-locked (induced) activity associated with ERD and ERS [15]. In case of evoked (ERP) activity, both phase-relevant (ITC and PIC) and power-related (ERSP) measures show significant change, whereas only phase-relevant measures are sensitive to phase resetting. In order to provide a measure of induced activity directly related to signal consistency over trials, the phase-shift intertrial coherence (PsIC) measure has been introduced. It considers the amplitude (or power) consistency of the post-stimulus part over the trials, without reference to the pre-stimulus. As such, PsIC is sensitive to phase and also frequency resetting, which reorganize and synchronize ongoing processes in terms of either phase or frequency attributes. Also, a method to extract the induced activity out of the global energy is to subtract the ERP out of each trial before applying the measures [16] is evaluated in terms of separate evaluation of the evoked and induced energy.

Taking under consideration the above measures that quantify the different signal features using average values across trials we introduce a measure that enables the quantification and comparison of the distribution of these features across trials, the so called intertrial entropy. The motive for developing such a measure is that a very interesting and quite neglected aspect of ERP, trial to trial variability of the activations, is important for the physiological interpretation of the activations that take place during the experiment. In general measures using the trial mean of a metric do not characterize the variability of the metric across trials. Trial to trial variability may give more insight into the generation of the brain processes that arise during the response. Reference [7] exhibits significant results identifying processes that despite their equivalent mean dynamics, present strong amplitude variability in the single trial data. Identifying such processes effectively alters the interpretation of their physiological function. We introduce and propose the notion of intertrial entropy as a measure of power variability between trials. Entropy has been used

before for ERP data in the form of wavelet entropy [17],[18] in order to characterize the degree of order/ disorder in a multi-frequency response and its time evolution. We also describe and demonstrate the use of the different measures and how they can be used in conjunction in order to reveal different aspects of the underlying activity. In addition we demonstrate the additional information provided by using intertrial entropy.

It is apparent that the brain functions and activations that take place during an ERP experiment are much more than the simple averaging can reveal. These activations arise independently in different brain locations and coordinate in order to deliver the required response. The mixed activations of these functions are captured in the electrodes, sometimes masked by task irrelevant ongoing EEG oscillations. It is apparent that we need a method to decouple the mixed activity captured in the electrodes into the underlying generating sources. Independent component analysis (ICA) is a widely adopted method which has gained great focus recently. It is able to decompose the ERP signal into independent components which reflect meaningful physiological properties and separates extra cerebral noise from brain activations. ICA has been found useful even though the EEG data do not follow the ICA model exactly and the number of brain sources activating probably exceeds the number of sensors used in the recording. Also since ICA is a stochastic algorithm that takes advantage of the data distribution is affected by the sample size used to calculate the independent components and the starting initial weight matrix used. It is possible that the algorithm may converge in different local maximum/ minimum when run multiple times, producing components with different activations and scalp topography.

The stability of ICA is less explored in the bibliography and is often neglected when ICA used in analysis. Empirical observations have reported that ICA decomposes the data into maximally independent components, separating large sources into single components and tends to group together small activations. We present results in simulated and real EEG data that demonstrate the effect of sample size and starting initial weights in the stability of the algorithm. We present a methodology to evaluate the stability of the independent components before used in analysis.

In this work we present measures that quantify phase-locked and induced activations in an ERP experiment and propose a methodology for exploring and analyzing ERP data. We also propose a methodology for working with ICA taking into account the stability of the resulting components. We applied the developed techniques to a dataset of 9 control and 9 Alzheimer subjects and evaluate the differences in the activations of the two groups.

In chapter 1 we describe the basic bands and activations found in EEG and present the basic assumption behind event related potentials. We also describe the Alzheimer disease and the effects in the patients' resulting ERP activations.

In chapter 2 we present the measures used for identifying phase-locked and non-phase-locked activity and an application of the techniques in a single control subject, illustrating the purpose the usefulness of the measures.

In chapter 3 we present in detail the ICA algorithm and its application in EEG data. We also present techniques used to discriminate independent components that contain interesting activity from unrelated EEG activity and noise.

Chapter 4 deals with the stability of the ICA algorithm. We present a method for evaluating the stability of independent components based in bootstrap statistics and multiple runs of the algorithm in the same data. We cluster together activations with similar topographies that represent stable components found in each run of the algorithm. We apply this technique in different scenarios changing the sample size of the data and initial weights of the algorithm in order to illustrate the effect of each in the stability of the resulting components.

Finally, a procedure based in a simplification of the notion of bootstrap statistics is presented. This method is based in a well known technique for evaluating stability and reliability, called split-half comparison. The main advantage of this technique is that it needs much less resources and computational time than the bootstrap statistics and makes it easy in application in real life experiments.

In Chapter 6 we apply the split-half comparison technique in a dataset of 9 control and 9 Alzheimer subjects. We evaluate the resulting stable components and common activation patterns were found for each group. The results found are in agreement with the results in the bibliography and significant differentiations were found in the activations between control and Alzheimer subjects. This method was compared with the results obtained when using all independent components despite being stable or not. The results found in both cases are in agreement enhancing the effectiveness of the proposed methodology.

Chapter 2. Electroencephalography and Event Related Potentials

2.1 Electroencephalography

The electroencephalogram (EEG) is a record of the oscillations of brain electric potentials recorded from electrodes attached to the human scalp. The first human EEG recordings were accomplished by the German psychiatrist Hans Berger in 1924[1]. Berger's subjects revealed robust changes when the eyes were closed and when the subjects performed mental arithmetic. Since then, electric patterns of the EEG have been correlated to neurological disorders, as epilepsy and Alzheimer.

Electric potentials can be acquired either between pairs of active electrodes, also known as bipolar montage or in respect to a passive electrode (reference electrode), known as monopolar montage. The location of the electrodes placed on the scalp is defined by established standards since 1949. An international committee was assembled under Rasmussen and designed the international 10-20 electrode placement system. Such system consists of 20 electrodes uniformly distributed along the head generally referenced to two earlobe (A1, A2) electrodes[2]. EEG is usually taken under specified conditions and subject mental state, like relaxed (i) with eyes closed, (ii) with eyes open or (iii) while the subject performs predefined cognitive tasks. The major difficulty one encounters while performing scalp EEG recordings are the artifacts[2]. A variety of different in morphology artifacts arise mainly due to head movements, eye blinking, muscle activity, electrocardiogram (ECG) recording and so on.

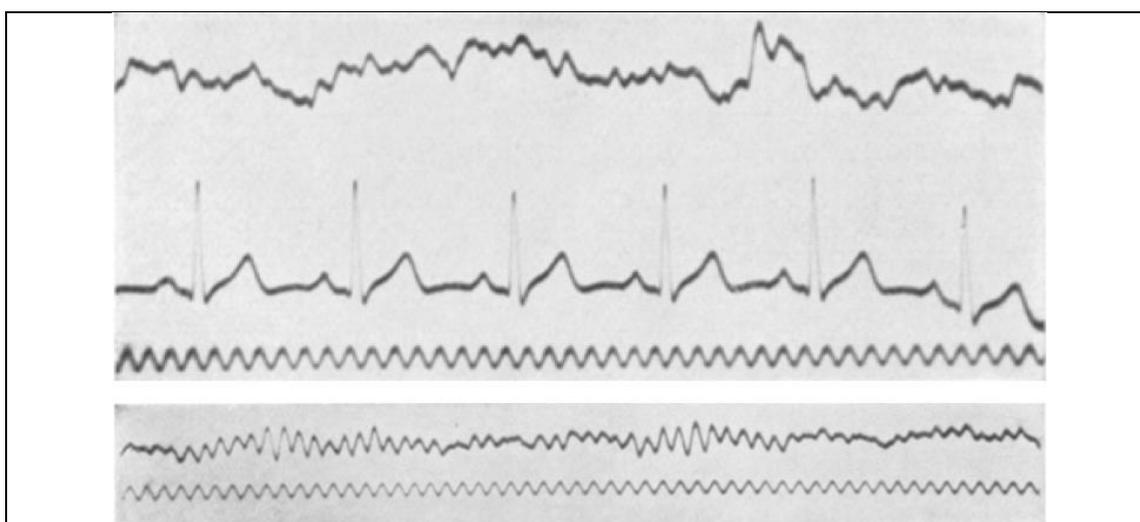


Figure 2.1. EEG recordings appeared in Berger's publication in 1929. These are samples of EEG recorded from his son Klaus (16 years old). The bottom figure is an EEG sample with 10Hz frequency what it is now well known as the 'alpha rhythm'. The figure above what he would later call the 'beta rhythm'.

2.2 Brain Activity patterns

Even from his first publication on EEG, Berger[1] mentioned the presence of certain signal patterns (rhythms) that he called alpha and beta oscillations (Fig. 2.1). The frequency patterns seen in the EEG, have been divided into the well known bands of delta, theta, alpha and beta, which are empirically determined, frequency limits. Since the beginnings of electroencephalography these patterns have been related with different brain arousal states, functions or pathologies[2]. Although these patterns may found to vary across different individuals, their frequency remained relatively the same. However, as we noted these bands were empirically defined and to this day there exists no specific standard. For this reason the detailed frequency limits are often noted alongside with the band names. The most "popular" band definitions are presented here:

- **Delta (δ) band (<4Hz):** Delta band starts as low as the band pass filter allows and extends up to 3/4Hz. 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. It is usually most prominent frontally in adults (e.g. FIRDA - Frontal Intermittent Rhythmic Delta) and posteriorly in children (e.g. OIRDA - Occipital Intermittent Rhythmic Delta).
- **Theta (θ) band (3 - 8HZ):** is the frequency range from 4 Hz to 7 Hz. 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.[38] 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. On the contrary this range has been associated with reports of relaxed, meditative, and creative states.
- **Alpha (α) band (8-13Hz):** is the frequency range from 8 Hz to 12 Hz. Hans Berger named the first rhythmic EEG activity he saw as the "alpha wave". This was the "posterior basic rhythm" (also called the "posterior dominant rhythm" or the "posterior alpha rhythm"), seen in the posterior regions of the head on both sides, higher in amplitude on the dominant side. It emerges with closing of the eyes and with relaxation, and attenuates with eye opening or mental exertion. The posterior basic rhythm is actually slower than 8 Hz in young children (therefore technically in the theta range)[39][40]. 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".
- **Beta (β) band (13-30Hz):** is the frequency range from 12 Hz to about 30 Hz. It is seen usually on both sides in symmetrical distribution and is most evident frontally. Beta activity is closely linked to motor behavior and is generally

attenuated during active movements.[41] Low amplitude beta with multiple and varying frequencies is often associated with active, busy or anxious thinking and active concentration. Rhythmic beta with a dominant set of frequencies is associated with various pathologies and drug effects, especially benzodiazepines. 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.

- **Gamma (γ) band (>30Hz):** is the frequency range approximately 30–100 Hz. Gamma rhythms are thought to represent binding of different populations of neurons together into a network for the purpose of carrying out a certain cognitive or motor function.

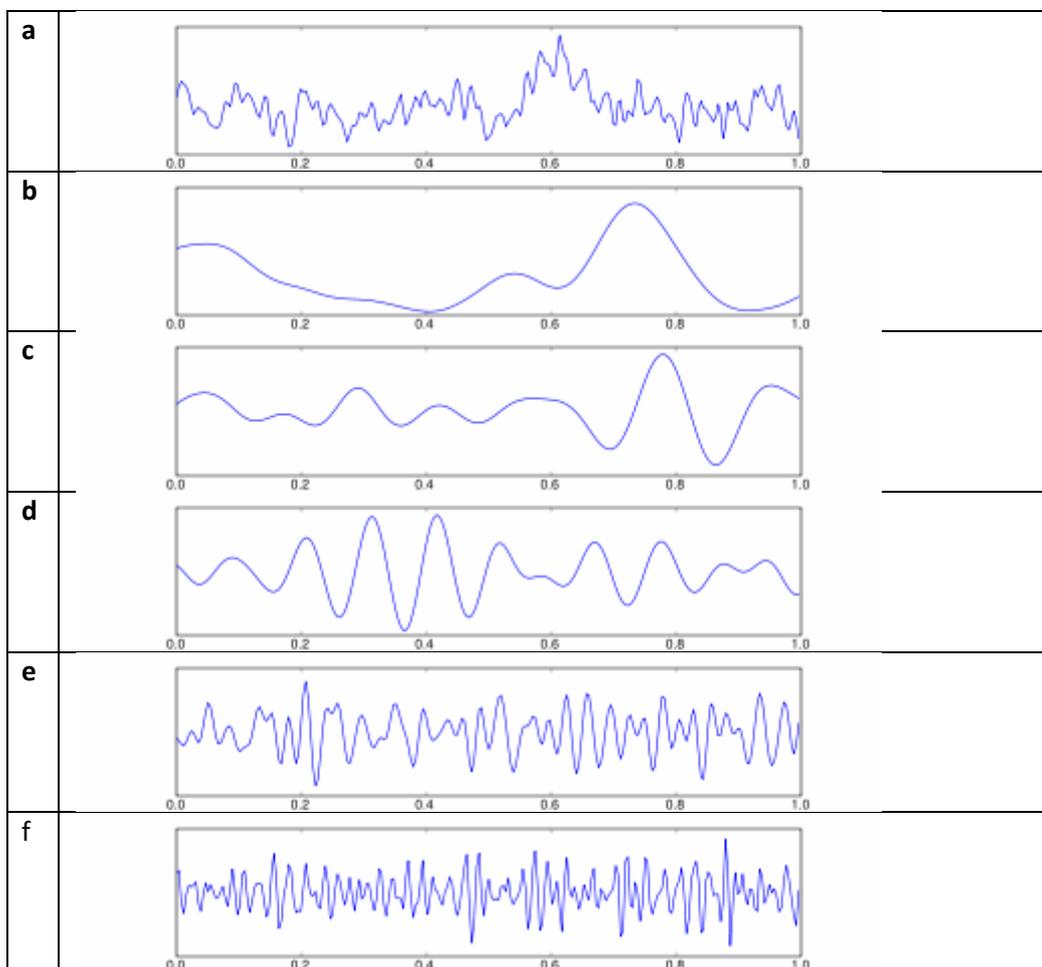


Figure 2.2: Examples of different EEG waveform patterns. From a to f respectively: part of an EEG recording (single electrode), delta, theta alpha, beta and gamma band activity.

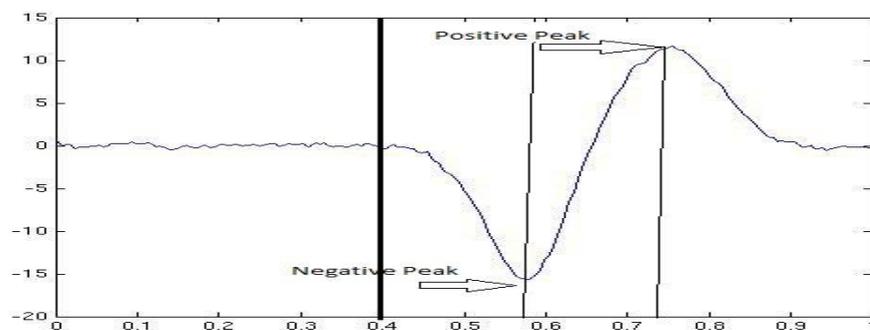
Type	Frequency (Hz)	Location	Normally	Pathologically

Delta	up to 4	frontally in adults, posteriorly in children; high amplitude waves	<ul style="list-style-type: none"> Adults slow wave sleep in babies Has been found during some continuous attention tasks^[33] 	<ul style="list-style-type: none"> subcortical lesions diffuse lesions metabolic encephalopathy hydrocephalus deep midline lesions
Theta	4 – 8	Found in locations not related to task at hand	<ul style="list-style-type: none"> young children drowsiness or arousal in older children and adults idling Associated with inhibition of elicited responses (has been found to spike in situations where a person is actively trying to repress a response or action).^[33] 	<ul style="list-style-type: none"> focal subcortical lesions metabolic encephalopathy deep midline disorders some instances of hydrocephalus
Alpha	8 – 13	posterior regions of head, both sides, higher in amplitude on non-dominant side. Central sites (c3-c4) at rest	<ul style="list-style-type: none"> relaxed/reflecting closing the eyes Also associated with inhibition control, seemingly with the purpose of timing inhibitory activity in different locations across the brain. 	<ul style="list-style-type: none"> coma
Beta	>13 – 30	both sides, symmetrical distribution, most evident frontally; low amplitude waves	<ul style="list-style-type: none"> alert/working active, busy or anxious thinking, active concentration 	<ul style="list-style-type: none"> benzodiazepines
Gamma	30 – 100+	Somatosensory cortex	<ul style="list-style-type: none"> Displays during cross-modal sensory processing (perception that combines two different senses, such as sound and sight) Also is shown during short term memory matching of recognized objects, sounds, or tactile 	<ul style="list-style-type: none"> A decrease in gamma band activity may be associated with cognitive decline, especially when related the theta band; however, this has not been proven for use as a clinical diagnostic measurement yet

2.3 Event Related Potentials

Brain activity induced by an internal or external stimulus involves the organized activation of different neural assemblies at different brain locations. Recording of EEG during an event related experiment captures electrical activity arising from the synchronized activation of neuronal assemblies each contributing to observable EEG characteristics[3]. These activations are captured in the scalp electrodes either as phase-locked (evoked) or non-phase-locked (induced) oscillations. Phase-locked activity can be directly extracted by averaging of the EEG single trials, resulting in the well known Event Related Potential (ERP).

The ERP represents an important and extensively studied brain response. It has been proven extremely useful in clinical and physiological research. There is a rich literature about the functional meaning of the different peaks of ERP (such as the P1, N1 and P3), which are thought to reflect different aspects of information processing in the brain [4]. This classical point of view states that ERPs are generated by fixed latency, phase-locked responses [4]. Its underlying assumption implies that the interesting ERP response is evoked by the task and can be detected by averaging the recorded signals over trials, which increases the signal-to-noise ratio (SNR) in the average signal [5].



2.3: Example of an averaged ERP waveform. The valleys and peaks represent task relevant activity, which activates in the same manner through trials. Amplitude and latency of these peaks are considered important features for analysis.

Event related potentials (ERPs) have been considered as originating from stable phase-locking due to transient synchronization of underlying neural substrates caused by the stimulus onset. A technique applied for the detection of the ERPs is averaging of the single trial responses which increases SNR of the low amplitude responses in each single trial[5]. Underlying this approach is the assumption that due to the event, a brain process arises independently from ongoing EEG at the moment,

in order to serve the response[4]. Such activations are known as *evoked* responses[6]. Averaging of the single trials and inspection of the characteristics of the resulting ERP waveform, has given a lot of information about the brain processes that take place during an ERP experiment and has been proven extremely useful, from biomarkers for brain diseases [7] to brain computer interfaces[8]. FIGURE presents an overview of frequent ERP peaks that have been identified and linked to brain processes.

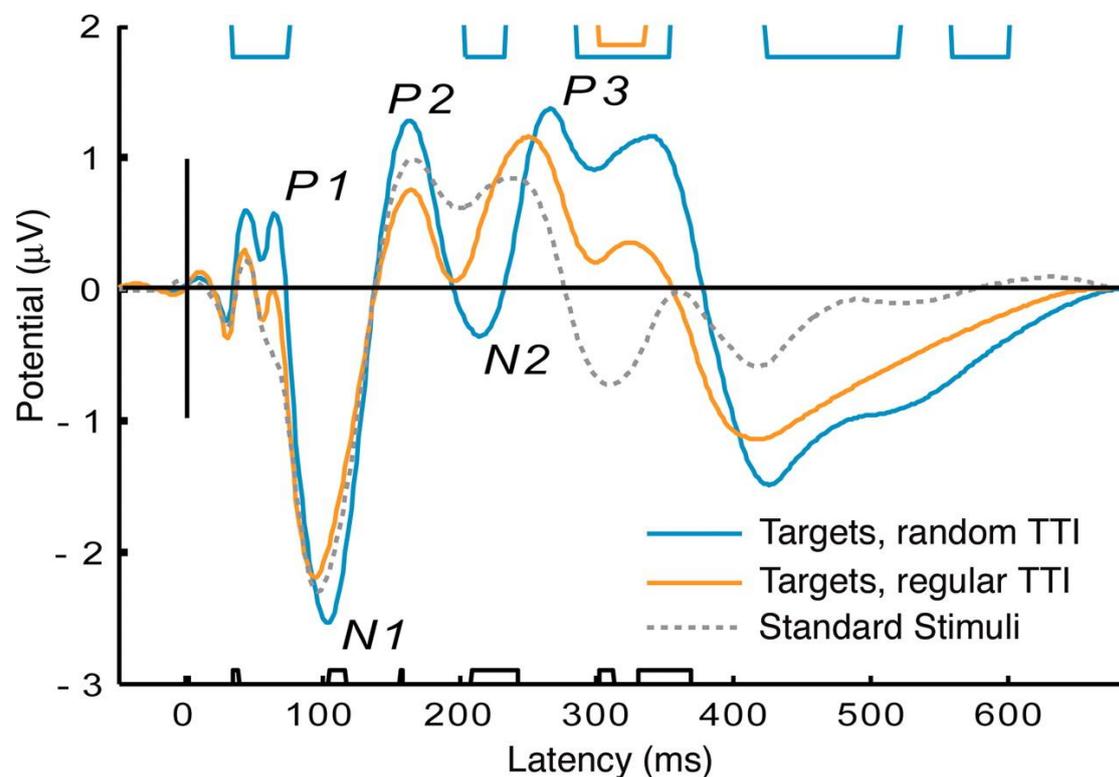


Figure 2.4: Illustration of different ERP peaks that have been identified. Notice how the peak polarity and latency determines the name of the peak.

Several studies [5, 9] though, revealed associations between the ongoing EEG and the ERP, putting the validity of the model under question. For example in[10], the ERP amplitude was associated with the alpha phase on stimulus onset. Alongside with such findings, studies revealed that during an ERP experiment, processes take place that cannot be observed through the averaging procedure. Specifically, increase or decrease in the post-stimulus energy of specific EEG bands compared to the pre-stimulus were observed, which was time-locked but not phase locked to the event and is referred to as *induced* activity[6]. Actually, the transient event-related activity seems to be elicited from evoked and induced response sources, each having a distinct topographic organization [11, 12].

Event related brain activity has been considered to contain many different types of oscillations such as alpha and theta activity caused by phase resetting of ongoing

EEG activity [13] or delta, theta and gamma activity observed in phase-locked and induced (non phase-locked) brain activity[5, 9, 14]. This kind of activations takes place in conditions of multitasking in which during the event presentation and execution of its consequent actions many stimulus types take place[11], which represents a set of different brain functions, some operating distinctly and some operating pair wise[15]. Phase and non phase-locked oscillations originate from different neurophysiologic brain activations, so even though they may occur during similar cognitive procedures, they have different nature [9] and functional roles[11, 12]. More specifically, they can be considered as coupled processes in time, each originating from a different spatial topography and having partially overlapping frequency content[11, 12]. Thus, it is of great importance to observe alterations in EEG so as to match their neurophysiologic origin and pathological cases.

Separating and identifying the contribution of EEG activity 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.) [6, 12] can reveal aspects of the brain functions that take place during the execution of a task and lead to possible biomarkers for different pathologies.

2.4 Alzheimer's Disease.

Alzheimer's disease is an irreversible, progressive brain disease that slowly destroys memory and thinking skills and, eventually, the ability to carry out the simplest tasks. Memory problems are one of the first signs of Alzheimer's. People may have trouble remembering things that happened recently or names of people they know. Over time, symptoms will most often get worse, and problems can include getting lost, repeating questions, and taking longer than normal to finish daily tasks. As the disease progresses, people may have trouble learning new things, recognizing family and friends, and communicating.

The disease is named after Dr. Alois Alzheimer. In 1906, Dr. Alzheimer noticed changes in the brain tissue of a woman who had died of an unusual mental illness. Her symptoms included memory loss, language problems, and unpredictable behavior. After she died, he examined her brain and found many abnormal clumps (now called amyloid plaques) and tangled bundles of fibers (now called neurofibrillary tangles).

Alzheimer's disease disrupts critical metabolic processes that keep neurons healthy. These disruptions cause nerve cells in the brain to stop working, lose connections with other nerve cells, and finally die. The destruction and death of nerve cells causes the memory failure, personality changes, problems in carrying out daily activities, and other features of the disease.

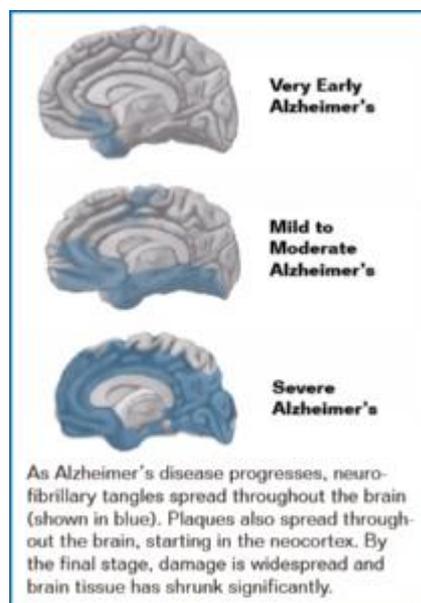
The brains of people with AD have an abundance of two abnormal structures—amyloid plaques and neurofibrillary tangles—that are made of misfolded proteins. This is especially true in certain regions of the brain that are important in memory. The third main feature of AD is the loss of connections between cells. This 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.

We now know 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. We still do not know whether amyloid plaques themselves cause AD or whether they are a by-product of the AD process.

For many years, scientists thought that plaques might cause all of the damage to neurons that is seen in AD. However, that concept has evolved greatly in the past few years. Many scientists now think that oligomers may be a major culprit. Many scientists also think that plaques actually may be a late-stage attempt by the brain to get this harmful beta-amyloid away from neurons.



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.

LOSS OF CONNECTION BETWEEN CELLS AND CELL DEATH

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.

2.4.2 Alzheimer Disease Symptoms and stages.

Very Early Signs and Symptoms

Memory problems are typically one of the first warning signs of cognitive loss, possibly due to the development of Alzheimer's disease. Some people with memory problems have a condition called amnesic mild cognitive impairment (MCI). People with this condition have more memory problems than normal for people their age, but their symptoms are not as severe as those seen in people with Alzheimer's disease. Other recent studies have found links between some movement difficulties and MCI. Researchers also have seen links between MCI and some problems with the sense of smell. The ability of people with MCI to perform normal daily activities is not significantly impaired. However, more older people with MCI, compared with those without MCI, go on to develop Alzheimer's.

A decline in other aspects of cognition, such as word-finding, vision/spatial issues, and impaired reasoning or judgment, may also signal the very early stages of Alzheimer's disease. Scientists are looking to see whether brain imaging and biomarker studies, for example, of people with MCI and those with a family history of Alzheimer's, can detect early changes in the brain like those seen in Alzheimer's. Initial studies indicate that early detection using biomarkers and imaging may be possible, but findings will need to be confirmed by other studies before these techniques can be used to help with diagnosis in everyday medical practice.

These and other studies offer hope that someday we may have tools that could help detect Alzheimer's early, track the course of the disease, and monitor response to treatments.

Mild Alzheimer's Disease

As Alzheimer's disease progresses, memory loss worsens, and changes in other cognitive abilities are evident. Problems can include, for example, getting lost, trouble handling money and paying bills, repeating questions, taking longer to complete normal daily tasks, using poor judgment, and having some mood and personality changes. People often are diagnosed in this stage.

Moderate Alzheimer's Disease

In this stage, damage occurs in areas of the brain that control language, reasoning, sensory processing, and conscious thought. Memory loss and confusion grow worse, and people begin to have problems recognizing family and friends. They may be unable to learn new things, carry out tasks that involve multiple steps (such as getting dressed), or cope with new situations. They may have hallucinations, delusions, and paranoia, and may behave impulsively.

Severe Alzheimer's Disease

By the final stage, plaques and tangles have spread throughout the brain, and brain tissue has shrunk significantly. People with severe Alzheimer's cannot communicate and are completely dependent on others for their care. Near the end, the person may be in bed most or all of the time as the body shuts down.

2.4.3 Alzheimer's effects on Event Related Activations.

Alzheimer disease has several effects on the brain which reflect on the resulting EEG. The degeneration of brain areas forces Alzheimer patients to revoke different brain resources to correspond to the task given. This results in altered activations to emerge in the recording EEG and evoked related potentials. We will describe such effects in the activations presented in typical tasks and the related differences in Alzheimer in the following paragraphs.

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 [16-18]. 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 [10]. Theta generally precedes delta in the P300 response and is more anterior in topography, while delta has rather posterior response [19, 20]. 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 [9, 18]. Furthermore, both fast and slow alpha have been related to specific aspects of cognitive processing indexed by P300 [9]. Several studies have also demonstrated that encoding typically induces alpha ERS activity, whereas retrieval elicits alpha ERD [17, 18]. 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)[18]. Similar results are reported in other studies, where theta and alpha synchronization is observed during the encoding of new items in short-term memory [21-23]. 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[18].

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[9]. 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[24] 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 [25]. 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[18]. 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[18].

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[26]. 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[27]. 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 [16, 28]. Significant group differences between controls and ADs have been reported in [27], 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 MEG[29]. Furthermore, the increase in relative theta power has been used to predict MCI patients that will progress to AD[30].

The multiple activations that take place during an ERP experiment introduce the need for a consistent approach in the analysis and quantification of such phenomena. In the next chapter we will introduce the methods and measures that will enable us to explore the different ERP activations.

Chapter 3. Methods and Measures of ERP analysis.

3.1 Introduction

A variety of methods to characterize the nature of EEG activity in terms of their major time/frequency activity and topographic origin has been employed. Phase-relevant techniques as intertrial coherence (ITC) and phase intertribal coherence (PIC) have been used as measures to characterize the phase consistency of the detailed TF content throughout trials [31]. Alternatively, the techniques based on average power relative to prestimulus reveal non-phase-locked (induced) activity associated with ERD and ERS[32]. In case of evoked (ERP) activity, both phase-relevant (ITC and PIC) and power-related (ERSP) measures show significant change, whereas only phase-relevant measures are sensitive to phase resetting. In order to provide a measure of induced activity directly related to signal consistency over trials, the phase-shift intertribal coherence (PsIC) measure has been introduced. It considers the amplitude (or power) consistency of the post-stimulus part over the trials, without reference to the pre-stimulus. As such, PsIC is sensitive to phase and also frequency resetting, which reorganize and synchronize ongoing processes in terms of either phase or frequency attributes. Also, a method to extract the induced activity out of the global energy is to subtract the ERP out of each trial before applying the measures[33].

Another very interesting aspect of ERP is the trial to trial variability of different activations that take place during the experiment. In general measures using the trial mean of a metric do not characterize the variability of the metric across trials. Trial to trial variability may give more insight into the generation of the brain processes that arise during the response. Reference [34] exhibits significant results identifying processes that despite their equivalent mean dynamics, present strong amplitude variability in the single trial data. Identifying such processes effectively alters the interpretation of their physiological function.

3.2 Methods

3.2.1 Independent Component Analysis on EEG data

Independent component analysis (ICA) is increasing in popularity in the field of biomedical signal processing. It is generally used when it is required to separate measured multi-channel biomedical signals into their constituent underlying components. Independent Component Analysis(ICA) 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[6, 12, 35].

In the next chapter we will discuss in detail the application of ICA in EEG data and the related problems that need to be addressed.

3.2.2 Time-Frequency Transforms

EEG signal analysis provides the advantage of high time resolution and thus it can deduce information related to both local and widespread neuronal activations in short time periods, as well as their time evolution. FFT basic idea is the decomposition of the signal into complex sinusoids that extend through the whole time domain. While, it reveals significant frequency information, there is no information concerning the time evolution of those frequencies. EEG is known to be a highly non-stationary signal, with its composing frequencies changing through time. The classical FFT cannot represent the evolution of frequency content through time and estimates the frequency change over time as a broad frequency spectrum.

A solution to this problem is based on the time-varying spectral analysis that takes into account the non-stationary nature of EEG [20]. The short-time Fourier (STFT) and the wavelet transforms are the most prevalent analysis frameworks of this class. The first approach uses a sliding time window, in which the classic FFT is computed. The second one forms the projection of the signal onto several oscillatory kernel-based wavelets matching different frequency bands. Currently, such time-varying methods have been widely applied in event-related potential (ERP) data, where distinct waveforms are associated with an event related to some brain function[22]. Both time-frequency transforms are in fact mathematically equivalent, since they both use windows that under certain conditions can provide the same results [23].

The essential difference is that the Wavelet transform (WT) is typically applied with the relative bandwidth ($\Delta f/f$) held constant, whereas the Fourier approach preserves the absolute bandwidth (Δf) constant. In other words, STFT uses a constant window length. The main implication is that according to the Gabor time-frequency uncertainty principle [25] one cannot accurately discriminate frequencies in small time intervals. Wavelet transform attempts to overcome the resolution problem treating different frequencies with different resolutions achieving good time resolution but poor frequency resolution at high frequencies and good frequency resolution but poor time resolution at low frequencies.

Owing to the extensive use of TF decomposition in our methodology, we examine two of the most widely used approaches, i.e. the wavelet decomposition [36, 37] and the Cohen's class of energy distributions[38].

The application of wavelet transforms in TF analysis is limited by the tradeoff between frequency and time, since wavelets compute small scale (high-frequency) intervals with shorter time windows and large scale regions (low-frequency) with longer time windows. As a result, they resolve energy in higher scales with high resolution in time but not in frequency. Alternatively, in lower scales they resolve energy in detailed frequency ranges but in larger time intervals.

In contrast to linear time-frequency representations, the purpose of the energy distributions is to distribute the *energy* of the signal over time and frequency. The working assumption is that since the energy of a signal can be computed from the squared modulus of either the signal or its Fourier transform, we can interpret these representations as energy densities in time and in frequency separately. From this point of view, the Cohen's class derives a joint time and frequency energy density. A time-frequency energy distribution which is particularly interesting is the *Wigner-Ville distribution* (WVD). This distribution satisfies a large number of desirable mathematical properties, as it is always real-valued, preserves time and frequency shifts and satisfies the marginal properties, i.e. if the time-frequency energy density is integrated along one variable, the energy density corresponding to the other variable is obtained. Next, we will briefly present the principles behind these transforms.

3.2.3 The Short-Time Fourier Transform

In order to introduce time-dependency in the Fourier transform, a simple and intuitive solution consists in pre-windowing the signal $x(u)$ around a particular time t , calculating its Fourier transform, and doing that for each time instant t . The resulting transform, called the short-time Fourier transform (STFT, or short-time spectrum), is

$$f_x(t, v; h) = \int_{-\infty}^{+\infty} x(u)h^*(u - t)e^{-j2\pi v u} du$$

Where $h(t)$ is a short time analysis window localized around $t=0$ and $v=0$. Because multiplication by the relatively short window $h^*(u - t)$ effectively suppresses the signal outside a neighborhood around the analysis time point $u = t$, the STFT is a "local" spectrum of the signal $x(u)$ around t . Provided that the short-time window is of finite energy, the STFT is invertible according to

$$F_x(t, v; h) = \int_{-\infty}^{+\infty} X(\xi)H^*(\xi - v)\exp[-j2\pi v u] d\xi$$

where X and H are respectively the Fourier transforms of x and h . Thus, the STFT $F_x(t, v; h)$ can be considered as the result of passing the signal $x(u)$ through a band-pass filter whose frequency response is $H^*(\xi - v)$, and is therefore deduced

from a mother filter $H(\xi)$ by a translation of ν . So the STFT is similar to a bank of band-pass filters with constant bandwidth.

The time resolution of the STFT is proportional to the effective duration of the analysis window h and frequency-resolution of the STFT is proportional to the effective bandwidth of the analysis window h . we have a trade-off between time and frequency resolutions: on one hand, a good time resolution requires a short window $h(t)$; on the other hand, a good frequency resolution requires a narrow-band filter i.e. a long window $h(t)$. We cannot satisfy both requirements because we are limited by the Heisenberg-Gabor inequality.

3.2.4 Wavelet Transform

Wavelet approaches decompose signals into constituent time–frequency ranges of energy based on the notion of scale applied to a set of basis functions. A set of wavelet functions is constructed from a single prototype wavelet, called the mother wavelet, by scaling (contraction, dilation) and shifts. By varying the so-called wavelet scale and translating along the time index, it is possible to identify the temporal evolution of various frequencies.

The continuous wavelet transform (CWT) of a discrete sequence x_n with time spacing δ_τ and N data points ($n = 0 \dots N - 1$) is defined as the convolution of x_n with consecutive scaled and translated versions of the wavelet function $\psi_0(n)$:

$$W_n(s) = \sum_{n=0}^{N-1} x_n \psi^*[(n - k)\delta t/s] \quad \text{Eq. 1}$$

$$\psi_0(k) = \pi^{-1/4} e^{i\omega_0 k} e^{-k^2/2} \quad \text{Eq. 2}$$

where k and ω_0 is a non dimensional time parameter and frequency, respectively.

As we noted before, WT is applied with the relative bandwidth ($\Delta f/f$) held constant. The time resolution becomes arbitrarily good at high frequencies, while the frequency resolution becomes arbitrarily good at low frequencies. In general wavelets perform best if the signal is composed of high-frequency components of short duration plus low-frequency components of long duration.

3.2.5 Wigner-Ville distribution

In contrast with the linear time-frequency representations which decompose the signal on elementary components (the atoms), the purpose of the energy distributions is to distribute the energy of the signal over the two description variables: time and frequency.

The starting point is that since the energy of a signal x can be deduced from the squared modulus of either the signal or its Fourier transform

$$E_x = \int_{-\infty}^{+\infty} |x(t)|^2 dt = \int_{-\infty}^{+\infty} |X(v)|^2 dv$$

we can interpret $|x(t)|^2$ and $|X(v)|^2$ as energy densities, respectively in time and in frequency. It is then natural to look for a joint time and frequency energy density $\rho_x(t, v)$, such that

$$E_x = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |x(t)|^2 dt dv$$

As the energy is a quadratic function of the signal, the time-frequency energy distributions will be in general quadratic representations.

The other properties that an energy density should satisfy are the following marginal properties:

$$\int_{-\infty}^{+\infty} \rho_x(t, v) dv = |X(v)|^2$$

$$\int_{-\infty}^{+\infty} \rho_x(t, v) dt = |x(t)|^2$$

which mean that if we integrate the time-frequency energy density along one variable, we obtain the energy density corresponding to the other variable.

The classes of time-frequency energy distributions covariant by translations in time and in frequency are known as Cohen classes. The Wigner-Ville distribution is an interesting distribution defined as:

$$W_x(t, v) = \int_{-\infty}^{+\infty} x(t + \tau/2) x^*(t - \tau/2) e^{-j2\pi v \tau} d\tau$$

Or

$$W_x(t, v) = \int_{-\infty}^{+\infty} X\left(v + \xi/2\right) X^*\left(v - \xi/2\right) e^{-j2\pi \xi t} d\xi$$

This distribution satisfies a large number of desirable mathematical properties, as is always real-valued, it preserves time and frequency shifts and satisfies the marginal properties.

3.2.6 Application and comparison of time-frequency transforms.

Both methods are capable of producing high-resolution TF distributions, but they differ significantly in their approach, implementation, and the inferences that can be drawn from their considerations. A difference between wavelets and WVD is that wavelets do not satisfy the time and frequency marginal conditions[39]. On the other hand, the WVD for the representation of two signals introduces interference terms, which are non-zero regardless of the time-frequency distance between the two signal terms. This interference renders the visual inspection and algorithmic interpretation of the time-frequency surface quite difficult. Modified techniques address the problems of interference terms, but they come at a high implementation cost. The computational time increases significantly and the marginal properties are no longer satisfied, making the advantages of WV less attractive compared to wavelets. The interference effect is depicted in Figure 3.1b(i), where we observe that although the WVD captures the two-chirp signal with high resolution in time and frequency, the interference terms are evident making the interpretation of results more difficult than the corresponding representation with wavelets Figure 3.1a(i). The Figure 3.1a(ii) presents a simulated signal composed of a set of four sinusoids without noise. We observe that the Wigner-Ville decomposition (1b(ii)) represents the signals more compact in time and frequency than the wavelets (1a(ii)), which at low frequencies diffuse the energy of signals (1b(ii)). The addition of EEG-like noise makes the effect of interference terms even more apparent, rendering the result very difficult to evaluate. Wavelet transform has the advantage that retains the phase information of the signal, an attribute that makes the use of wavelet transform very attractive. Following these considerations, we promote the use of wavelets in our study and implement TF decompositions by means of the Wavelet transform using the complex Morlet wavelet functions.

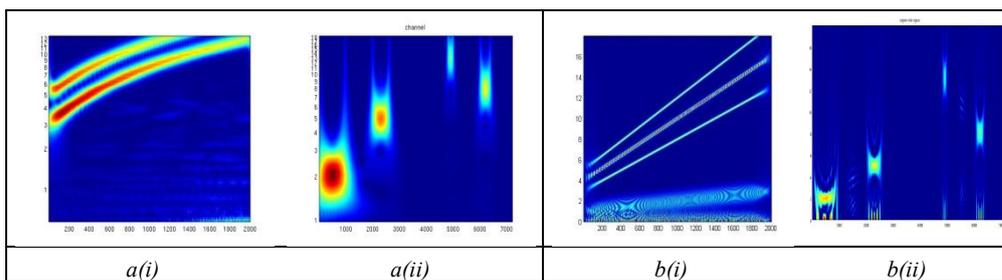


Figure 3.1: Time Frequency representations of (i) a two-chirp signal and (ii) a simulated EEG signal: 1a) Left figure presents the time-frequency transform with wavelets. 1b) Right figure depicts the time frequency representation using the Wigner-Ville distribution.

3.2.7 Decomposition of Time-Frequency Data

As we previously mentioned, several neurophysiological studies indicate that the evoked processes possibly originate from stable phase-locking transient synchronization of brain regions, with different signal peaks being evoked from specific brain regions at distinct frequency bands. Furthermore, the induced activity has been attributed to phase-resetting of ongoing EEG activity at various topological

areas. Because of their different neurophysiological origin of evoked and induced activity, the analysis of both types of signal waveforms is useful in the analysis of event related recordings [1].

In general the approach that is frequently used is to use Independent component analysis in order to decompose the EEG signals into spatially fixed Independent Components, which reflect the activity of physiological sources. Then, by means of time-frequency transforms and the activity measures presented earlier evaluate their time and frequency activation, especially on particular frequency bands of interest (mainly delta, theta and alpha) in order to choose a subset of interesting components. This approach increases the amount of information that has to be processed for the analysis by splitting the signal into a number of independent components that need to be evaluated separately and by transforming the signals in the time-frequency domain which enhances the information content.

In the above context it is quite important to provide a mean for efficient analysis of the overall EEG content over all trials. We propose the use of PCA decomposition of the TF maps for a detailed analysis of the content of EEG channels or ICA components. It is most appropriate for the analysis of summarized information content of the EEG over all channels and trials. Principal Component Analysis decomposes the energy content of the entire set of EEG signals into orthogonal, spatially localized components, which are consistently induced by all channels. In this form, the PCA decomposition acts as a well established data reduction scheme in order to extract the major characteristics composing the entire ERP data, from the wealth of information embedded in its multichannel TF representation. The main advantage of using PCA over other techniques is that since PCA essentially analyzes variance it the most effective technique for summarizing and simplifying the captured time-frequency content into as few components as possible. Our goal is to decompose the time frequency content across all channels into a few, easily interpretable components in order to summarize important activations. In contrast with techniques as ICA which would decompose the signal into many components with rather equal variance, PCA is useful for compressing the information into a few components which would represent the activations embedded in the channels.

The PCA approach used here is a general data reduction technique for TF signal representations. The PCA method employed here was recently developed[14], offering a data driven method for decomposing a dataset of TF surfaces. The application of PCA to time–frequency energy is much the same as its application to signals specified in the time or frequency domain. Each time–frequency surface is rearranged into a vector, recasting the time–frequency energy into concatenated time segments each of different frequency content. In this form, the PCA data is formulated into a matrix of trials in rows and different points of activity (different

time–frequency point) in columns. This arrangement is still amendable to decomposition, since PCA makes no assumption about the ordering of the columns for decomposition.

Overall, starting from the time-frequency surface of each channel, we form a three-dimensional matrix of channel \times time \times frequency. Then, we concatenate the time and frequency dimensions into a single dimension, obtaining the representation of the two-dimensional data matrix \mathbf{X} (in the dimensions of: channels/ components \times time–frequency). The decomposition can be applied in the original channels or in channels that are reconstructed by projection of selected Independent components. The PCA analysis is performed on this domain, resulting in the principal-components matrix \mathbf{S} . Finally, by folding this matrix back to three dimensions, we obtain the time–frequency surfaces of the principal components. The number of principal components can be decided in terms of the singular values of the decomposition.

Toward this direction, we created a dataset consisting of five sources, each sampled at 1024Hz, which are mixed to only four channels using a 5×5 mixing matrix. The mixing weights for each channel were calculated as to reflect sources arriving from different origins (different topographies). The first four sources simulate signal peaks at different time locations and at 2, 5, 8 and 12HZ, respectively, whereas the fifth source simulates noise with ongoing EEG power spectrum. The TF energy maps of the four mixed channels are depicted in the first row of figure 1. Following PCA decomposition of the TF maps, the resulting principal components are depicted in figure 1 (second row), where we observe that the information content cannot be efficiently unmixed; the principal components form a mixture of the different sources in the TF surface. In the sequel, we apply ICA decomposition on the dataset. The independent components can separate the EEG-like noise, but the other components are mixture of the initial sources. By removing the noise-like component and back-projecting the remaining components to the channels, we obtain a filtered dataset, whose TF maps are depicted in the third row of figure 1. Despite the remaining effects, the channels appear as much simpler mixtures. Applying PCA decomposition on these TF surfaces provides the results of figure 1 (last row), which separate well each single source utilized in the mixture. The color-maps for all representations range from minimum to maximum values individually for each component; the actual values of color-bars are not important, since we only consider the content of each component and do not compare components themselves. The way the colormap is arranged depends in the application at hand. We can use a global range for all components based in the minimum/ maximum values of the subject or group or we can use a colormap per component. The former enables comparison between components while the later focuses only in the content of each component. When applied in standardized measures we generally select to

present the colormap from the minimum to maximum value of the measure. Otherwise, in case another colormap scheme is selected it will be noted.

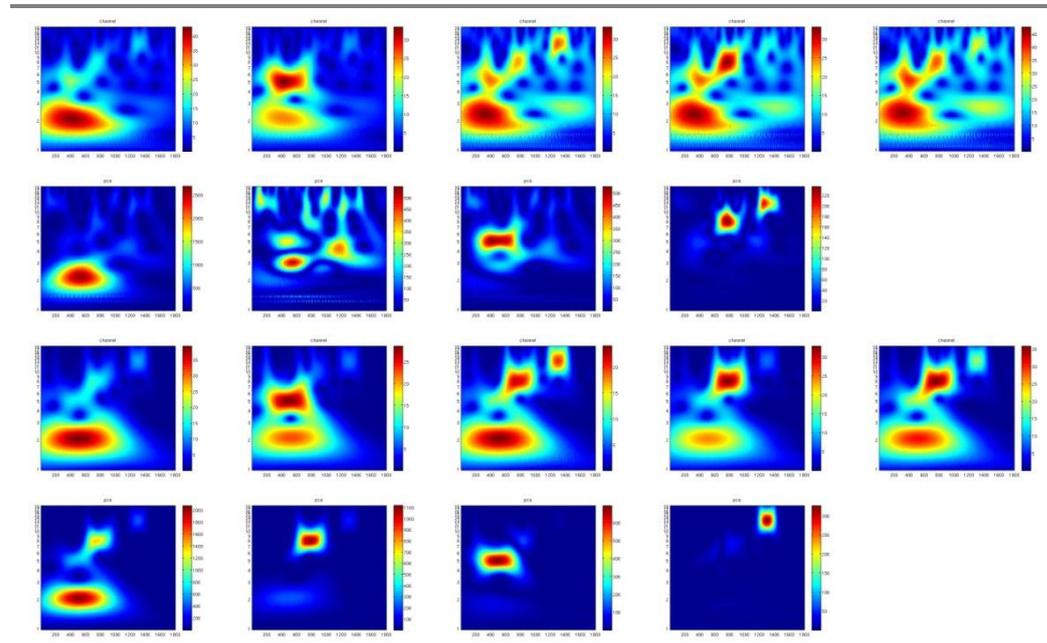
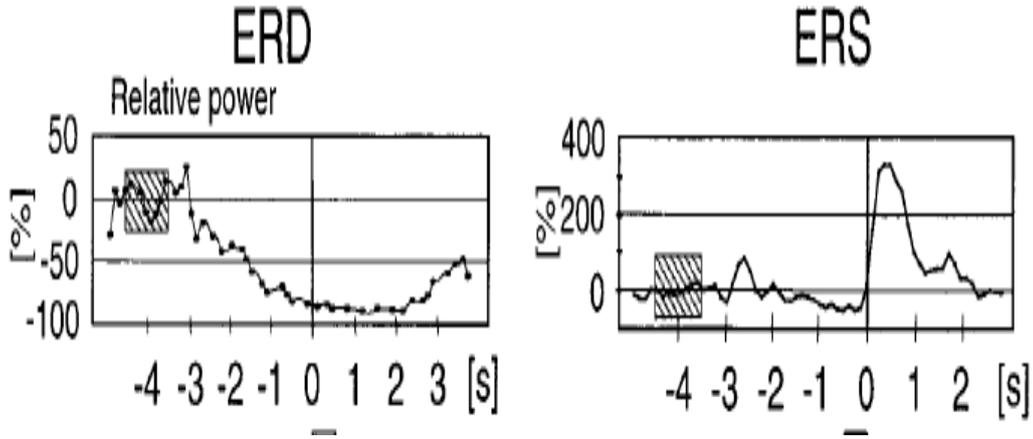


Figure 3.2: Time frequency analysis of the simulated data. 2a) First row: TF energy maps of the five original channels; 2b) second row: four PCA components of original TF energy maps 2c) third row: TF energy maps of the five filtered channels; 2d) fourth row: four PCA components of filtered TF maps.

3.3 Measures of evoked and induced activity

Many efforts have been aimed at the characterization of the generating mechanisms of ERP. In particular, [40] considers the theoretical evaluation of the involvement of various frequencies in ERP and studies the functional meaning of delta, theta and alpha oscillations in terms of frequency, amplitude and phase. It studies the event related changes in average energy revealing the ERD/ERS effects and the phase locking index of evoked delta/theta and induced alpha oscillations. [41] introduces statistical t-like tests for evaluating mean amplitude and energy changes attributed to the event. It considers simple tests for the presence of evoked and induced activity and other promising tests for the identification of phase resetting. Nevertheless, the power of various statistical tests for amplitude, phase or energy significance has been questioned in terms of the ability to characterize the nature of event related activity [2].



3.3: Example of ERD/ ERS calculation. Average power in pre-stimulus is selected (gray area) and subtracted from post-stimulus power. The increase or decrease in power is expressed as a percentage.

The similarity or consistency of components across trials has been initially addressed in ERP studies through the average signal across trials or the spectral energy of the intertrial average. We define the consistency measures on the TF representation of a component derived using the Continuous Wavelet Transform (CWT). Let $x_i[n]$ represent the i -th trial of the component with time spacing δt and duration N data points ($n=1\dots N$). The CWT is defined as the convolution of $x_i[n]$ with consecutive scaled versions of a wavelet function $\psi(n)$, where the concept of scale is used as an alternative to frequency [37]. In our analysis we use the normalized complex Morlet wavelet, which has been extensively studied in EEG analysis [19].

We first test a normalized version of the spectral energy of the intertrial average, referred to as spectral energy (SE) measure. For the TF representation of a component $X_i[k, n]$, with k and n indicating the frequency and time ticks,

respectively, the SE measure is defined as: $c_{SE}[k, n] = \frac{\left[\sum_i X_i[k, n] \right]^2}{\max_{k, n} \left[\sum_i X_i[k, n] \right]^2} \leq 1$. This

measure emphasizes the evoked response of phase locked nature, which increases the energy of the average signal without necessarily increasing the amplitude or energy of individual trials. It can be plotted as a two-dimensional TF map, indicating the temporal and spectral areas of component activation.

Through its definition, the SE measure is normalized in each subject. Quantitative and qualitative comparisons among bands of the same subject are still valid, but only qualitative comparisons can be performed among subjects. Thus, one can compare the bands of maximum activation in two subjects (or populations), but cannot compare the levels of activation, since they are normalized differently for each subject. In order to also enable the direct energy comparison among subjects we can

utilize the direct SE measure as $c_{dSE}[k,n] = \left[\sum_i X_i[k,n] \right]^2$. In the experimental

results we clearly refer to this direct SE definition, whenever required. Furthermore, an overall value c_{SE}^b can be defined as a global index for the spectral energy within a specific band b as the average in time and scale of $c_{SE}[k,n]$, i.e

$$c_{SE}^b = \frac{1}{N} \frac{1}{N_b} \sum_{n=1}^N \sum_{k=1}^{N_b} c_{SE}[k,n] \leq 1, \text{ where } N \text{ denotes the length of the signal and } N_b \text{ the}$$

number of frequency ticks of interest.

We exploit yet another approach for component consistency across trials, which consider phase alignment of each component instead of its energy activation over trials. For phase locked synchronization we use a variant of the intertrial coherence (ITC) measure, which is originally defined as

$$ITC[k,n] = \left| \frac{1}{T} \sum_i \frac{X_i[k,n]}{|X_i[k,n]|} \right| \quad \text{Eq. 3}$$

with T being the total number of trials [12, 32]. On the other hand, for non-phase locked activity we introduce the so-called phase-shift intertrial coherence. The former is based on the phase similarity of phase-locked components, whereas the latter is based on the energy similarity of same-structure but not phase-locked components across trials. Notice that the latter is an extension of the energy measure used in ERD/ERS detection [9], based only on the response after the event. The intertrial coherence measures of components are defined on their TF decompositions, deriving two maps complementary to that of the SE map, which has been often used in component analysis [32]. Examples of measures and their trial-synchronization maps are presented in Figure 3.4.

3.3.1 Event Related Synchronization and Desynchronization.

Event-related desynchronization (ERD) and event-related synchronization (ERS) is the change of signal's power occurring in a given band, relative to a reference interval, in this experiment chosen between 4.5 and 3.5 seconds before the movement onset. Movement-reactive frequency components were detected by comparing power spectral densities in the reference interval (-4.5 s to -3.5 s) and in a 1 s-period around the movement (-0.5 to +0.5 s). Frequency bands, showing significant differences between the reference interval and the activity interval, were chosen for ERD/ERS computations. The EEG was digitally band pass filtered in the selected bands (alpha band 10-12 Hz; beta band 14-18 Hz; gamma band 36-40 Hz) and power values were computed by squaring the samples and averaging these power samples across trials. To reduce the number of power values (original signal was sampled 128Hz), 16 consecutive samples were averaged producing an 8 Hz sampling rate in the resulting power time series. Absolute band power was

converted to percentage power using the reference interval as 100%. By convention, an ERD corresponds to power decrease and an ERS to a power increase[42]

3.3.2 Phase Intertrial Coherence (PIC)

For the i -th trial, the phase shift is reflected as an exponential phase term in the TF representation. We define the metric referred to as Phase Intertrial Coherence (PIC):

$$c_{PIC}[k, n] = \frac{\left| \sum_i X_i[k, n] \right|}{\sum_i |X_i[k, n]|} \leq 1 \quad \text{Eq. 4}$$

with equality holding if and only if the trials resonate in phase. Thus, if all trials “ i ” at a certain frequency band are phase-locked to the same phase shift, then the PIC measure will have large values at the corresponding frequency range. Notice that

$$c_{PIC}[k, n] = \left| \sum_i \frac{X_i[k, n]}{|X_i[k, n]|} \frac{|X_i[k, n]|}{\sum_i |X_i[k, n]|} \right|, \text{ where the first part relates to the ITC}$$

measure. In this form, PIC measures the uniformity of distribution of angles (first part) weighted by the relative amplitudes (second part), in contrast to the ITC that measures the uniformity of pure angles. Based on the study performed in [41], the ITC measure alters the distribution of TF wavelet coefficients ($X_i[k, n]$) from multiple trials in the complex plain, projecting them all to the unit circle with no respect to amplitude. Alternatively the PIC measure preserves the structure of the cloud of coefficients (amplitude and phase) performing just scaling, so that it measures uniformity on a mixed product term involving both the angle and amplitude of coefficients. Certain trials of little amplitude in a frequency band affect the ITC measure exactly the same as trials with significant activity (amplitude), but this is not true for PIC which is proportionally affected by the amplitude of phase-locked trials. Thus, PIC is less sensitive to intertrial variations and forms a more stable measure than ITC for the discrimination of phase-locked oscillatory activity in ERP.

3.3.3 Phase-shift Intertrial coherence (PsIC)

In case of the same basic signal with different shifts from trial to trial, there is a different exponential (phase) term remaining in the TF wavelet coefficients of each trial. Similar to metrics defined for ERD/ERS activity [6, 32], the proposed PsIC metric in non phase-locked responses utilizes the energy of activity instead of the signal’s value. For phase-shift responses, this metric eliminates the complex phase effects and compares the intertrial content of the signal based only on its energy in specific frequency bands. Along these lines, we define the so-called phase-shift intertrial coherence (PsIC) metric based on the energy of the wavelet transform averaged over trials and normalized to one, as:

$$c_{PsIC}[k, n] = \frac{\left[\sum_i |X_i[k, n]|^2 \right]}{\max_{k,n} \sum_i |X_i[k, n]|^2} \leq 1. \quad \text{Eq. 5}$$

Notice that we do not perform a direct evaluation of the ERD/ERS effects with our measure, as we are interested in evaluating the phase or non phase-locked nature of oscillations and not the changes in energy with respect to a baseline or pre-stimulus period. One point of caution related to this definition of PsIC is that, due to the low temporal resolution at low frequencies (delta band), the wavelet coefficients present little variation in amplitude, so that the specificity of the measure decreases. Hence, the reliability of this measure at very low frequencies must be handled with care, similar to the cone of interest for the energy analysis of WT coefficients [43].

3.3.4 Intertrial Power Entropy

The Shannon entropy [8] gives a useful criterion for analyzing and comparing probability distributions. It provides a measure of the information of any distribution. We define the intertrial power entropy as:

$$S_{IE}[k, n] = - \sum_{i=1}^J p_i[k, n] \log_2 p_i[k, n] \quad \text{Eq. 6}$$

Where $p_i[k, n]$ is the probability of trial power for each frequency k time n and J is the number of bins used in the histogram. Probability is calculated from the histogram of power in trials for each k and n using as many bins as trials. The intertrial entropy reflects the order/ disorder of the power levels for each frequency among trials. Equal levels of power among trials would be the equivalent of an ordered state while random power fluctuation from trial to trial can be perceived as a state of disorder. Low values of intertrial entropy reflect the same activation across trials, while high values reflect variability in the power or latency of the activation across trials.

It is apparent that differences between trials are natural to arise in terms of power. Since the brain is always working, receiving information from different sources, we can claim that the recordings entail an inherent level of entropy. Thus, the value of entropy for each time and frequency tick is difficult to be evaluated as an absolute value as is. We define the baseline probability distribution, reflecting the natural disorder in the power between trials, as the mean entropy in the prestimulus time interval. Base in this we define the relative entropy or Kullback - Leibler divergence between trials as

$$S_{RIE}[k, n] = \sum_{j=1}^{trials} p_j[k, n] \log_2 \left(\frac{p_j[k, n]}{q_j[k, n]} \right) \quad \text{Eq. 7}$$

where $q_j[k, n]$ is the baseline power distribution. Relative intertrial entropy gives a measure of similarity between two distributions, between $p_j[k, n]$ in respect with

$q_j[k, n]$. Relative intertrial entropy is positive and zero only when $p_j[k, n] \equiv q[k, n]$. This measure can be used to evaluate divergence of the power distribution post stimulus compared to the power distribution prestimulus. Activations with significant difference to the prestimulus distribution can be identified. In order to evaluate whether this difference refers to an increase or decrease of entropy relative to prestimulus we use the change of intertrial entropy as:

$$S_{IEc}[k, n] = \frac{S_{IE}^{(post)} - S_{IE}^{(pre)}}{S_{IE}^{(pre)}} \quad \text{Eq. 8}$$

Used in parallel with the measure of relative entropy we can assess whereas the activation under consideration presents an increase or decrease in entropy.

3.3.5 Induced and Global Energy

Induced activity is time locked to the event but the exact latency and frequency that this activation occurs in not known a priori. Thus the measures that use the mean power of trials, actually take under consideration the mixed activity of evoked and induced activations. A method to extract the induced activity out of the global energy is to subtract the average ERP out of each trial before applying the measures[33]. This way we can discriminate between power changes attributed directly to induced and those to evoked activations by comparing results from ERP removed and original data.

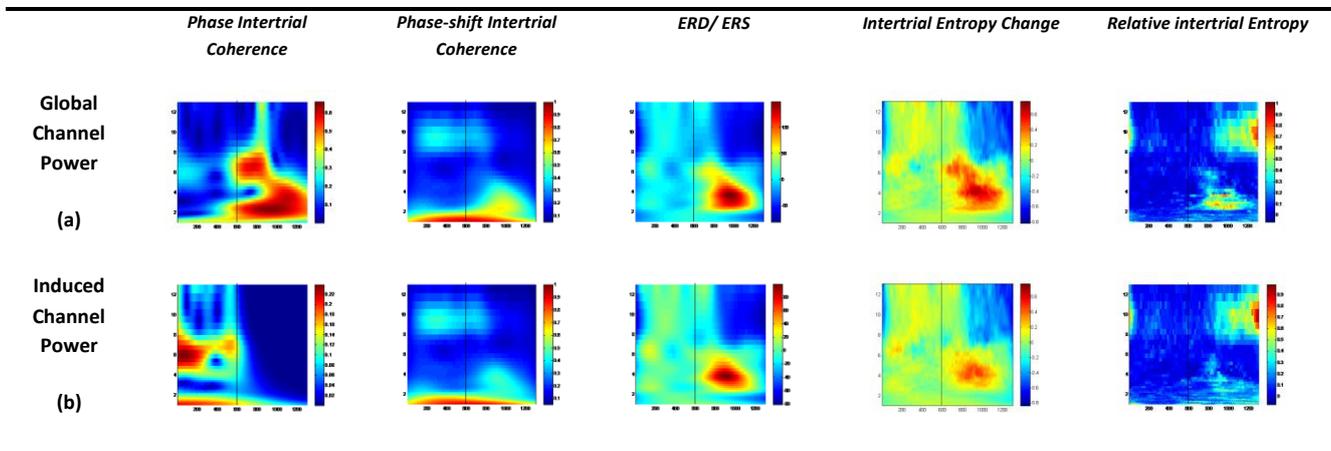


Figure 3.4: Results for measures in Channel Cz. First row (a) displays the measures on the original channel, involving both evoked and induced energy. Second Row (b) displays the results on the processed (ERP-free) data where the average ERP has been subtracted from the post-stimulus.

3.3.6 Application of ERP activity measures.

We applied the measures described in 27-channel recordings from a healthy subject performing an auditory oddball experiment. The subject comes from a dataset that we will describe in detail later.

The time-frequency transform for the processed data as well the original was computed and each of the measures was calculated. Figure 3.4 presents the results on channel CZ for the considered dataset. The latency of the event is presented with

black line. We can observe that the data present increased phase locking activity in delta and theta bands following the event. Alpha band presents ERD starting 200ms after the event. Notice the weak phase locking in alpha band at the same latency as the ERD onset. PsIC measure shows that delta and lower theta bands present significantly coherent activity in all trials. PsIC measure accounts for the global power and cannot be associated with evoked or induced activity, as they overlap in time and frequency (Fig. 1a and Fig. 2b). The same applies for the ERS that occurs 300ms (P300) after the event at theta (mainly) and upper delta band. We cannot distinctly associate the power increase with an evoked activation.

During the ERD in alpha band, the relative entropy shows significant divergence from the prestimulus distribution which can be identified as an entropy decrease (80% decrease). This finding indicates that the alpha ERD presents small variations in power and latency from trial to trial and is directly modulated by the event onset. On the other hand, the entropy divergence identified in the same region with the ERS, in delta and theta bands, corresponds to an increase (40%) of entropy compared to the baseline. The mean divergence is 0.47. This indicates that this activation is variable from trial to trial and is affected from other factors as fatigue or level of attention.

By linking the above with the results from the ERP-free data we can identify whether the nature of the activity can be identified as evoked or induced[33]. The PIC measure confirms that all evoked activity has been removed. A very interesting finding is that ERS is slightly reduced for the delta and theta bands compared to the original data. This indicates that the increase in power on these bands is mostly related with a non-phase locked (induced) process that takes place at the same time.

Relative entropy shows that the divergence in delta and theta bands is less for the induced activity. More specifically the mean value of relative entropy is 0.16 for the region corresponding to delta and lower theta ERS and corresponds to an increase of 19% compared to baseline. Notice the absence of entropy increase in high theta band for the induced activity. These results indicate that entropy increase is mainly due, to activity of evoked nature.

3.4 Discussion

Application of the described measures provides a way to explore the wealth of information that otherwise would go unnoticed if the well established technique of averaging was applied.

There is an ongoing debate regarding the generative model of ERP activations. The models that have been proposed can be grouped in two categories. The additive model considers the evoked response as completely independent from the ongoing background EEG, whereas the phase reset model suggests a phase re-organization of

ongoing EEG oscillations as the generative mechanism of the ERP. Although this debate has attracted a lot of attention, there is not enough evidence which model is more valid. The evoked model has enabled the extraction of useful information about the neurophysiological origins of ERP, mainly through averaging over trials. Underlying the evoked model approach is the assumption that the event itself gives rise to certain brain processes at fixed latency with the same phase, independently from ongoing EEG. This activation serves the response to the event and then vanishes [4, 35]. Averaging of the single trials and inspection of the characteristics of the resulting ERP waveform results in a significant increase of the signal-to-noise ratio, so that useful information can be easily extracted about the brain process and its characteristic markers. In respect to the methods presented above the evoked model would result in high PIC and PsIC values in the specific time interval along with an increase in power as described by the ERD/S measure. The intertrial entropy would show a significant decrease since the evoked model assumes that the response is identical from trial to trial.

On the other hand, the phase-resetting model assumes that the phase-locked phenomenon is not different from the induced activations that appear to be time-locked but not phase-locked. This model states that the different ongoing oscillations either synchronize in phase or time modulated by the stimulus. The pure phase-resetting model would produce high PIC values in the case of phase-locking and high PsIC values that would be and possibly remain present before and after the phase synchronization. Since the oscillation is already there and did not emerge as response to the event exclusively the ERD/S measure would not display any increase or decrease. The same applies for intertrial entropy; it would not show any difference in the entropy since only the phase of oscillation would have been affected.

Until now there is not a single answer to this ongoing debate. Indications that show the validity of one or the other model have been reported in many cases and each is strong enough to rule the other model out. On top of that there exist indications that maybe both of these models are valid under different circumstances, representing different types of activations. We tried to formulate the representation in the measures of the different assumptions about the generation of the ERP covering the spectrum between the two models. The organized representation can be seen in Figure 3.5. We can see that even though we can distinguish evoked from induced activations through the PIC measure the underlying generative model cannot be distinguished using only these measures. Further research is needed in order to shed light in this complex debate.

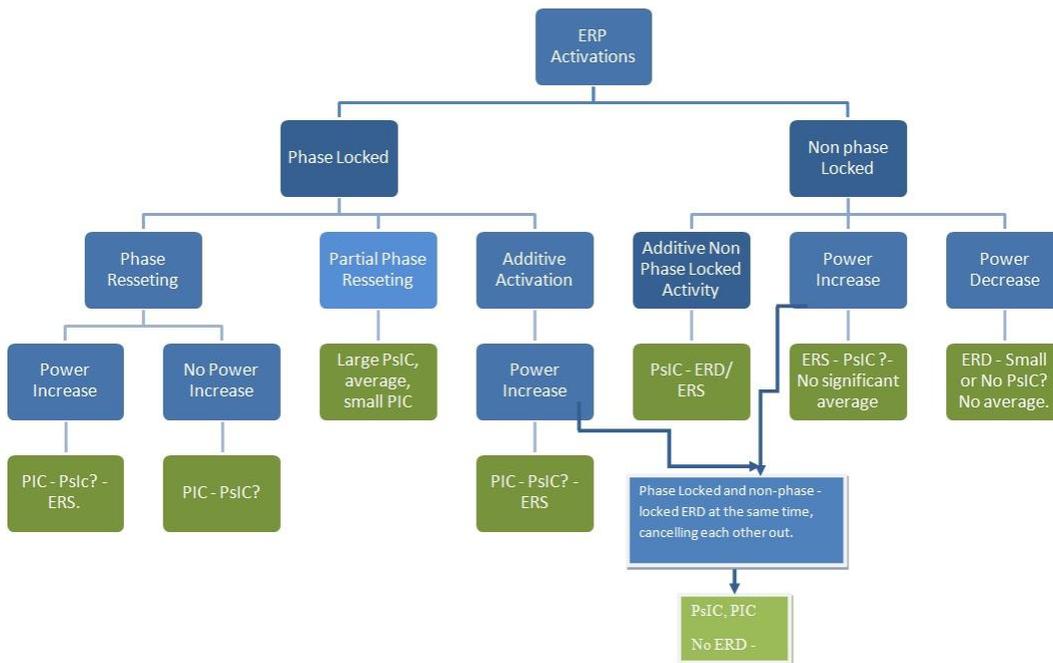


Figure 3.5: Categorization of the different activations that can take place under the different assumptions of the generative model of the ERP.

Due to the nature and complexity of EEG data, it is important to decouple the mixed brain activations, in order to take advantage of the described methods in their full extend. Independent component analysis, briefly introduced above, is a technique which has gained a lot of attention lately, due to its ability to efficiently separate the mixed EEG signals into interpretable, physiologically valid independent sources. In the next chapter we will present in depth the ICA algorithm and the possible pitfalls that introduces to our analysis.

Chapter 4. Independent component analysis application in EEG data.

4.1 Independent Component Analysis

Independent component analysis (ICA) is a statistical signal processing technique which tries to solve the so called blind source separation problem (BSS). In the BSS problem the goal is to recover/ estimate the original sources, by using only the information available from the observed mixed signals, with unknown mixing and no information(or at least minimal) on the sources[44]. Any approach used to solve this kind of problem has to make some assumptions about the source signals, the mixing process or both.

In general, the classic blind source separation model where we observe k signals which are generated by the linear mixture of m source signals can be described as

$$x(t) = A * s(t) + n(t) \quad \text{Eq. 9}$$

where $n(t)$ is considered noise. Noise can be inserted to the observations through limitations of the measuring instruments or due to the fact that the model does not explain the generation of data, adequately.

The blind source separation problem consists in estimating the unknown mixing matrix A and the source signals s using only the observed data x .

One of the most widely used methods for solving the BSS problem is Independent Component analysis[44]. It has different implementations, all of them based on the same basic assumption: that the source signals are statistically independent. Its wide adoption is based in the fact that such an assumption seems natural in many applications, especially when applied in biological signals. Another great advantage of the method is that in practice the assumption of independence is not necessary to hold exactly (strict) in order for ICA to work.

4.2 ICA model.

If we assume the random observed vector $X = [X_1, X_2, X_3, \dots, X_m]$ where each element is a linear mixture of m independent sources of a random vector $S = [S_1, S_2, S_3, \dots, S_m]$ given by the linear equation

$$X = AS \quad \text{Eq. 10}$$

where A is the $m \times m$ mixing matrix. Each row of S holds one independent component, while each column a_k of the mixing matrix A holds the corresponding mixing weights, for all of the k sources.

An ICA algorithm searches for an inverse mapping from X so that $\mathbf{S}' = \mathbf{W}\mathbf{X}$ will be the best possible approximation of \mathbf{S} by minimizing some form of independence criterion. In order to be able to estimate the sources \mathbf{S} the following assumptions must be met:

- The sources S_i are statistical independent between each other and they do not have Gaussian distribution.
- The mixing matrix must be square and full rank. This means that the number of sensors must be as many as the number of sources, at least.
- At most one of the sources can have a Gaussian probability density function.
- The data are assumed to be zero mean and with unit variance. Before applying ICA the observed data are transformed into zero mean and unit variance vectors.

It is apparent that, since the mixing matrix \mathbf{A} and the initial sources \mathbf{S} are unknown, the exact signs and scaling of the sources cannot be identified. Multiplying a mixing vector A_k and divide the source vector S_k by any coefficient, the result is the same. In addition it is not possible to determine the order of the components. These ubiquities of the ICA model, are not so crucial and do not pose problems in most applications.

4.3 Statistical Independence

ICA is based in the notion of statistical independence. In theory, statistical independence means that sources do not contain any information on each other. This means that the joint probability density function (pdf) of the sources is equal to the product of its marginal probability densities. Let assume $s_1, s_2, s_3, \dots, s_m$ be random variables, then the variables s_i are mutually independent if

$$p(s_1, s_2, \dots, s_i) = \prod_i p(s_i) \quad \text{Eq. 11}$$

Sometimes, uncorrelated variables can be confusingly called linearly independent, but that only means that their covariance is zero:

$$E\{s_i, s_j\} - E\{s_i\}E\{s_j\} = 0 \quad \text{Eq. 12}$$

We can see that statistical independence is a stronger restriction than uncorrelation, since it applies for any arbitrary function h_1, h_2 sot that

$$E\{h_1(s_i), h_2(s_j)\} = E\{h_1(s_i)\}E\{h_2(s_j)\} \quad \text{Eq. 13}$$

A direct result from equations (4) and (5) is that independence implies uncorrelation. A common procedure is transforming the data into uncorrelated vectors with zero mean and unit variance before applying ICA, also known as whitening. This

procedure will not alter the resulting independent components, but will constrain the free parameters of the problem and will provide more accurate estimation.

A strict solution of the ICA problem would require the exact determination of the probability distribution functions (pdf), which in general are not known, the sources have to be estimated by approximating independence with an objective function.

4.4 Estimating Independence.

Since a closed form solution would require exact determination of the pdf's, which is generally not possible, the independence has to be approximated. Objective functions that can help this task can be mutual information or negentropy, that essentially measure how non-Gaussian the estimated sources are.

In order to understand how we can solve the ICA problem, we need to explore the relation between gaussianity and statistical independence. A good starting point is given by the common statistical measures of information theory. First, since mutual information measures the amount of information shared between random variables, it can be used as a natural measure of independence. Mutual information I between k random variables is defined using entropy:

$$I(s_1, \dots, s_k) = \sum_{i=1}^k H(s_i) - H(s) \quad \text{Eq. 14}$$

where the entropy H of a discrete random variable s_i , is defined as:

$$H(s_i) = - \sum_i p(s_i = a_i) \log p(s_i = a_i) \quad \text{Eq. 15}$$

Clearly, mutual information between independent variables should be 0. Thus, estimating the independent components is possible by minimizing the mutual information between them. However, doing this in practice can be computationally difficult. Fortunately mutual information, in turn, is closely tied to negentropy, which basically compares a given density to a Gaussian and is defined as:

$$J(s) = H(s_{gauss}) - H(s) \quad \text{Eq. 16}$$

where s_{gauss} is a Gaussian random variable with the same covariance as s . The important point is that as s is considered to be white, negentropy differs from mutual information only by a constant:

$$I(s_1, \dots, s_k) = C - \sum_{i=1}^k J(s_i) \quad \text{Eq. 17}$$

Therefore, maximizing negentropy equals minimizing mutual information when estimating independence. Although negentropy is simpler to calculate than mutual information, it is also based on the exact pdfs. Since negentropy is defined as the difference to a Gaussian, we can calculate an approximation directly from the

observations by using measures of non-Gaussianity. Such measures are skewness and kurtosis, the third and fourth order cumulants. Kurtosis is defined as:

$$kurt(s) = E\{s^4\} - 3E\{s^2\}^2 \quad \text{Eq. 18}$$

Another, more intuitive, explanation is offered by the central limit theorem. Basically, it states that the distribution of a mixture of independent and identically distributed random variables tends to be more Gaussian than the original ones. This means that, when the sources are made more non-Gaussian, they must become more unmixed, or independent.

4.5 ICA preprocessing

As we mentioned earlier, before estimating the independent components, the observed data \mathbf{X} can be whitened, that is, the samples can be uncorrelated with their variances one. Whitening is a linear transformation and can be implemented, for example, using principal component analysis (PCA)[45]. The direction \mathbf{V} of the first principal component is defined as the direction that maximizes the variance of the projection $\mathbf{V}^T x$, where x is a column of \mathbf{X} . Generally, the i th principal component is along the direction of highest variance that is orthogonal to the previous $(i - 1)$ components. One way of finding all the principal components is based on the eigenvalue decomposition (EVD) of the covariance matrix of the data, which gives two matrices, \mathbf{D} (\mathbf{D} is a diagonal matrix of the eigenvalues), and \mathbf{E} , the corresponding eigenvectors.

$$E\{xx^T\} = \mathbf{E}\mathbf{D}\mathbf{E}^T \quad \text{Eq. 19}$$

Our goal is to transform our data X into uncorrelated variables Y such that $Y = PX$ and the covariance matrix of Y is a diagonal matrix. We can write:

$$E\{xx^T\} = \mathbf{E}\mathbf{D}\mathbf{E}^T = \mathbf{E}\mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{E}^T \quad \text{Eq. 20}$$

$$E\{xx^T\} = E\{P^{-1}YY^T P^{-T}\} = P^{-1}E\{YY^T\}P^{-T} \quad \text{Eq. 21}$$

We can see that $P = \mathbf{D}^{-1/2}\mathbf{E}^T$ from Eq. 20

Finally, the whitened matrix Y can be written as $Y = \mathbf{D}^{-1/2}\mathbf{E}^T X$.

The directions of the principal components define the uncorrelated basis and the corresponding variances give the scaling for the whitening transformation. As we noted before, whitening the data does not constrain the ICA in any way, since the scaling is ambiguous and independence implies un-correlation. Un-correlation means

that the covariance is the identity, which always holds when the joint pdf is factorisable as shown before.

The end result of the whitening is that the original mixing matrix \mathbf{A} is also transformed in the following way:

$$\hat{A} = D^{-1/2} E^T A \quad \text{Eq. 22}$$

Estimating the independent components from the whitened data matrix \mathbf{X} is easier, since the number of free parameters is reduced. For example, the mixing matrix \mathbf{A} is orthogonal and has unit length (orthonormal), making its inverse easy to calculate. Instead of having to estimate the n^2 parameters that are the elements of the original matrix \mathbf{A} , we only need to estimate the new, orthogonal mixing matrix $\hat{\mathbf{A}}$. An orthogonal matrix contains $n(n - 1)/2$ degrees of freedom. Because whitening is a standard procedure, much simpler than any ICA algorithms, it is a good idea to reduce the complexity of the problem this way. It may also be quite useful to reduce the dimension of the data at the same time as we do the whitening. Then we look at the eigenvalues d_j of $E[\mathbf{x}\mathbf{x}^T]$ and discard those that are too small. This has often the effect of reducing noise. Moreover, dimension reduction prevents overlearning[46] in case there is no adequate sample size.

Figure 4.1 displays the steps of an ICA algorithm. In fig1a in first and second row we display two independent sources with their scatter plot, respectively. We linearly mix these two sources and we can see the corresponding scatter plot in fig1b. We can see how preprocessing with PCA (whitening) of the data helps to solve ICA. We can see that in order to solve find the initial sources we only have to find the correct rotation, if we start with uncorrelated data (fig1c). In the second row of fig1c we can see that the scatter plot defining the distribution is now clearly a rotated version of the original distribution in fig1a. All that is left is the estimation of a single angle that gives the rotation.

PCA is the optimal method for reducing the dimensions of the data, in terms of retained variance. The dimension reduction also reduces noise, that is, improves the signal-to-noise ratio. Additive noise, for example Gaussian, often appears distributed along several directions with a small energy. If the selected eigenvalues are the largest ones, which identify the directions with the highest variances, the dimension reduction is optimal in the sense that it retains as much of the original signal power as possible.

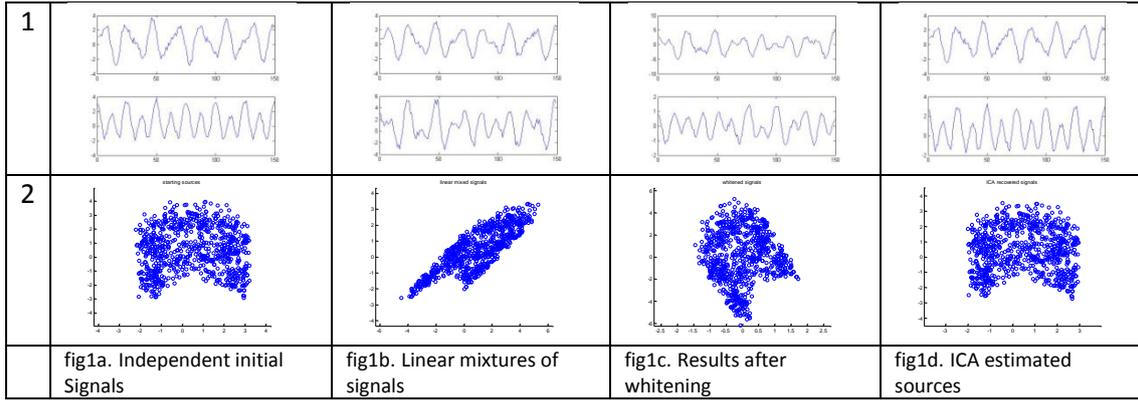


Figure 4.1: Illustration of the whitening effect and the final independent sources in terms of the joint distribution of the two sources (2nd row).

4.6 Infomax Algorithm

Different algorithms have been developed that implement a solution of the ICA problem. There are two main families of ICA algorithms[47]. While some algorithms are based in the minimization of mutual information, others are based in the maximization of non-Gaussianity. For different reasons, algorithm that uses the Infomax principle, in order to minimize the mutual information for solving ICA, is the most popular for EEG analysis. A prominent reason is that it comes bundled with EEGLAB toolbox[7] which facilitates and makes EEG analysis accessible to users who are not familiar with Matlab programming.

The approach used in Infomax algorithm is a generalization of Linsker's infomax principle[48] to nonlinear units with arbitrarily distributed inputs. The principle is that described by Laughlin[49]: when inputs are to be passed through a sigmoid function, maximum information transmission can be achieved when the sloping part of the sigmoid is optimally lined up with the high density parts of the inputs. This can be achieved in an adaptive manner using a stochastic gradient ascent rule. The generalization of this rule to multiple units leads to a system which in maximizing information transfer also reduces the redundancy between the units in the output layer. The approach used was to maximize by stochastic gradient ascent, the joint entropy $H[g(u)]$ of the linear transform squashed by a sigmoidal function, g . When the nonlinear function is the same up to scaling and shifting as the cumulative density functions of the underlying independent components it can be shown[50] that such a nonlinear Infomax procedure also minimizes the mutual information between the u_i exactly what is required for ICA.

The joint entropy at the outputs of a neural network is:

$$H(y_1, \dots, y_n) = H(y_1) + \dots + H(y_n) - I(y_1, \dots, y_n) \quad \text{Eq. 23}$$

where $H(y_n)$ are the marginal entropies of the outputs and $I(y_1, \dots, y_n)$ is their mutual information. Each marginal entropy can be written as:

$$H(y_n) = -E\{\log(p(y_n))\} \quad \text{Eq. 24}$$

The nonlinear mapping between the output density $p(y_n)$ and source estimate density $p(u_n)$ can be described by the absolute value of the derivative with respect to u_i [ref]as:

$$p(y_n) = \frac{p(u_n)}{\left| \frac{\partial y_n}{\partial u_n} \right|} \quad \text{Eq. 25}$$

Taking the derivative of the joint entropy Eq. 23 and rewriting, we obtain:

$$\frac{\partial H(y)}{\partial W} = \frac{\partial}{\partial W}(-I(y)) - \frac{\partial}{\partial W} \sum_{i=1}^n E \left\{ \log \frac{p(u_n)}{\left| \frac{\partial y_n}{\partial u_n} \right|} \right\} \quad \text{Eq. 26}$$

This equation shows a relation between maximizing the joint entropy and minimizing the mutual information between the components at the outputs. A direct minimizations of the mutual information is achieved when $p(y_n) = \left| \frac{\partial y_n}{\partial u_n} \right|$ is satisfied. The mutual information will be minimized when the nonlinearity $y_n = g_n(u_n)$ is the cumulative density function of the source estimates u_n .

The maximum of the joint entropy $H(y)$ can be found by deriving $H(y)$ with respect to W , that is computing the gradient of $H(y)$. The natural gradient rescales the entropy by post-multiplying the entropy gradient by $W^T W$ giving:

$$\Delta W \propto \frac{\partial H(y)}{\partial W} W^T W = (I + \widehat{y}_n u^T) W \quad \text{Eq. 27}$$

Where I denotes the identity matrix and \widehat{y}_n is

$$\widehat{y}_n = \frac{\partial}{\partial y_n} \frac{\partial y_n}{\partial u_n} = \frac{\partial}{\partial u_n} \ln \frac{\partial y_n}{\partial u_n} \quad \text{Eq. 28}$$

Eq. 28 is the original Infomax algorithm.

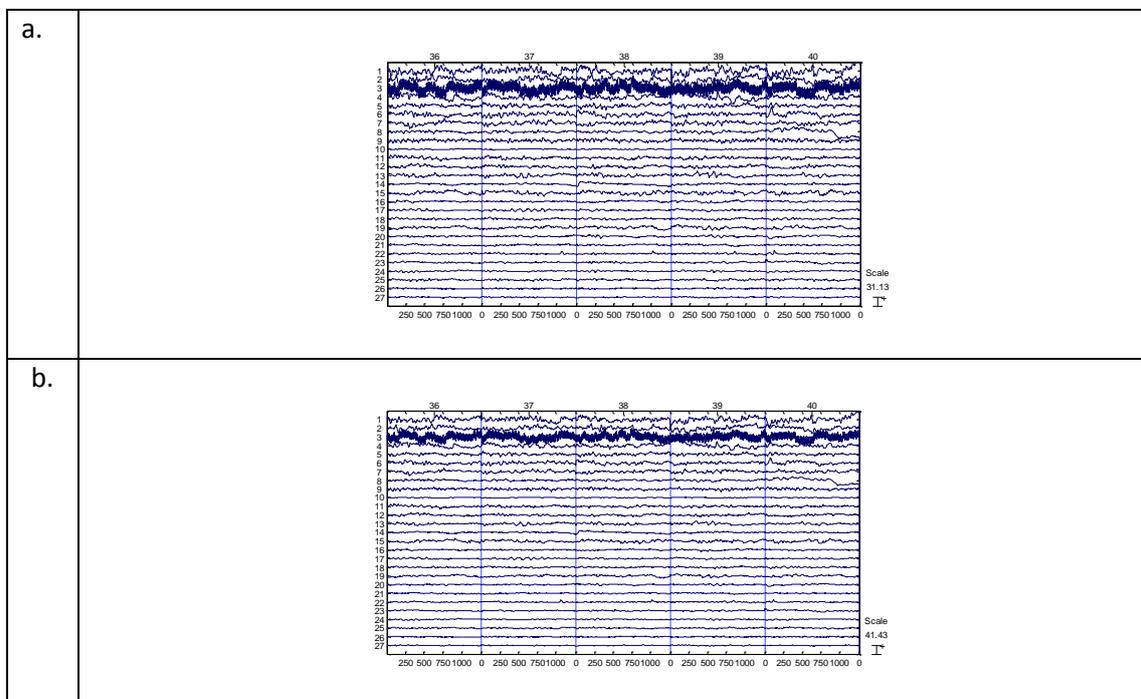
It is clear that with such estimation schemes, the solutions are only approximates and the true solution may not always be found. For example, in addition to the ambiguities in the ICA model, the strict assumption of statistical independence of the sources may not hold for a given data. Still, the model in is noise free, and noise often makes finding the optimal solution harder. Additionally, a problem is that the solutions can be affected by the parameters of the algorithm, for example, the initial conditions. Although a robust algorithm should not be affected by the noise, it can get stuck on a local minimum and even the optimal solutions reached along different paths can be slightly different. Furthermore, since the number of free parameters

can be very high, it is relatively easy for the ICA estimation to overfit the data. Overfitting can cause severe estimation errors.

4.7 Application to EEG data.

Independent component analysis (ICA) is increasing in popularity in the field of biomedical signal processing. It is generally used when it is required to separate measured multi-channel biomedical signals into their constituent underlying components. EEG data consist of a set of multi-channel measurements which are usually recorded using a known and standardized spatial distribution of the recording sensors with respect to the human body, hence giving rise to a set of temporally and spatially correlated measurements. The information inherent in the measurements depends on the specific application domain (which of course influences the number and position of recording sensors or electrodes). The signal(s) of interest is seldom recorded in isolation and is generally mixed with other ongoing 'background' activity and sensor noise, and is almost certainly contaminated by artifacts of either physiological or environmental origins; furthermore, the signal-to-noise (SNR) ratio of the desired signal is generally quite poor.

This automation may be a simple artifact extraction algorithm [51] or the detection of event-related regions of activity in functional magnetic resonance imaging (fMRI) experiments [52]. An example of removing line noise from an EEG recording using ICA is presented in FIGURE.



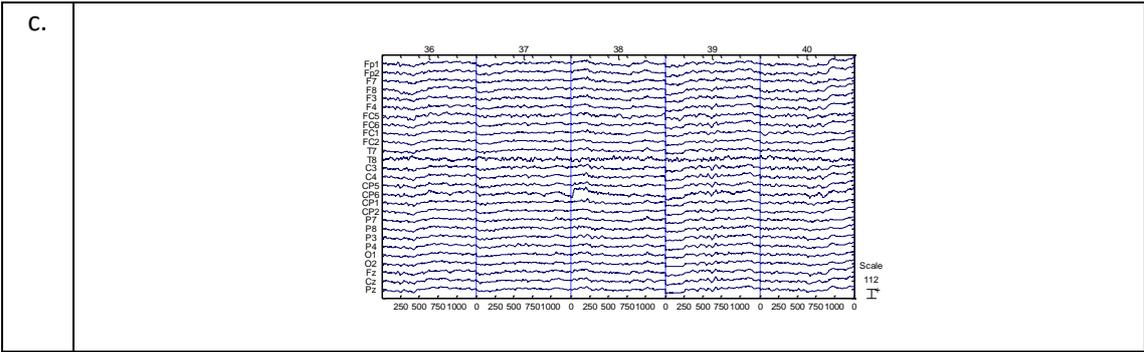


Figure 4.2: Removal of line noise contaminated EEG recording. In row a. the original EEG signal. Note that channel 3 is heavily contaminated with line noise. In row b. the independent components derived from the above figure. Independent component 3 has separated the line noise from the rest of EEG activity. Removing the corresponding IC and back projecting the remaining components, we can see in row c. the resulting noise free EEG signal.

Algorithms that could successfully perform linear ICA appeared in the early to mid 1980s [53] and in the 1990s [54], [55] there were a large number of papers in the literature utilizing ICA for many applications. Application fields of ICA include digital imaging, economic and financial markets and psychometric testing, along with analysis of biomedical signals such as neurophysiologic and cardiac signals as well as analysis of biomedical images such as in fMRI[52]. ICA is also used in feature extraction, where underlying sources are used as a basis to represent the measured data [56]. Most applications of ICA are to multi-channel (or ensemble) time series measurements, in this case the term blind source separation (BSS) can be used as a more generic identifier of the separation process.

One of the biggest advantages in favor of using ICA in brain signal analysis is the fact that multisource activity can be naturally separated into neurophysiologically meaningful components. Standard signal processing techniques such as matched and/or adaptive filters can be used to detect and extract activity of interest, but these generally require much detailed a priori knowledge about the characteristics of each of the signals in question. Furthermore, such techniques are never as discriminative as ICA can be, because there are usually residuals in performing unmixing in this way. ICA also unmixes signals by making very basic assumptions about the data (those of independence foremost) and it makes little difference if the signals are artifactual in origin or brain-signals, for example, for the technique to work.

Let the n EEG channels be arranged as rows of a matrix \mathbf{X} with dimensions $n \times t$, where t denotes the number of signal samples. Independent component analysis performs blind separation of the observed data \mathbf{X} using the restriction that the resulting components arranged in a similar form in a component matrix \mathbf{S} are maximally independent. Alternatively, ICA computes an unmixing matrix \mathbf{W} , which multiplied with the observed data \mathbf{X} results in a matrix \mathbf{S} of independent components. Mapping the weights of \mathbf{W}^{-1} on the electrodes provides scalp topography of the projection of

each component. This presumes that the source locations are spatially fixed and the independent components reveal the time-course activation of each source. In the examples section we utilize the scalp topography of each components in order to infer the brain area of its origin. Another fundamental assumption in ICA decomposition is that the number of sources is the same as the number of electrodes, which is questionable given the wealth of information encoded into the EEG signal. Applying ICA decomposition to few data channels should, thus, result in some or all extracted components being mixture of sources, summing up the activity from more than one neuronal assembly. Even in this case, however, ICA should efficiently arrange for these mixtures to have minimal common or mutual information[57]. Another aspect is the number of samples needed to perform ICA. Since ICA depends in the statistic properties of the data it is apparent that we need enough samples to provide as an accurate estimation of the data statistics as possible. A general rule of thumb is to use more than $2 * (\text{number of channels})^2$. We employ the concatenated trials scheme for ICA decomposition, with the EEG signal extended by one trial following the other and in the same way for each channel, under the assumption of spatially consistent sources,. 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, while it is effective in recovering the inter-trial variability of sources (derived components)[31]. Thus, the content of each ICA component can be subsequently analyzed in several perspectives including its topological origin, the time and frequency distribution of its energy, as well as its coherence over trials.

4.7.1 **Selecting interesting Independent Components**

Selecting independent that capture interesting brain activity is not a trivial task. In application to ERP data two methods have been applied, one that takes under consideration only the evoked nature of the ERP activity and one that uses statistics of both measures that characterize evoked and induced activity.

a) Correlation with average waveform.

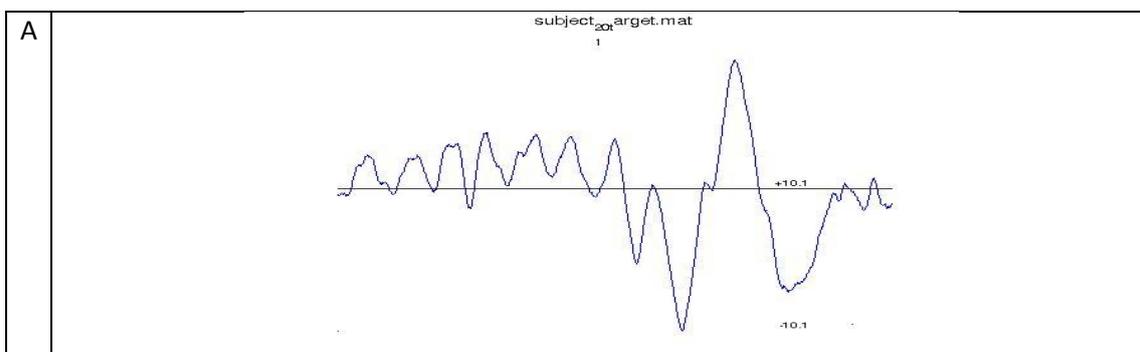
The first and simplest method to extract independent components that contain phase locked activity is by comparing the average independent component across trials with the average signal in a selected channel. This technique is based on the assumption that Independent components that contain a phase locked ERP activity, will resemble to some degree the initial averaged signal. Therefore, each independent component is averaged and the cross-correlation with the average signal at lag 0 is computed. Independent components that have a correlation value above a certain threshold are retained for further visual evaluation. This technique can only detect components that contain phase locked activity that is expressed in the average signal.

This technique provides an intuitive way to select few interesting components based only on the phase locked features of the data. Selected components are further evaluated based on their time activation and their topography. It is apparent that this technique missed a lot of interesting information and could lead to misleading results since the average waveform of the component does not always reflect phase locked activity.

In Figure 4.3 we present an example of the procedure. Figure 4.3A presents the average waveform at channel Cz used as reference. Correlation with this waveform is computed for all components at the time interval of interest, in our case 700 to 1000ms (note that stimulus occurs at 600ms). In Figure 4.3B we present two independent components that present high correlation value with the average, above 0.7. The reconstructed waveform in channel Cz is presented in Figure 4.3C. We can see that the main P300 peak in the data was adequately reconstructed using only the selected components. In Figure 4.4 the normalized spectral energy, phase intertrial coherence and phase-shift intertrial coherence confirms the phase locked nature of these components.

Selecting only phase locked independent components provide only half of the available information hidden in ERP data. Also, this method is based in the average representation of the data which can be misleading. Since the EEG data have a spectrum in which frequency and power are inversely analogous. Therefore the energy of the average waveform is dominated by low frequencies which in turn results in false positive findings by the correlation measure. An independent component with high amplitude peak in a single trial, possibly an artifact, will be present in the final average waveform will provide a high correlation value for the component. In figure we present the time-frequency maps of a component that could be misleading regarded to contain phase locked activity. Although it presents a high amplitude peak in the delta band, no phase-locked activity can be detected.

In order to be select components that represent both kind of activities, evoked and induced we prefer to employ a technique which utilizes all the measures we have defined and we present next.



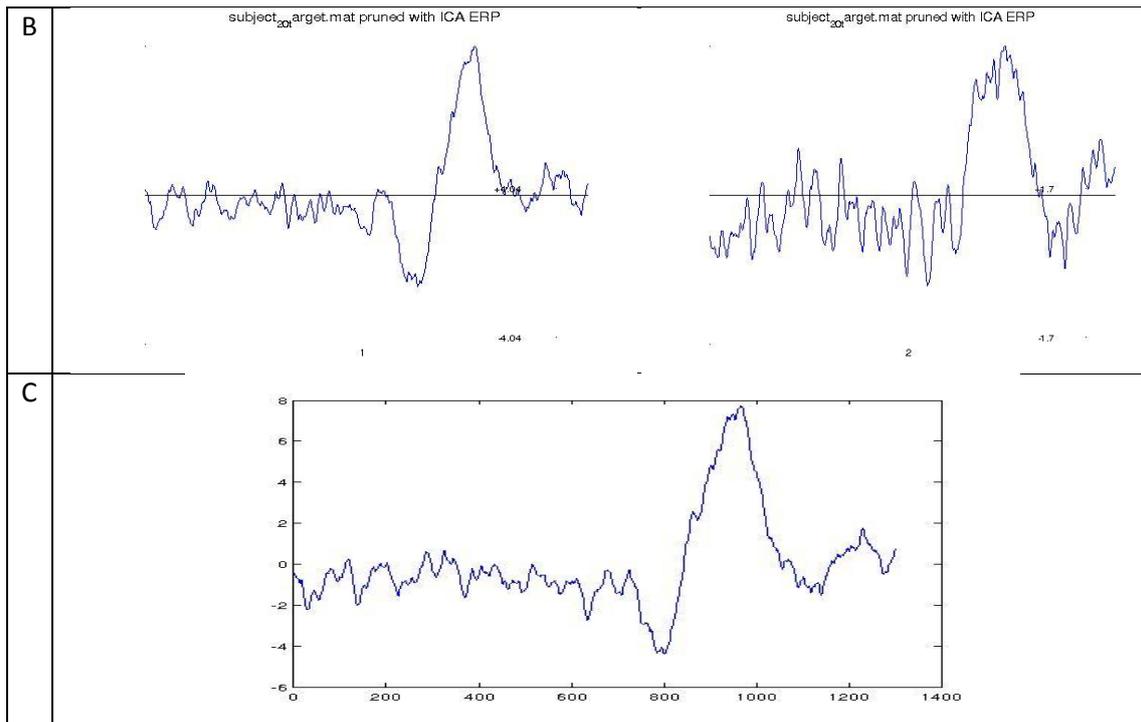


Figure 4.3. Illustration of the components selected using the correlation method. In Row A we can see the average waveform at Channel Cz with which the independent components will be compared. Row B presents two selected independent components with correlation values above 0.7. Row C shows the reconstructed average waveform using the selected components in Channel Cz.

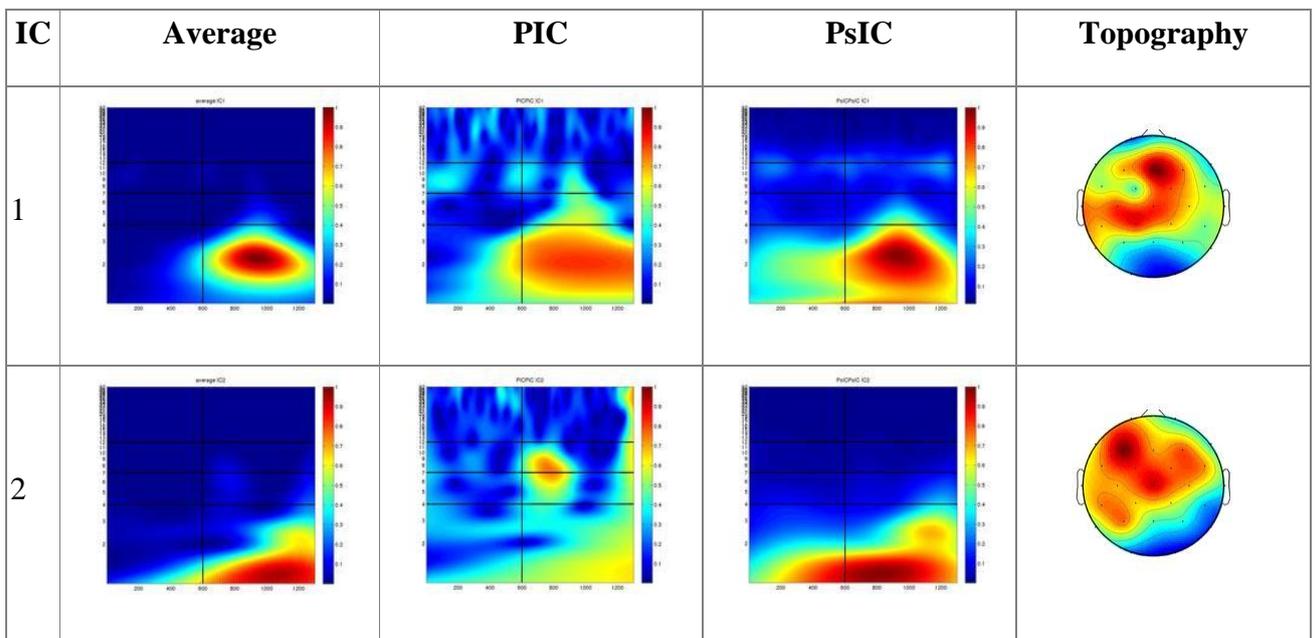


Figure 4.4. Time Frequency maps of the components presented in Figure 4.3. We can see that both contain significant intertrial coherence post-stimulus.

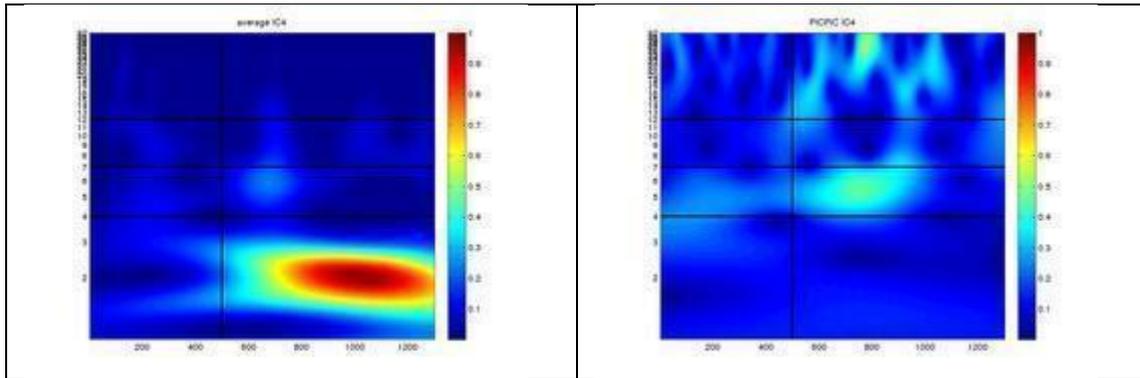


Figure 4.5. Illustration of a component with significant average power in Delta band(left image) but with no phase locked activity as represented in the second figure(right figure).

b) Selecting components based on the intertrial measures.

Using the measure of Phase intertrial coherence and the measure of Phase-shift Intertial coherence we initially calculate the average value of the measures per band, which can be plotted and compared in a single diagram for all derived components. This stage identifies those components that express significant activation lasting for a significant temporal interval, justifying the integration over time. For each component we obtain a single value for each measure and band. Components that present activity above a certain threshold are selected for further visual evaluation.

The PIC maps reflect the phase-locked component over all trials, whereas the PsIC map reflects the non-phase-locked activity, which implies similar structure of the signal but without phasic coherence across trials. For the detailed analysis of selected components, we then consider the two-dimensional TF maps of the three measures that reflect the component consistency (in energy and/or phase) along trials.

Using this technique we are able to initially select interesting components that contain both phase locked and non phase locked activity relevant to the task. We prefer to use a conservative threshold, usually the upper quartile of the data, in order to capture components that present partial phase-locked and non-phase locked activity, that otherwise would be discarded. Then each component is visually evaluated and characterized according to the values of the measures, latency and band of activation and topography. In Figure 4.6 we present the average measures per band of a single subject and interesting components derived from the diagrams for demonstration.

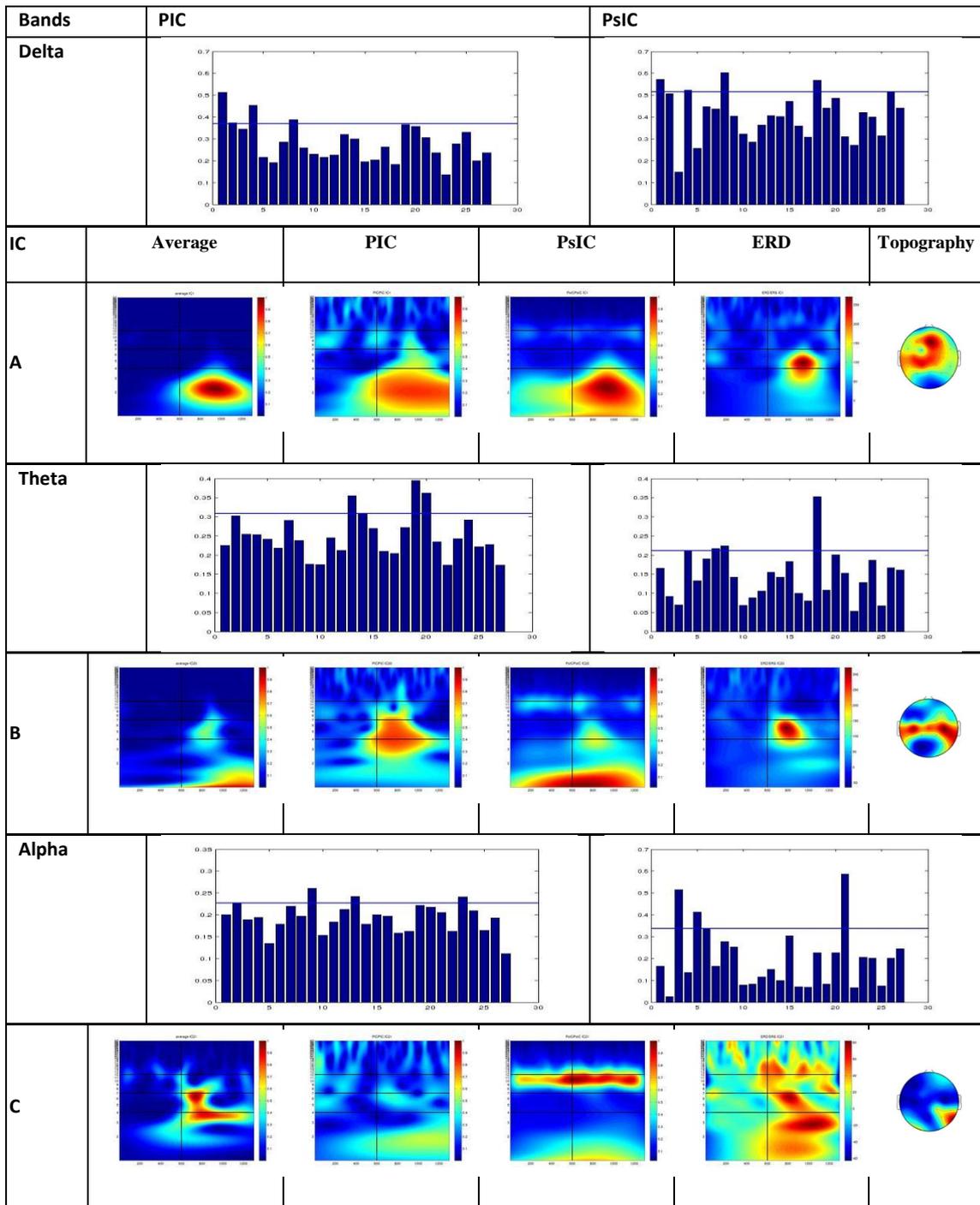


Figure 4.6. Using the measures as global features per band for selecting interesting components. A component from each band is illustrated sequentially demonstrating the visual evaluation using the time frequency maps of the measures.

We picked a phase locked delta band component Figure 4.6A, which shows significant phase locked activity. The peak presented in spectral energy measure is also reflected phase intertrial coherence, meaning that the average peak is of phase-locked nature. The phase-shift intertrial coherence shows significant activity also. The PsIC measure detects coherent oscillations across trials. We cannot distinguish if the oscillation is phase locked to the event or not. In this case the oscillation synchronizes its phase after the event occurrence. Lower theta presents some phase

locking activation and a weak non phase locking one. We can observe an interesting phenomenon, if we take under account the increase in energy in theta band compared to the prestimulus. We can see that only theta displays an increase in amplitude, whereas delta remains unchanged. This indicates that this independent component capture a central-frontal delta oscillation which synchronizes its phase after the event without any increase in amplitude, whereas at the same time an additive partial phase-locked theta emerges.

A phase-locked theta component is presented next Figure 4.6B and we can see this time that it presents clearly a phase locked additive activation which results in theta power increase after the stimulus. The relatively low PsIC value shows that emerge of theta activation is not at the same power level in all trials.

Finally an alpha component Figure 4.6C is illustrated where we can see that average spectral map has completely erased all alpha activation. This is in accordance with PIC measure which shows that alpha activity presents no phase locking. Note that using only the average we could erroneously classify this component as theta. The PsIC measure though, successfully captures the alpha activity which also presents a slight power increase as noted by the ERD/S measure, parietal.

4.8 Discussion

There is an ongoing debate about the nature of phase-locked activity of whether emerges from brain activations additive to the EEG or reorganization of phase of ongoing oscillations. We will further discuss such matters in detail in the results section. Never though, we can see the effectiveness and usefulness of ICA in analyzing such phenomena. Using ICA we can decouple the brain processes that take place during an ERP experiment and extract useful features from the activations of the independent components. As we seen though, application of ICA has its limitations and problems, whether they are algorithmic or due to the dataset and we need the means to characterize the stability of the derived components besides their content. This is a crucial factor in order to be able to confidently analyze and characterize the content of Independent components. We will deal with this matter and provide means to accomplish this in the next chapter.

Chapter 5. Stability Estimation methods

5.1 Introduction

Over the past years, many researchers have used independent component analysis (ICA) methods for blind source separation and feature extraction. Most prominent applications are arguably in the field of biomedical engineering. A major problem in application of ICA is that the stability of the estimated independent components is not known. Since the ICA techniques are stochastic, their results may be somewhat different in different runs of the algorithm. In fact, most algorithms give different components when run multiple times, due to different initial conditions. This is in direct contrast to such methods as principal component analysis, whose results are unique.

Another problem is that as real data never follows the ICA model exactly, the contrast function used in the estimation may have many local minima which are all equally good. Even in the extreme where the data is exactly generated according to the ICA model, the finite sample size induces statistical errors in the estimation [13]. In extreme cases the algorithm may overlearn the data, rendering the results completely invalid. Understanding the causes of the variability and how to overcome the consequences is essential for consistent analysis of EEG using ICA. In this chapter we will discuss and present different methods used to tackle this problem.

5.2 Definition of Stability

The two most important and fundamental characteristics of any measurement procedure are reliability and validity. In this study, we are most interested in the reliability of the results of ICA.

Reliability is defined as the extent to which a questionnaire, test, observation or any measurement procedure produces the same results on repeated trials. In short, it is the stability or consistency of scores over time or across raters. The extent to which they agree on the scores for each contestant is an indication of reliability. Similarly, the degree to which an individual's responses (i.e., their scores) on a survey would stay the same over time is also a sign of reliability.

An important point to understand is that a measure can be perfectly reliable and yet not be valid. Consider a bathroom scale that always weighs you as being 5 lbs. heavier than your true weight. This scale (though invalid as it incorrectly assesses weight) is perfectly reliable as it consistently weighs you as being 5 lbs. heavier than you truly are.

There are four *general classes of reliability estimates*, each of which estimates reliability in a different way. They are:

- **Inter-Rater or Inter-Observer Reliability:** Used to assess the degree to which different raters/observers give consistent estimates of the same phenomenon.
- **Test-Retest Reliability:** Used to assess the consistency of a measure from one time to another.
- **Parallel-Forms Reliability:** Used to assess the consistency of the results of two tests constructed in the same way from the same content domain.
- **Internal Consistency Reliability:** Used to assess the consistency of results across items within a test.

In our case, we are interested in the internal consistency reliability measurement, where we use our single measurement instrument administered to a group of people on one occasion to estimate reliability. In effect we judge the reliability of the instrument by estimating how well the items that reflect the same construct yield similar results. We are looking at how consistent the results are for different items for the same construct within the measure. When it is impractical or inadvisable to administer two tests to the same participants, it is possible to assess the reliability of some measurement procedures by examining their **internal consistency**. This type of reliability assessment is useful with tests that contain a series of items intended to measure the same attribute. In our case we want to evaluate the reliability of ICA on calculating independence in the provided dataset.

5.3 Approaches to stability estimation

As with any statistical method, it is necessary to analyze the statistical reliability of the obtained components since the finite sample size induces statistical errors in the estimation [58].

Such an analysis can be accomplished using bootstrapping, a well-known method for computing the statistical reliability in the case where a simple mathematical formula cannot be found[59]. Bootstrapping is a resampling method; the data sample is randomly changed by simulating the sampling process, and the algorithm is run many times with the bootstrapped samples that are somewhat different from each other. The reliability of the original estimate can then be analyzed by looking at the spread of the obtained estimates. A further problem typical of ICA is that most algorithms have random (stochastic) elements, i.e. the algorithms give somewhat different results at every run of the algorithm. Most ICA algorithms such as FastICA[60] or the natural gradient (infomax) algorithm[55, 61] are based on minimization or maximization of an objective function, such as the likelihood, mutual information. The randomness of the algorithm stems from the fact that it is not

possible to find the point that globally maximizes/ minimizes (for simplicity, we shall refer only to maximization) the objective function.

The basic principle is to start in some initial point, and then make steps in a direction that decreases the objective function, until one finds a point in which the objective function is locally minimized. Depending on the starting conditions, the algorithm may converge in different local minima. In the case of a very high-dimensional signal space, the probability of finding the global minimum may be very small and the situation is even more complicated, since at each local minimum a subset of the components may be accurately estimated. This is plausible because the estimation of the individual components is partly decoupled from each other. In fact, the components can be estimated one at a time by using the criterion of negentropy or nongaussianity[60].

To assess the algorithmic stability of the estimated components, it is reasonable to run the estimation algorithm many times, using different initial values, and assessing which of the components are found in almost every run. Different methods have been proposed that provide reliability estimation for unsupervised learning algorithms, using well-known resampling methods from statistics[59, 62]. The approach is based on estimating a large number of candidate independent components by running an ICA algorithm many times by implementing a resampling stage that generates surrogate data to test the ICA results. Then a clustering stage follows that groups together the estimated parameters. We will discuss the different steps and approaches in detail in the following sections.

5.4 Bootstrapping

In the typical unsupervised learning scenario, we want to learn or estimate a set of parameters from observed data that characterize the process that generated the data. Resampling is a statistical method which gives, e.g., the bias and the variance of estimators only from one set of data at hand by virtue of modern computer power. Among such procedures, the Jackknife and the Bootstrap are most well known (see, e.g., [3] and [4]). The Jackknife produces surrogate data sets by just deleting one datum each time from the original data set. There are generalizations of this approach like the delete - Jackknife which delete more than one datum at a time. The Bootstrap is a more general approach and is widely used in data analysis recently.

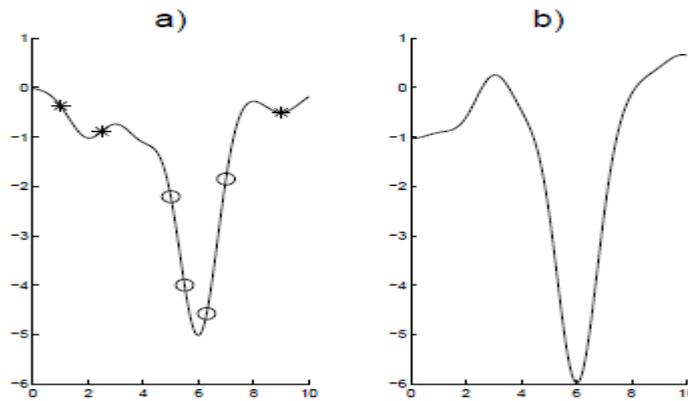


Figure 5.1: Illustration of the problem of local minima[63]. a) An objective function that is to be minimized. There is a global minimum at 6, and small local minima at 2 and 9. If a gradient descent algorithm starts at the points marked by circles, it will probably find the global minimum. However, if it starts at the points marked by stars, it will probably converge the nearest local minimum, at 2 or 9. Running the algorithm many times, it will converge to 6 most of the time. b) When the data is resampled in the bootstrapping method, the optimized function changes a bit. The smaller local minima at 2 and 9 disappear, and a new local minimum appears at 0. The stable minimum at 6 is still a local (and global) minimum.

Bootstrapping is the practice of estimating properties of an estimator (such as its variance) by measuring those properties when sampling from an approximating distribution. One standard choice for an approximating distribution is the empirical distribution of the observed data. In the case where a set of observations can be assumed to be from an independent and identically distributed population, this can be implemented by constructing a number of resamples of the observed dataset (and of equal size to the observed dataset), each of which is obtained by random sampling with replacement from the original dataset[59]. The basic idea of bootstrapping is that the sample we have collected is often the best guess we have as to the shape of the population from which the sample was taken. For instance, a sample of observations with two peaks in its histogram would not be well approximated by a Gaussian or normal bell curve, which has only one peak. Therefore, instead of assuming a mathematical shape (like the normal curve or some other) for the population, we instead use the shape of the sample.

Consider the case in which we get T independent and identically distributed samples $x = [x_1, x_2, \dots, x_T]$ from a distribution F . Suppose we want to evaluate the root mean square error (RMSE) of the estimator. Let \hat{F} be an empirical distribution of the data x . A random variable from \hat{F} takes values x_t , where $t = 1, \dots, T$ with equal probabilities $1/T$. Consider B new surrogate datasets such that $x^b = [x^b_1, \dots, x^b_T]$ where $b = 1, \dots, B$, that are generated by taking T i.i.d random variables x^b_1, \dots, x^b_T from the empirical distribution \hat{F} . We remark that some data points might occur several times, while others may not occur at all in a particular bootstrap

sample. On each surrogate x^b the estimator $\widehat{\theta}_i^b = \widehat{\theta}_i(x^b)$ is calculated so we have B estimators $\widehat{\theta}_i^1, \dots, \widehat{\theta}_i^B$. The bootstrap estimator of the RMSE is calculated as

$$\sigma_i(B) = \sqrt{\frac{1}{B} \sum_{\beta=1}^B (\widehat{\theta}_i - \widehat{\theta}_i^\beta)^2}$$

This quantity measure, how robust our estimation is against small (resampling) changes to the data; in other words, how stable the learning algorithm is with regard to the estimated solution $\widehat{\theta}_i$. This can be extended to other measures beyond RMSE that can be used as measures of reliability. These measures help us to select between different algorithmic solutions or to choose parameters for a single algorithm. Furthermore, an assumption about the data generating model can be accepted or rejected (in the sense of mathematical testing theory).

5.4.1 Using Bootstrap for estimating ICA stability

The point is that the resampling should not affect the global statistical properties of the data, but the solution, defined by the resampled dataset, would be slightly different on each run. The global minimum/ maximum stays relatively stable, but the shape of local minima and noise on the solution space can change freely. Thus, it allows the ICA algorithm to converge to different solutions and even reach solutions otherwise difficult to identify due to surrounding local minima or a high level of noise.

The solutions of multiple runs are reached along different paths based on the initial conditions and the resampled dataset. In theory, when the same solution is reached in more than one run, the estimates would match perfectly. This is often not the case, but the estimates should still form consistent groups with high similarity. Thus, one may expect that the true solution can be found as a mean of the consistent group and that the spread of the group can be used to analyze the reliability of that solution. On the other hand, if the spread is too high, one expects the solution not to be reliable. The mean representatives of the groups can also depart somewhat from the strict assumption of independence, which may actually lead to a more natural decomposition of the data.

Assuming the general model $s = Wx$, ICA runs multiple times n on the data matrix $X = [x_1, x_2, \dots, x_n]$ consisting of n estimates with k channels. The estimates of demixing matrices \widehat{W}_i from each run $i = 1, \dots, n$ are collected into a single matrix $\widehat{W} = [\widehat{W}_1^T \dots \widehat{W}_n^T]^T$.

5.4.2 Grouping Estimates

A natural measure of similarity between the estimated independent components is the absolute value of their mutual correlation coefficients $r_{i,j}, i, j = 1, \dots, K$. First the estimates are normalized to account for the scaling ambiguities of ICA. One good way of doing this is to make the estimates have zero mean and unit variance. If the normalized solutions are defined as \widehat{W} , with columns w_i the normalization equals:

$$w_i = \frac{w_i - \bar{w}_i}{\|w_i - \bar{w}_i\|}$$

where w_i is a column of total matrix A and \bar{w}_i is the mean of the column. We still have to deal with the ubiquities of sign and ordering of the components. For the sign ubiquity we will use the absolute of the correlation coefficient. The final similarity matrix then has the elements σ_{ij} defined by:

$$\sigma_{ij} = |r_{ij}|$$

In order to deal with the random ordering of the components we will have to employ some sort of component clustering.

5.4.3 Clustering the estimates

Different approaches have been used in order to group the estimated components. A specialized software package for Matlab, *Icasso* [63] uses agglomerative hierarchical clustering to cluster the estimates. Agglomerative hierarchical clustering is a well-known method for a modest number of objects [4, 5]. This is a "bottom up" approach; each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy. The tree-like hierarchy (dendrogram) produced by agglomeration is intuitively appealing in the sense that all clusters implied by lower levels of the tree are always subsets of clusters at the higher levels. As a result, the user is able to explore and compare the different level(s) of clustering that are readily computed. The simplest way to obtain a partition of L clusters from a dendrogram is to cut it at level where L clusters are present. There are numerous reviews and studies on the multitude of agglomeration strategies and cluster validity indices, see at [64, 65].

Unfortunately, there is no easy way of selecting the optimal agglomeration strategy for a specific set of data, and the selection must be based on problem specific considerations. The same applies also to selecting a clustering validity index for determining a "natural" number of clusters [65].

Three basic agglomeration strategies that operate directly on the similarity matrix are single-link (SL), complete-link (CL), and group average-link (AL). *Icasso* uses AL as the default choice of agglomeration strategy. This is because, firstly, SL is in general reported to be more sensitive to noise than AL and CL.

In order to assess the interesting clusters, a cluster quality measure is introduced that reflects the compactness and isolation of a cluster. It is computed as the difference between the average intra-cluster similarities and average extra-cluster similarities. Let us denote by C the set of indices of all the estimated components, by C_m the set of indices that belong to the m cluster, and by $|C_m|$ the size of the m -th cluster. The cluster quality index is defined as:

$$I_q(C_m) = \frac{1}{|C_m|^2} \sum_{i,j \in C_m} \sigma_{ij} - \frac{1}{|C_m| |C_{-m}|} \sum_{i \in C_m} \sum_{j \in C_{-m}} \sigma_{ij} \quad \text{Eq. 29}$$

where C_{-m} is the set of indices that do not belong to cluster m . $I_q(C_m)$ equals 1 for an ideal cluster and decreases when C_m becomes less compact and isolated. *Icasso* leaves the number the final selection of clusters to the user, who can experiment with the results from different levels of the dendrogram. It is reasonable to start studying the clustering from the number of clusters L equal to the data dimension k .

We need a quantitative metric for evaluating the optimal number of clusters. *Icasso* shows the R-index (I_R) in the user interface. The index is defined as

$$I_R = \frac{1}{L} \sum_{m=1}^L \frac{S_m^{in}}{S_m^{ex}} \quad \text{Eq. 30}$$

Where $S_m^{in} = \frac{1}{|C_m|^2} \sum_{i,j \in C_m} d_{ij}$ and $S_m^{ex} = \min_{m' \neq m} \frac{1}{|C_m| |C_{m'}|} \sum_{i \in C_m} \sum_{j \in C_{m'}} d_{ij}$. The index looks for compact and well-separated clusters and the minimum of I_R suggests the best partition.

5.4.4 Visualization by nonlinear 2D projection

In order to visually inspect the resulting clusters ICASSO provides a tool that projects the estimates into a 2D surface and plots it as a point on the display. A convex hull bounds the estimates belonging to the same cluster. The projection method used is Curvilinear Component analysis. It is important that in the resulting projection the smaller a convex hull is, the more compact the corresponding cluster is. An ideal cluster should contract into a single point the same cluster. Spatial proximity is one of the strongest visual indicators of grouping[66]. In order to be trustworthy, a projection should be such that one can trust the visual proximity as an indicator of similarity. The trustworthiness index in [67] is a function of the visual neighborhood size, and it must be evaluated for the neighborhood sizes of interest: according to [67] it is especially important that the trustworthiness is retained for small neighborhoods.

Curvilinear Component Analysis (CCA) aims to reproduce the topology of the original data in the projection subspace but without fixing in a static way the configuration of the topology. The nonlinear mapping adapts itself to the shape of the joint

distribution of the data. In the CCA, the topology is defined by the distances between all pairs of vectors of the original data. Since the topology cannot be entirely reproduced in the projection subspace, which has a lower dimension than the original subspace, the local topology, the most important, is favored to the detriment of the global topology. CCA tries to reproduce short distances firstly, long distances being secondary. Formally, this reasoning leads to the following error function (without normalization):

$$E_{CCA} = \sum_{i,j=1}^N (d_{i,j}^n - d_{i,j}^p)^2 F_{\lambda}(d_{i,j}^p)$$

where $d_{i,j}^n$ are the distances between the i-th and j-th vector in the original space and $d_{i,j}^p$ are the distances in the p-dimensional projection space. The factor F is a decreasing function of its argument, so it is used to favor local topology preservation. Given this error function an optimal projection can be computed using a gradient descent algorithm.

5.4.5 Experimental Results

We tested the aforementioned methodology on simulated data and data from a single healthy subject. Our goal is to evaluate ICA under simulated conditions, knowing the sources beforehand, while altering the sample size, the starting conditions and shuffling the data. Then we test the performance of ICA in a real dataset.

5.4.5.1 Application in simulated data

We generated simulated data, resembling a real EEG signal, using the procedure described in[2]. We simulated an event related potential recording, consisting of 40 trials and 27 electrodes. The simulated trials, sampled at 1024Hz and lasting 1256ms, were constructed to match the spectrum of a real EEG recording. Then, two phasic peaks were added to uncorrelated EEG-like noise.

The peaks were created by a half cycle of a sinusoid. We chose two embed two peaks in the EEG like noise. The two peaks and their corresponding energy on the scalp electrodes are illustrated in Figure 5.2 and Figure 5.3. The first one is a negative going half-cycle of a 5Hz sinusoid, which peaked at 215ms. The latency and amplitude was constant across trials. The peak amplitude varied across electrodes in a way to mimic a frontocentral distribution of amplitude. The other peak is a half-cycle 9Hz sinusoid which peaks at 600ms and its amplitude distribution across electrodes can be described as centroparietal.

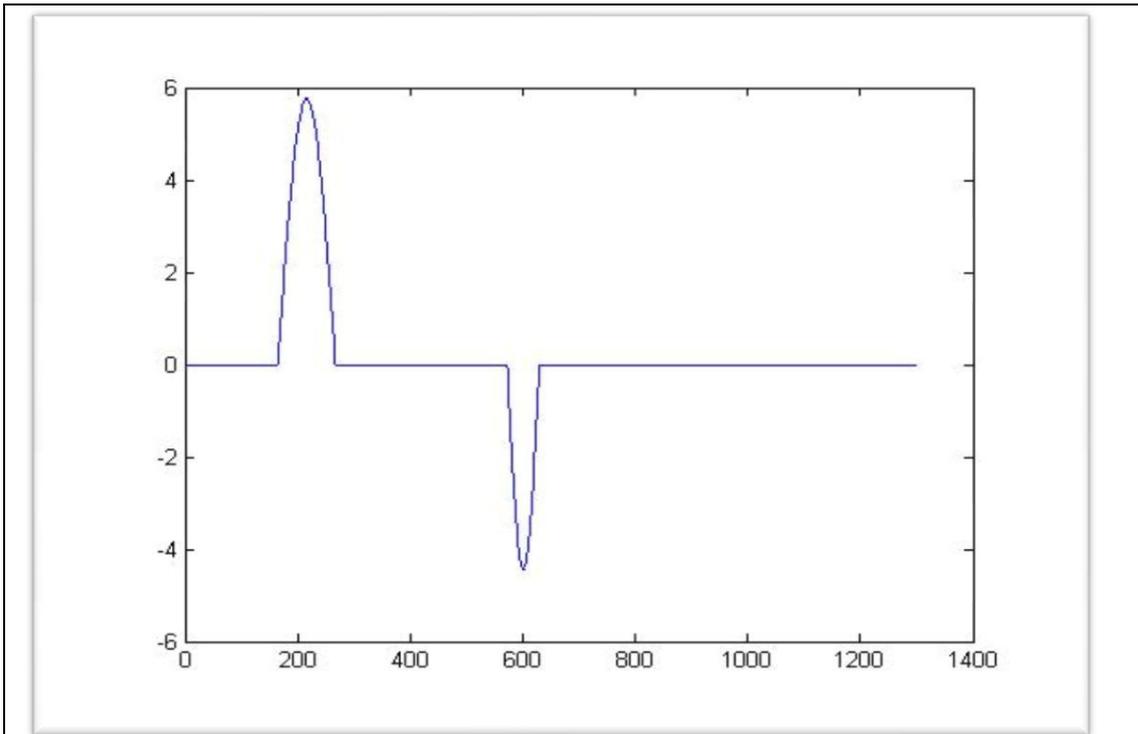


Figure 5.2: The two peaks used as sources for the simulation.

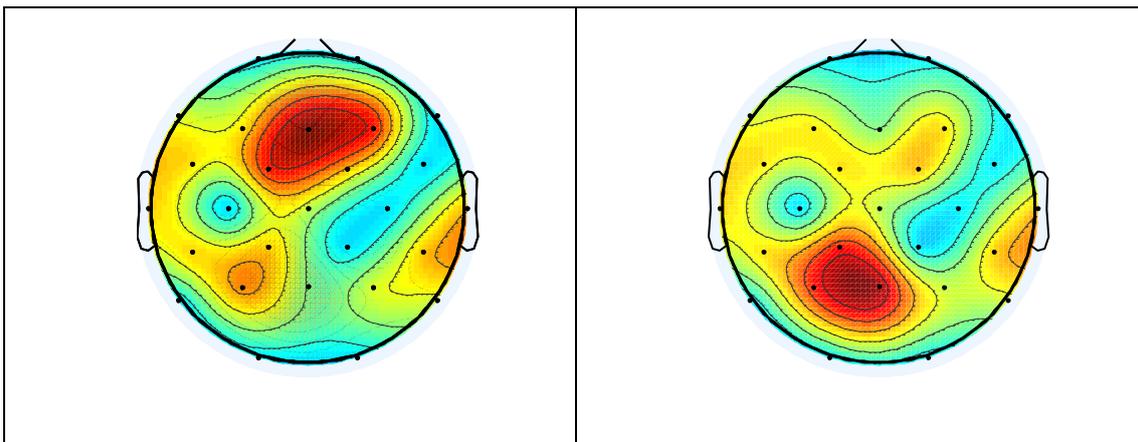


Figure 5.3: Amplitude distribution of the two sources in the electrodes. Peak at 200ms is depicted in left figure while peak at 615ms in the right picture.

The two phasic peaks were added to uncorrelated background EEG noise that was simulated by summing together 50 sinusoids of randomly varying frequency and phase (with different random values of phase and frequency generated for each simulated epoch). The frequencies were chosen to span the range from 0.1 to 125 Hz, the phases varied randomly between 0 and 2π . The maximum amplitude of any single frequency component of the background EEG (at 0.1 Hz) was set to be 20 mV. Within this constraint, the amplitude of the sinusoid at each frequency was scaled to match the power spectrum of the EEG (in this case we used the power spectrum from one healthy subject, which we will use also in the next section). The resulting

EEG signal is illustrated in Figure 5.4.

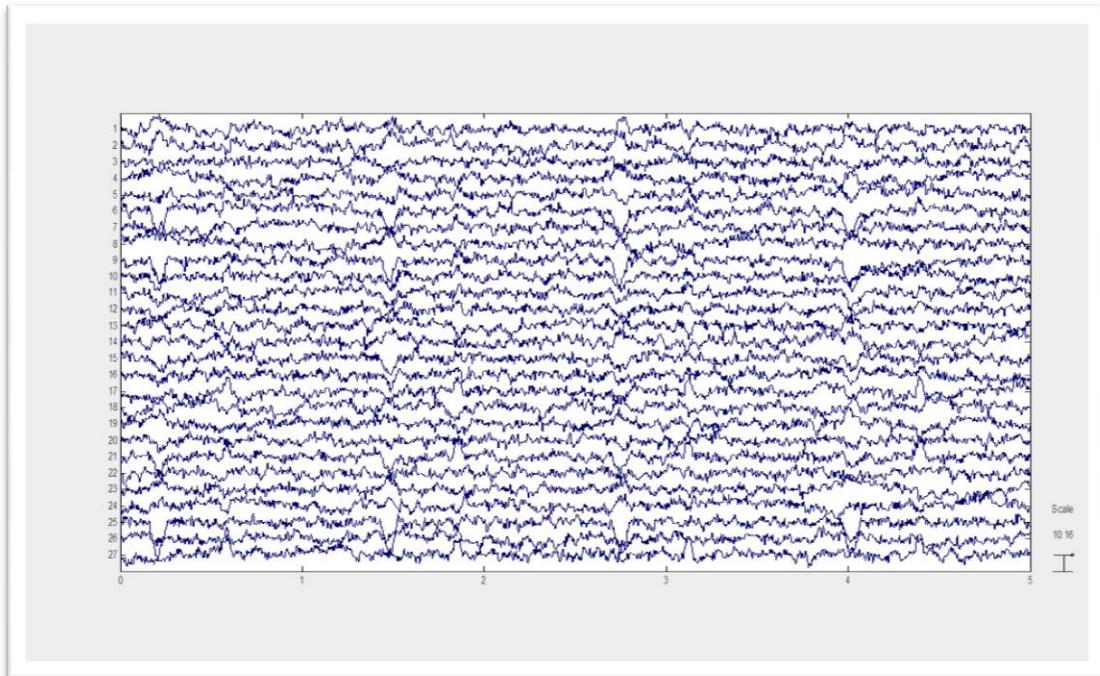


Figure 5.4: The simulated EEG signal used for evaluation of the ICA stability using the bootstrap method.

Since the dataset is composed from two sources with distinct and constant topographic maps and background EEG like noise, which has random energy distribution from trial to trial, we expect to reliably separate at least the two constant sources. The EEG-like noise is not Gaussian and does not violate the ICA model. We expect though, that the resulting components will vary greatly between the different runs. Therefore we will not take under consideration, whether the EEG-like noise is consistently separated or not. This approach is valid, as other researchers have found in EEG experiments, that ICA is able to separate EEG components that have a distinct topography, while other sources of short duration or low amplitude are grouped together into one component[1]. The way such sources are grouped is somewhat random and cannot be determined beforehand. This is what we expect from this experimental setup; two sources that can be separated and a number of components containing the EEG-like noise, without any meaningful ordering. We will concentrate on the efficiency of separating the two known phasic sources under different scenarios.

Simulation Results

We want to evaluate the reliability of the ICA results when starting from different initial conditions and see the results in comparison with sample size. Keeping the same sample size we run ICA multiple times using different initial conditions for each run, without using bootstrapped data. Applying this procedure to datasets of varying sample sizes we can evaluate the effect of starting from different initial conditions in respect with sample size. We used the estimates from 30 runs, using all channels

without dimensionality reduction. For each run we used random initial weights. The procedure was repeated for samples sizes of 1, 10, 25 and 40 trials (1 trial = 1300 samples). In Figure 5.5 we present the cluster quality index and the resulting independent components, using the mixing matrix of the centrotype in the case of sample size of 1 trial. We can see that the maximum value is $I_q = 0.75$. Only one source is separated (component number 4 in Figure 5.5 first row, right). Visually inspecting the projection of the clusters in 2D space confirms that there exists only one well separated cluster. Another interesting observation is that the number of components grouped in this cluster are 60 in total. This indicates that in most runs the two phasic sources must be combined together. Examining the other components included in the cluster we confirm that the two phasic sources are combined into one component.

As the sample size used, increases we see that the two phasic sources are efficiently separated from the rest of the noise into one component each. In Figure 5.6 we see the results using the full dataset of 40 trials.

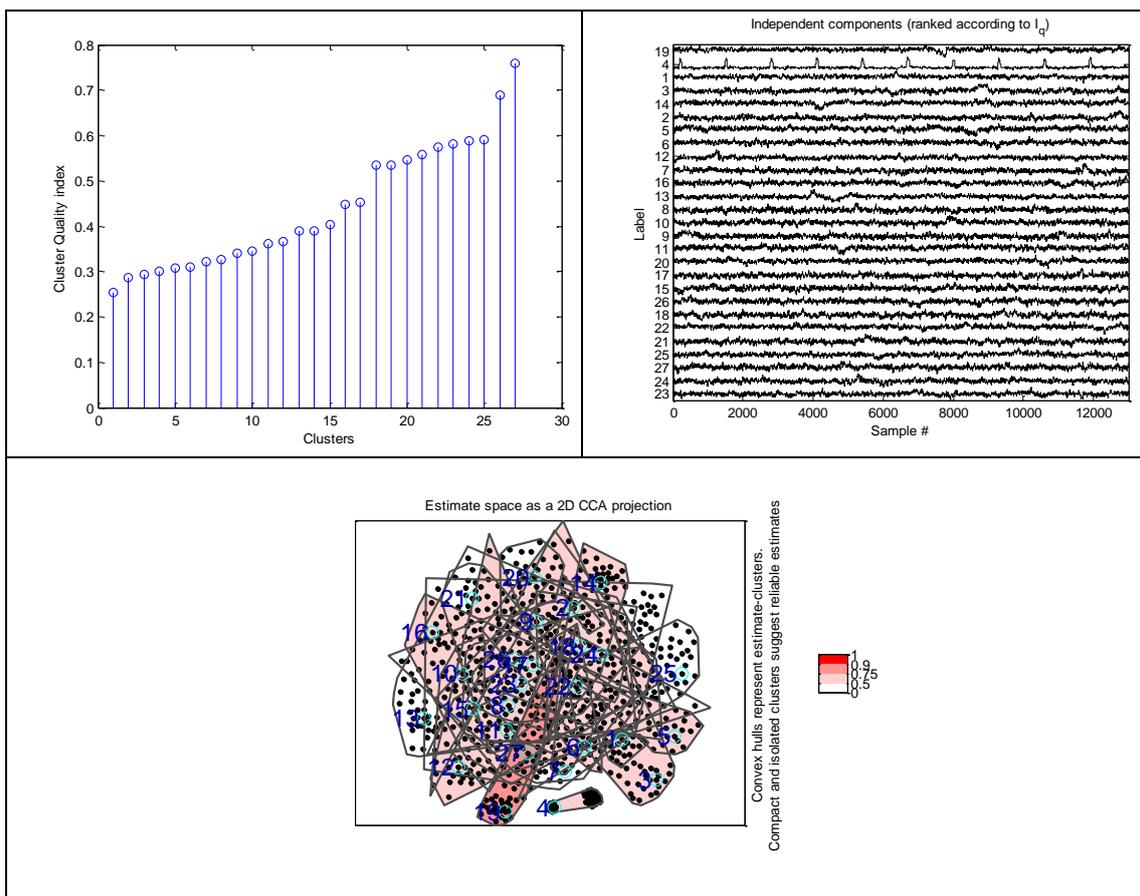


Figure 5.5: Results using data with sample size of 1 trial (1300 samples) as input to algorithm. We expect to reliably separate at least the two ERP-like sources. The results indicate that the two sources are not well separated. The low cluster quality index(max 0.75) as well visual inspection of the resulting components confirm that the sources were not separated into stable components Only one cluster is well separated from the rest as can be seen in the second row where the CCA projection of the clusters are presented

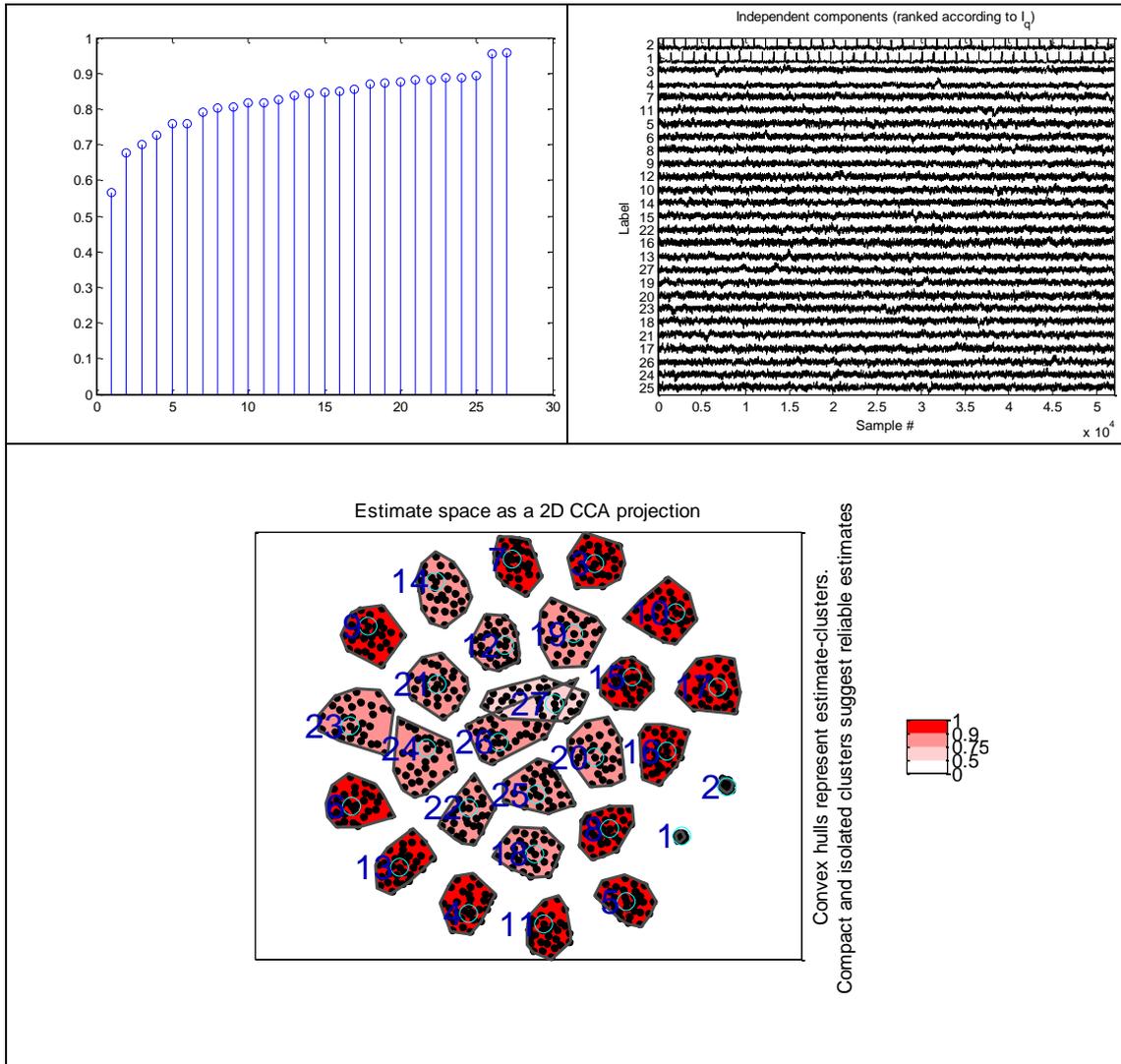


Figure 5.6: Results using data with sample size of 40 trial (1300 samples) as input to algorithm. The two phasic sources were separated into two stable components, as the high value of cluster quality index indicates. Visual inspection of the resulting components confirms that. The two clusters of the sources can be seen to be well separated from the others in the second row where the CCA projection of the clusters are presented.

5.4.5.2 Application in Real ERP data

We applied the above methodology to data from a single subject performing an auditory oddball event related experiment. In order to display the effect of sample size and different initial conditions we applied the method in recordings from a healthy (control) subject performing an oddball auditory experiment with 40 trials and 27 channels. If the algorithm provides consistent results in each run, then the calculated clusters will contain the estimates of one of the components obtained over multiple runs and the cluster will be compact; it will have a high cluster quality index. Here, we selected clusters that have an index of 0.8 and higher. We performed the following experiments in order to examine the algorithm performance under different conditions. In the case of real EEG data, we do not know the number of significant sources beforehand. The most likely is that the number of sources activating, greatly surpasses the number of electrodes. We expect though that there exists a number of sources activating in a consistent

manner from trial to trial. We expect these sources to be reliably separated. We also expect that inconsistent and small transient activations to be combined into the remaining components. As the sample size used increases the consistent sources will be separated in a few consistent reliable components, while the transient sources will be shared into the rest.

Running multiple times with random initial weights

We want to evaluate the reliability of the ICA results when starting from different initial conditions and see the results in comparison with sample size. Keeping the same sample size we run ICA multiple times using different initial conditions for each run, without using bootstrapped data. Applying this procedure to datasets of varying sample sizes we can evaluate the effect of starting from different initial conditions in respect with sample size. We used the estimates from 30 runs, using all channels without dimensionality reduction. For each run we used random initial weights. The procedure was repeated for samples sizes of 1, 10, 25 and 40 trials (1 trial = 1300 samples). In fig 3.2 we can observe the number of valid clusters for each run. Initializing the algorithm with different weights, affects the final result significantly. In the case of small sample size, we can see that we can reliably recover 4 independent components out of the 27.

As the number of samples increases so does the number of independent components that are consistently found by the algorithm. In Figure 5.7 we can see the cluster quality index for each run.

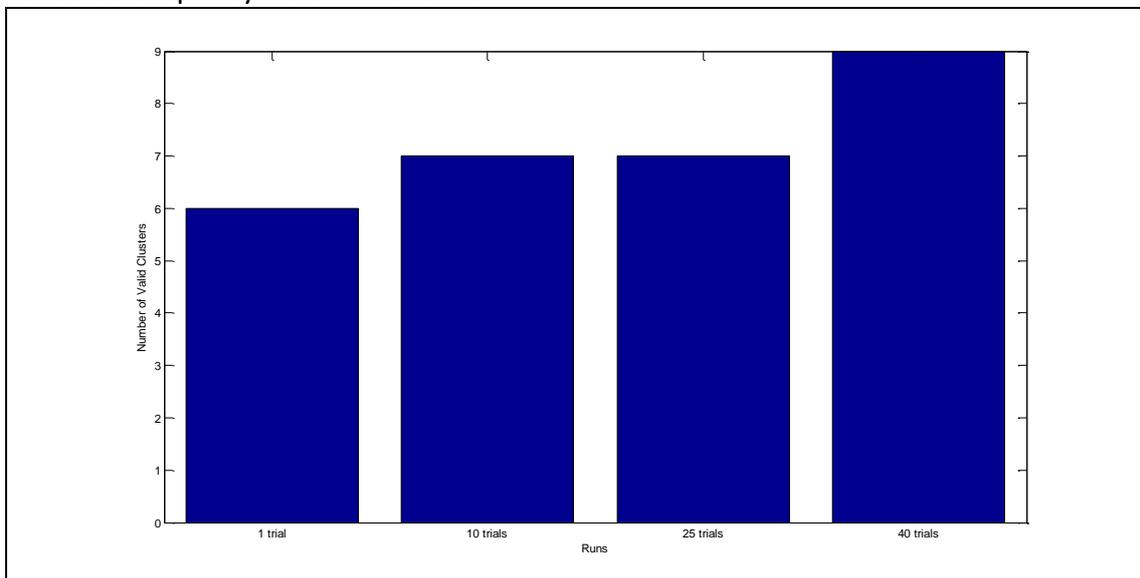


Figure 5.7: Number of reliable components recovered using different initial weights, using different sample sizes. Each bar presents the number of clusters that present cluster quality index above 0.8 for each run. We can see that as the number of samples used increases, the more stable components are recovered.

Running multiple times using bootstrapped data.

b) Using the same initial conditions, run ICA multiple times for the same sample size using bootstrapped data. We perform this experiment in different sample sizes in

order to evaluate the effect of sample size when starting from the same initial conditions.

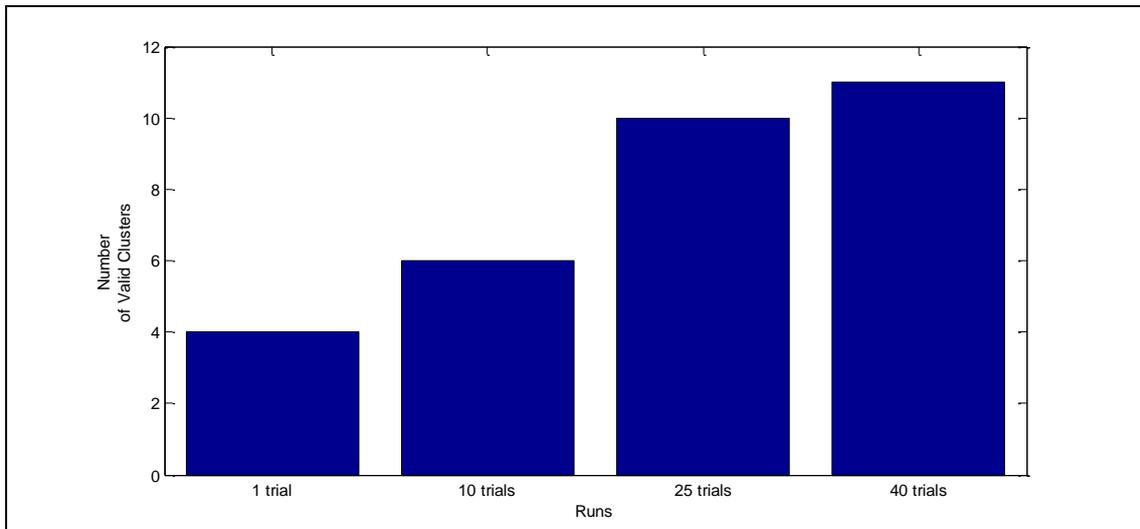


Figure 5.8: Number of reliable components recovered using different initial weights, using different sample sizes. We can see that as the number of samples used increases, the more stable components are recovered.

Running multiple times using both bootstrapped data and random initial conditions

Altering both initial conditions and using bootstrapped data in each ICA run. In this case we want to assess the reliability of ICA based only in the sample size.

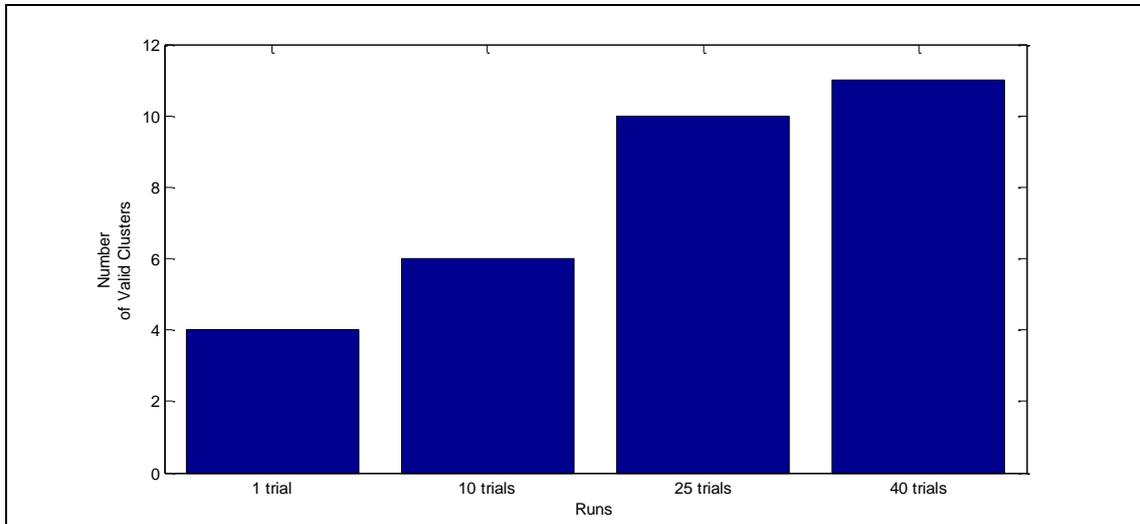


Figure 5.9: Number of reliable components recovered using both different initial weights and bootstrapped data, versus different sample sizes. We can see that as the number of samples used increases, the more stable components are recovered.

We calculated the cluster quality index for 27 clusters (we have 27 channels), for each case and for each sample size.

5.4.6 Advantages and disadvantages of the bootstrap method.

Using resampling techniques, like bootstrapping alongside with clustering can reveal the validity of the ICA results in a given dataset. We can see that the initial weight

matrix can affect significantly the produced result. Although, in the case of Infomax algorithm this does not seem to be a problem, since by default it will start from the identity matrix, it is important to have a notion of how a change in the initial weights will affect the algorithm's outcome. It was expected that the more data samples we use the more stable the algorithm becomes.

The main drawback of this technique is the amount of time that is needed. In average, when using the full dataset, the algorithm needed 79 minutes in order to perform 30 ICA decompositions, in an average laboratory personal computer (Intel Core 2 Duo 2.93GHz, Matlab 2010a). A typical ERP analysis involves many subjects and possibly many more trials than we used. This makes the use of this technique difficult to be implemented in practice. On top of that, it is not straightforward what a good threshold for the cluster quality index is. Different thresholds can be used that are equivalently acceptable. Another factor that affects the estimated clusters is the number of iterations. Although the number of bootstrap samples recommended in literature has increased as available computing power has increased, there is not any standardized number of iterations. Increasing the number of samples cannot increase the amount of information in the original data; it can only reduce the effects of random sampling errors which can arise from a bootstrap procedure itself.

In the case of ERP data, where we expect some sort of repeatability of the active sources, the more data (trials) we feed into the algorithm the more reliable sources we expect to get. Using simulated and a real dataset supports this conclusion. The more trials we use the more are the chances that consistent EEG activations will be separated into stable reliable components. We have to note that this conclusion applies into ERP data. It is not straightforward that this technique would provide better results in every case. EEG captures complex brain activations and therefore by supplying the algorithm with EEG data from different mental activities could easily have the exact opposite effect, too many active sources and very few degrees of freedom to efficiently separate them.

This observation leads to the use of an alternate method to estimate the reliability of independent components. Under the assumption that there exist consistent activations from trial to trial then we can use a well known procedure for estimating internal consistency of our results like split-half reliability.

5.5 Split-Half approach.

A well known technique to assess internal consistency stability is Split-half comparison. In split-half comparison we randomly divide all items that we measure into two sets. This is a cheap alternative to true replication of the measurement. While using only half of the data to calculate ICA diminishes the result, replication of components across the two datasets is a strong evidence of stability. This also means

that the extracted components are likely to contain interesting brain activity. Recently Groppe[39] used such a technique in order to estimate reliable components. The main advantage of this technique is the reduced computation time which acts as a burden when applying the bootstrap method.

The method can be described into five distinct steps:

1. Perform ICA on the full dataset.
2. Split the data into two comparable datasets.
3. Perform ICA on the two datasets.
4. Pair each IC from the initial dataset with an IC from each of the two subsets.
5. Find significantly similar triplets of ICs.

The concept is similar with the bootstrap approaches, which essentially perform ICA analysis on a large number of possible subsets of the data. Rather than using random permutation of the data as subsets we perform the analysis into 50% of the whole data, twice. This approach has the disadvantage that it can be biased by the method chosen to split the data. Also, it uses only half of the samples as we mentioned earlier. We could use the bootstrap approach presented above by shuffling the data trials instead of the individual samples and running the algorithm many times. We can say that the split-half approach is a simplification of the bootstrap technique which has the advantage of less computational time but with less accurate estimations.

5.5.1 Splitting the data

The sources of data variation in the case of ERP are the experimental conditions. In many cases the subjects have to respond in different events (e.g. target vs. non-target sounds in auditory oddball task). It is important to split the data so that only each experimental condition is equally represented in each set. Another source of variation is the level of alertness and fatigue of the subject during the experiment. The level of alertness and fatigue has large consequences in the performance of the subject in responding to the task at hand. For example the subject is more likely to blink, make mistakes and they tend to respond slower as the experiment progresses. In order to adequately represent these conditions in the two halves, the most intuitive way is to separate the dataset into odd and even trials, of each experimental condition.

5.5.2 Pairing the Independent Components

In order to pair together the n ICs from each half we need a similarity measure like in the case of bootstrapped data. Identifying a possible pair of ICs first requires an IC similarity or distance metric. Since the dimensionality of IC topographies is orders of magnitude less than that of IC activations and is easier to visualize, we use the one

minus the absolute value of the cosine similarity metric to compare IC scalp topographies as a first step in forming pairs of ICs. The metric used is

$$dist_{topo}(i, j) = 1 - \left| \frac{a'_i a_j}{\|a_i\| \|a_j\|} \right|$$

where a_j is the j th column of the mixing matrix A .

This metric deals with the IC scaling ambiguity by effectively normalizing the scalp topographies to unit root-mean square (RMS) magnitude and setting their polarities to minimize distance.

The final pairing algorithm is as follows:

1. Compute the scalp topography distance, $dist_{topo}$, between each possible pair of ICs to quantify their similarity.
2. Pair the two most similar ICs and remove them from further consideration.
3. Repeat Step 3 until all the ICs are paired.

5.5.3 Significantly similar Independent Components

The last step is to find which pairs are similar enough to be considered equivalent. We need to define null hypothesis and an appropriate critical region of the distribution of that metric. In [39], it is proposed to use the topography distance metric in conjunction with a similar distance metric for the independent activations. For the independent activations to be comparable, they need to be calculated using the same data. Therefore we will use the full dataset with the respective unmixing weights to calculate the activations, like this:

$$s_i = w_i x_t \quad \text{Eq. 31}$$

$$s_j = w_j x_t \quad \text{Eq. 32}$$

The distance metric for two independent activations u_i and u_j is:

$$dist_{act} = \max(f(i, j), f(j, i)) \quad \text{Eq. 33}$$

Where:

$$f(i, j) = \frac{\sum_t u_i(t) \|a_i\| \text{sign}(a'_i a_j) - u_i(t) \|a_j\|}{\sum_t (u_i(t) \|a_i\|)^2} \quad \text{Eq. 34}$$

This is the maximum normalized sum squared difference (i.e., residual variance) between the pair of IC activations. This measure equals zero for identical activations and grows towards infinity as the activations become increasingly dissimilar.

Using residual variance Eq. 34 as a similarity metric has the advantage that the difference between activations is scaled in proportion to the magnitude of one of the ICs. Thus small differences between large activations will be down-weighted relative to small differences between small activations. However, the residual

variance between two activations depends on which IC's activation is in the denominator. The metric is normalized by the smaller activation in Eq. 33. This way the smaller activation produces a larger value in Eq. 34). Doing otherwise would lead ICs with large activations to appear rather similar to any IC with a small activation.

We still, have to define an appropriate null hypothesis that two independent components are not similar. Using the distance measures for all possible independent component pairs, we form the empirical distributions for the two distance measures. Some of the independent component pairs are similar but the maximum number of similar pairs will be n out of n^2 . Therefore, the bias introduced by the similar pairs is relative small and in case this bias affects the approximation, it will only make it less permissive; it will be more difficult to find similar two similar pairs.

Given a distance distribution, we must select a critical region for rejecting the null hypothesis (i.e., that a pair of ICs are no more similar than a pair of ICs chosen at random) at a particular α level. Because it is necessary to test both IC features, the critical region is a segment of the two dimensional joint distance distribution. We prefer to select independent component pairs that present a highly similar measure (low values of the distance metric) rather than a pair that have two features with medium values. In [39] it is proposed the use of an L-shaped significant region, which favors pairs with a highly similar feature over the later.

We want all ICs from one decomposition to be significantly similar to exactly one IC from the other decomposition, if possible. The critical region should be big enough to contain $1/n$ of the empirical distribution. Also, independent components with a highly similar feature should have priority. The critical region that fulfills both these requirements is an L-shaped region constructed from two rectangles. Remember that the most similar features lie at the lower values of each distribution. Each rectangle is designed the rightmost edge and top edge, set at the 90th percentile of the scalp topography distance distribution and activation distribution, respectively (see FIGURE). The other edge of each rectangle is then grown, one sample at a time, until both rectangles together contain $1/n$ of the empirical distribution.

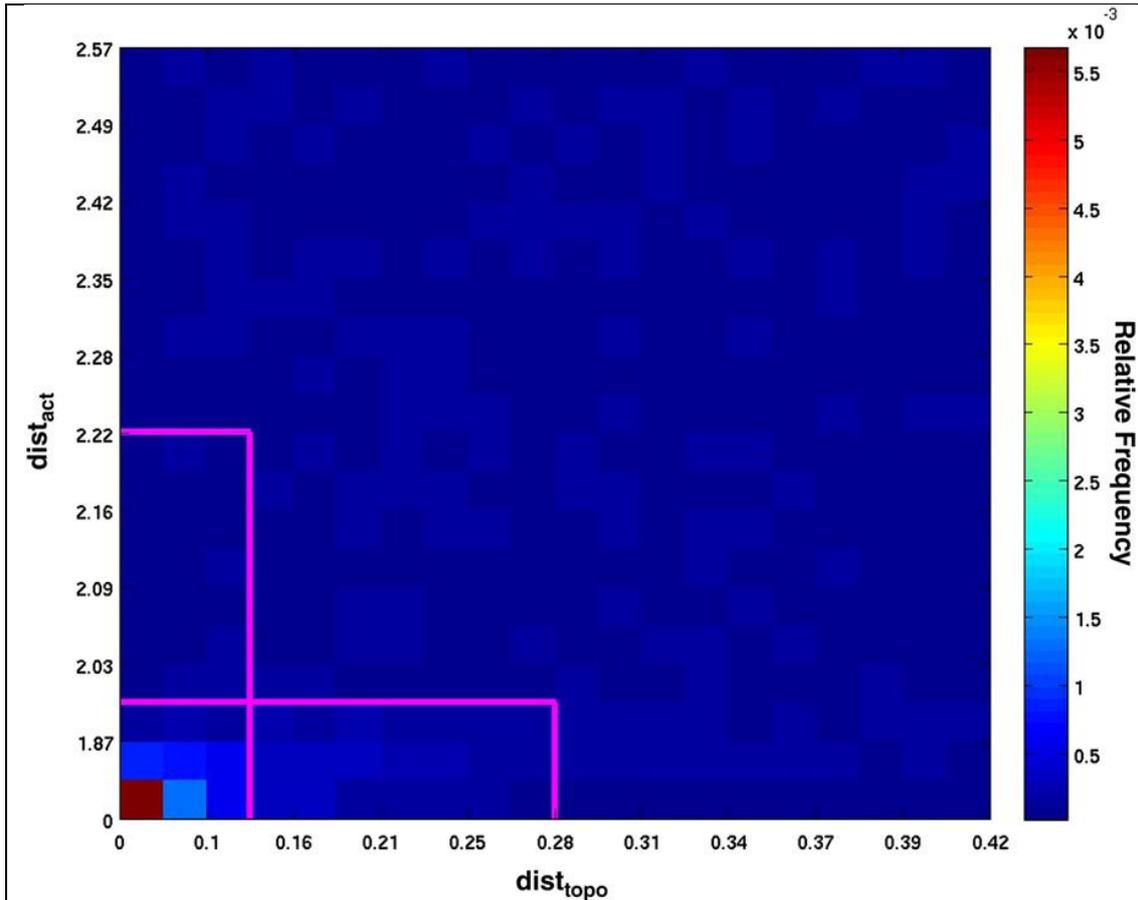


Figure 5.10: Image from [39]. Joint distribution of the IC pair topography and activation distances. The two distances are are binned in 1 percentile increment, in order to be comparable. The rectangles indicate “L” shaped critical region that contains 1/64(64 channels in example) of all samples.

5.5.4 Advantages and disadvantages of split-half method.

As we already mention, this technique is influenced by any bias in the method used to split the data, the bias from the similar components used to construct the empirical distribution and the fact that we are using only half of the data in the two subsamples. In [39], extensive comparisons between split-half and bootstrapped were performed. The results show that the two resampling methods produce highly correlated reliability estimates. In other words, the ICs whose scalp topographies and activations reproduce the best across the two split-halves also tend to be the most reproducible according to the bootstrap method. Split-half comparison approximates well the general results of the time and resources intensive bootstrap method. Overall, there appears to be no clear advantage to the more time intensive bootstrap resampling.

Since the quality of data and length of data are important, we compared the number of reliable independent components derived from the bootstrap test in 0 with the results of split-half in the same dataset. In the case of the bootstrap method we

collected 11 reliable independent components using a cluster quality threshold of 0.75. Using the split-half method we derived 8 reliable independent components. This is in line with the disadvantages of the method, that we mention earlier. The split-half method is less permissive than the bootstrap procedure.

The main advantage of the split-half procedure remains the limited requirements in time and computing resources. The convenience of the method is critical factor of whether it will be used in real life analysis.

Identifying reliable components is the first step using the ICA analysis. The EEG signal recorded in the different electrodes contains a magnitude of activations and there is little a priori regarding the brain processes that takes place. Using a technique that provides a number of reliable, stable components to evaluate gives credibility to the extracted results. The rest of the independent components that were found unreliable do not mean that they contain irrelevant or extra cerebral activity. On the contrary, components containing strong artifacts are more likely to be stable. Therefore a methodology that exploits the reliable components and also takes under consideration the activity contained in the remaining ones is needed in order to fully explore the wealth of information contained in the EEG signal. Under this perspective it is apparent that using reliability scheme that is less permissive is not necessarily a drawback. In the next chapter we will provide a methodology for working with ERP data using the split-half comparison for identifying reliable components and extracting as much information possible from the remaining sources.

Chapter 6. Proposed Methodology.

6.1 Introduction

Independent component analysis is being used extensively in the analysis of EEG. It is used to decouple the different brain activations and facilitate the characterization and evaluation of interesting brain activity. Brain activity is often considered on specific electrodes, whose positioning matches the expected localization of the generating brain source. Due to brain conduction effects, such activity is spread through a number of electrodes. Alternatively, the consideration of independent components appears to be appropriate for the consideration of brain activity, since the decomposition process aims at recovering (independent) brain signals. As we discussed in the previous chapter, an aspect of independent components analysis which is often neglected is how much we can trust the results of the algorithm applied on a certain dataset. Using the split-half approach, it is possible to derive a subset of the available components representing stable, reliable sources. Furthermore, using the different measures of ERP activity described in section 3.3, we can analyze and characterize the brain activations captured in these components. Knowing that the components represent stable activations we can use them to extract quantitative features, which could be used as biomarkers in a later stage.

In addition to stable ones, we also have the complementary set of components, which were not found stable enough. As we noted earlier, stability does not imply validity. Accordingly, although some components were found stable, they do not necessarily contain interesting brain activity and vice-versa. The non reliable components cannot be discarded easily, before evaluating their content. We can take advantage of the non-reliable independent components by reconstructing their contribution to the channels; a procedure also described as back-projecting the components to the channels. Then, we can analyze the combined back-projected activations on channels, using the same methods and measures as on stable components.

Using this approach we obtain two sets of components for each subject. One set contains the stable independent components which are evaluated separately. The other set contains the rest of independent components which were found to be unreliable. We choose to back-project the unstable components back to the channels where the matter of stability poses no problem and evaluate the remaining activity in the channels. The above methodology was applied to a dataset of Control and Alzheimer subjects as it is described in the following section. We can summarize the proposed methodology in the following steps:

1. Apply ICA using the split-half procedure.
2. Apply the ERP activity measure on the reliable ICs to evaluate their content.
3. The remaining ICs activations are evaluated separately.
4. Apply the activity measures on the channel projections of each component and evaluate content.

6.2 Experimental Setup of the Acoustic Experiment

EEG was recorded from 27 electrodes encompassing the largest possible area, recordable from 27 equidistant 10% positions using an EEG Recording Cap made by FMS (Falk Minnow Services), model EASYCAP, and the Large Equidistant 32-Channel-Arrangement, montage No. 23, giving, inter-electrode distances of 43 to 68 mm (for a head circumference of 58 cm). All EEG electrode impedances were lower than 5k Ω . Linked ears (A1-A2) were used as the recording reference with electrode AFZ as the ground. The recordings were made using the EMS CGmbH model Phoenix Clinical Lab Digital EEG machine. Following the guidelines in [46], signals were digitally sampled at 1024 Hz, with a high pass filter of cut-off frequency 0.016 Hz, a low pass filter of cut-off frequency 60 Hz, and a notch filter at 50 Hz (to remove electrical mains contamination). A hardware based multi channel ERP stimulator made by EMS-GmbH, model Helios II provided forty 2-kHz target tones (20%) and 160 1-kHz non-target tones (80%). The ISI was 1.29s. Subjects were seated with closed eyes, were relaxed, and were instructed to listen carefully and press a button immediately they heard the target tone.

A 10s epoch of EEG data was recorded for each subject, both before and after the total of 200 tones from the stimulator. Altogether 360 target trials were recorded from each of the normal and patient subjects. For each of the target stimuli, 600 samples before the stimulus and 700 samples after the stimulus (1300 samples) were taken to form a target single trial ERP waveform lasting 1269.5 ms.

The dataset consists of 9 healthy subjects and 9 AD patients performing an auditory oddball experiment. The dataset was provided by the Ecological University of Bucharest, Romania, obtained after an approved ethics protocol. The control participants (3 females and 6 males) had no history of neurological or psychiatric disorder and were between 37 and 74 years old. The AD patients (2 male, 7 female) were between 57 and 88 years old, and were recruited for P300 recordings after diagnosis and often after the commencement of therapy with a variety of drugs. All were of higher education, and were diagnosed with dementia of the Alzheimer's type in the early stage and mild form, by means of psychometric tests and Cerebral Tomography (CT). They had Mini Mental State Examination (MMSE) scores ranking from 18 to 29 and the CT examination showed cortical or cerebral atrophy. The

patients were on drug treatment with cholinesterase inhibitors. None had consumed alcohol. Even though aging affects the characteristics of activation, it does not alter the actual pattern of activation, making it reasonable to conclude that differences in the pattern of activation are due to the pathology itself.

It has been reported that the P300 ERP paradigm is directly related to mental efficiency and reflects brain processes that demand attentional allocation and fast memory processing [9]. Furthermore, the P300 response has been primarily located in the temporo-parietal cortex, which is the area most severely affected by Alzheimer's disease. According to previous studies of P300 [17, 24], delta and theta bands present phase-locked activity contributing directly to the ERP waveform, while alpha band characterizes both evoked and induced phenomena.

6.3 Evaluation of components of the 1st experiment

We applied the split-half comparison technique to the dataset presented above. ICA was applied in the concatenated dataset of 40 target tone trials, for each subject. Then, each dataset was split into even and odd trials, in order to avoid bias and ICA was calculated on the two half datasets. Following the procedure presented in section 5.5 we derived for each subject a set of stable components. Using the global indices of the measures for each component, as described in 4.7.1-b, we proceed to evaluate the components that present significant activity in each band.

In general there were no significant differences from Normal to Alzheimer subjects in the number of stable components. On the average 10 components were found stable for the subjects in each group.

Our goal is to find activation patterns that will be characterized as stable for each group. Since ICA tends to separate strong and consistent sources into single components, we expect that the activations found in the stable components will capture activity directly associated with the given task. Furthermore, we expect to find patterns of activation that are common across subjects of the same group. In order to accomplish the evaluation of activation patterns, we calculated the power of the spectral average, phase intertrial and phase-shift intertrial coherence in each stable component, as well as the ERD/S measure. Recall that all four measures summarize the content of each component throughout the trials. The global indices of the measures were used as a first step of selecting components, followed by visual inspection and statistical evaluation of the time-frequency maps for each measure.

More specifically, for each time-frequency map we computed the histogram and used the upper quartile in order to derive a threshold for defining significant regions in the map. Then we calculated the center mass of the significant region in order to classify the central band and latency that the activation occurs. Based on this

threshold, we segment regions of interest in the measures and derive their mean time, frequency and spectral power. For each subject we group the stable components that present significant activation in each measure (power of the mean signal, PIC, PsIC or ERD/S) according to the band that the activation takes place. Components that presented significant activations in multiple bands are assigned to all bands that presented significant activation. The statistics of these activations patterns for control and AD subjects are presented in Table 6-1.

To illustrate the statistical significance of our findings, which are presented in the results section, we perform statistical testing between the participant groups (normal controls and AD) for each band. Since the goal is to identify significant differences between two groups, the independent two-sample t-test is used to assess whether the means of the two groups are statistically different from each other. As a parametric test, it assumes that: i) data comes from normally distributed populations, ii) variances of the populations involved are homogenous and iii) all observations are mutually independent [68]. By using the D'Agostino Pearson test, our measurements were found to have a normal distribution, thus satisfying the first assumption. Homogeneity of variances was tested using Levene's test, based on the F-statistic [68]; it was found that our groups do not have equal variances. As the second assumption is violated, the t-test was applied assuming unequal variances (Behrens-Fisher problem). Finally, since our populations reflect two independent groups (controls and AD), the third assumption becomes reasonable.

6.3.1 Control Subjects.

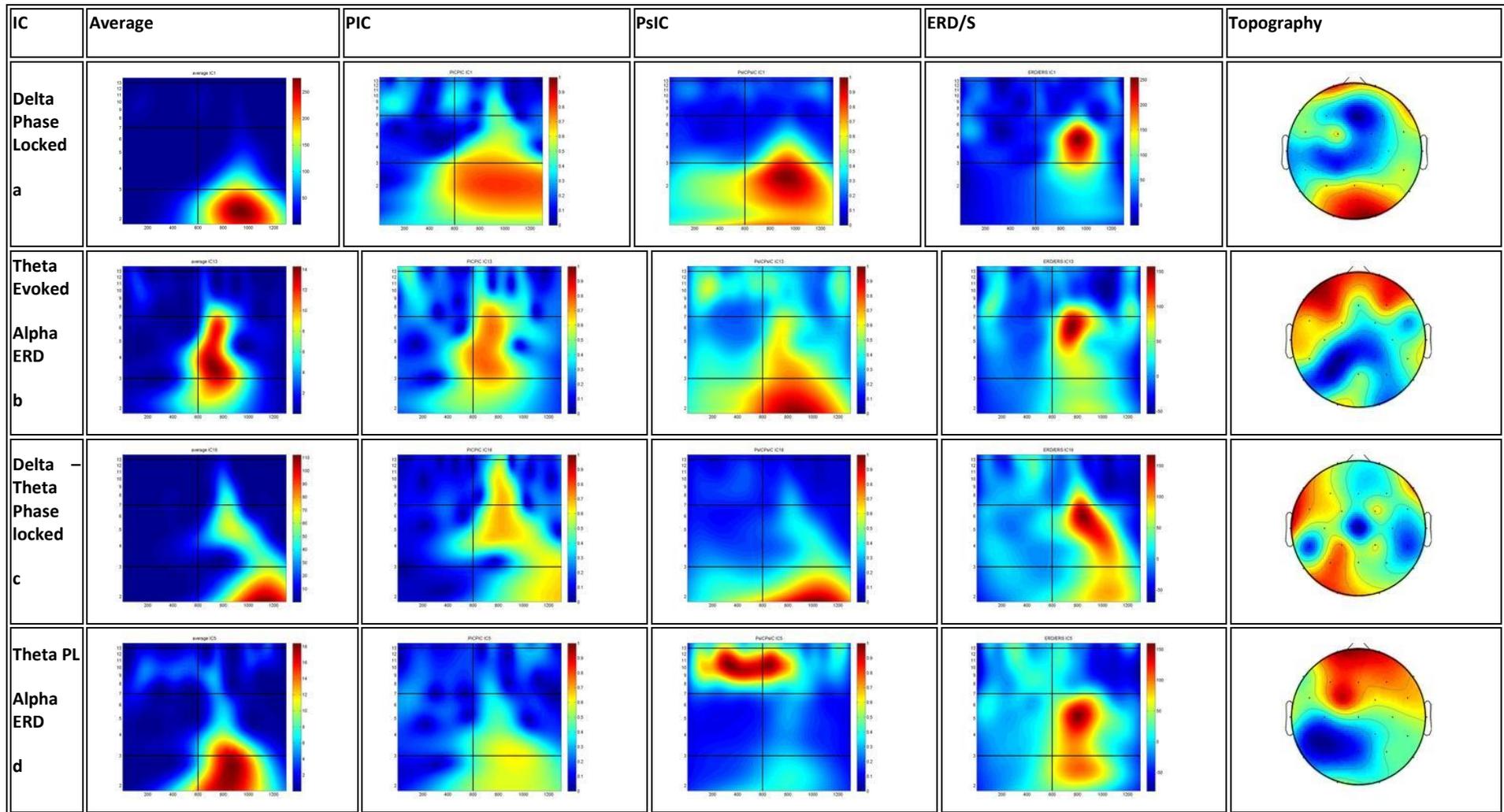
We found consistent activation patterns in subjects of the two population ensembles. Overall, we can report a general similarity of the patterns of activation in the delta and theta bands between the two groups. We first present some indicative results from one control subject in **Error! Reference source not found.** The horizontal axis in all illustrations of energy and coherence measures cover 600ms pre-stimulus and 700ms post-stimulus. The vertical axis spans frequencies from 1Hz to 14Hz in a logarithmic scale. The color map for the SE measure of the average is normalized within [0, 1] for each subject, whereas the color map of the other two measures is by definition constrained within the same interval. The colormap for the ERD/S measure is not constrained and represents the maximum decrease compared to pre-stimulus with blue colors and increase with red colors, per component and frequency: the mean prestimulus power is computed for each frequency. Finally, the color map for the topography map is normalized from 0 to 1 for each subject. These detailed TF maps provide information for the consideration and analysis of individual components, as well as their inter-relationship. They contain information to enable comparisons in both time and frequency content. Furthermore, the topographic map provides clues of the possible origin of each component.

The components presented in Figure 6.1 (controls) and Figure 6.2(ADs) do not necessarily correspond to a single source that reflects one specific type of activity. In fact, many components reflect mixed-type activity, or activations appearing in several bands. For instance, we can identify components expressing alpha desynchronization right after phase-locked theta activation, expressing semantic processing (alpha ERD) after some attention and memory activations (theta phase-locking). The derived components express separation of activity in temporal terms. One would also expect ICA to be able to decouple spectral activations. It is not clear yet if the result of mixed spectral activity is an inefficiency of ICA or it is indeed a characteristic of brain sources, which we would like to also reflect in derived ICs. In addition, these stable ICs capture the major patterns of activations engaged in the experiment and, thus, they can guide the spectral analysis and temporal sequencing of brain activities.

For the control group, components that isolated phase-locked delta activity appear with posterior topography and occur around 350ms after the stimulus directly contributing to the average P300 waveform, as can be seen in Figure 6.1a. Delta components often appear with theta phase-locked activations, but latter in time as in Figure 6.1c. Theta components preceded delta in time, displaying significant phase locking around 900ms or 300ms after the stimulus. In all control subjects delta and theta phase locked components also presented an increase in power post-stimulus indicating an additive activation emerging as a response to the task. **Error! Reference source not found.**c displays a theta phase locked component that presents strong phase locking and significant increase in power post-stimulus, as can be seen in the ERD/S measure. Some components also demonstrate induced (non phase-locked) delta activity exemplified by the PsIC measure, which appears at the same time as alpha ERD and possibly implies increased cognitive effort (**Error! Reference source not found.**b and c); notice the alpha ERD at the same time as alpha in the PsIC measure.

Alpha ERD is common in components that presented activity in the higher delta and theta band which also had central and frontal topography as in **Error! Reference source not found.**b, d and e. Figure 6.1d and e display components that capture theta phase locked activation with power increase and an alpha ERD phenomenon at the same time. Alpha presents significant PsIC activity pre-stimulus which diminishes as theta response emerges. The desynchronization reflected in dark blue in the ERD/S measure appears just after the decay of theta. This first group of alpha components appears with strong PsIC pre-stimulus and mostly parietal topography. It indicates an ongoing alpha oscillation, which reorganizes in order to facilitate the response to the event. Such strong alpha activity pre-stimulus has been found to be related to expectation for the stimulus[24]. Alpha activity reduces during ERD but a late alpha oscillation re-emerges after the cognitive activities (Figure 6.1b, d, e, g).

Some control subjects (6 out of 9) displayed a second group of alpha components with phase locking immediately after the stimulus, as in Figure 6.1f and c. Such a group of alpha components can be characterized as early evoked alpha oscillations, which desynchronize when delta and theta activations emerge and appear with strong frontal topography (Figure 6.1f). It has been reported that induced alpha band plays an inhibitory action and is increased in order for some other function (cognition) to emerge [13].



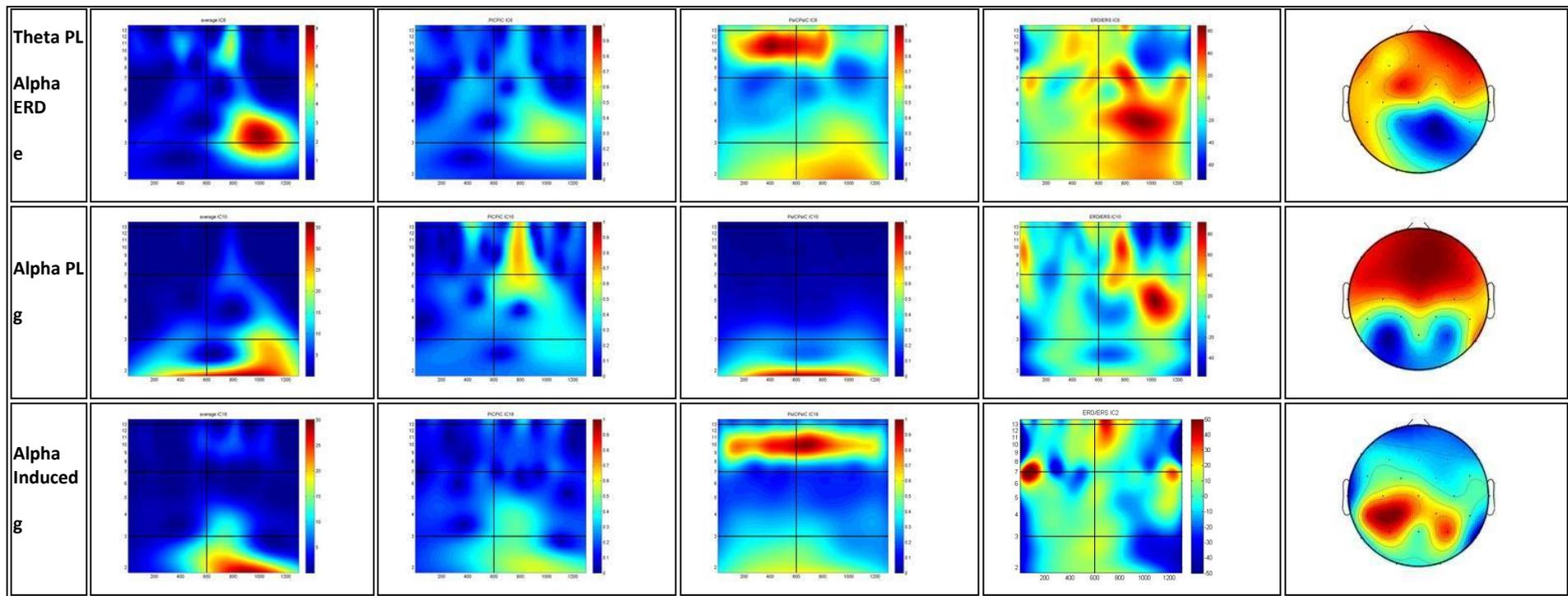


Figure 6.1: Example control components. Row a) displays a typical phase locked delta activation found in control subjects. It presents mixed activity with low theta which results in theta ERS. Row b,c) displays a theta phase locked component with energy increase expressed by the ERD/S measure. Rows e), f) and g) display components that present alpha activations. Row e) presents a mixed component with theta phase locked and alpha induced activity. Notice that Alpha activity desynchronizes when the phase locked activation emerges. Row f) shows an alpha induced oscillation which partial synchronizes in phase after the stimulus and then desynchronizes. Finally row g) shows a component with only an induced alpha oscillation which desynchronizes post-stimulus.

6.3.2 Alzheimer subjects

We follow here the same methodology as for the control subjects. We identified common patterns that appear in all subjects and categorized them per band. Activation patterns of stable components in Alzheimer subjects share some similar aspects in comparison to controls in delta and theta bands. For example, delta and theta activations appear in Figure 6.2a, that express parietal (posterior) topography dominated by delta, which also reflects on the average spectral energy.

Theta band components have frontal topography as in control subjects. Theta band activations occur in the upper limit of the theta and also enclose some parts of lower alpha band (Figure 6.2b). As in controls theta band precedes delta band in time. All theta phase locked activations are manifested with power increase in the ERD/S measure. Nevertheless, both theta and delta activations present lower phase locking values than the controls. This can be attributed to the Alzheimer disease which affects the brain mechanisms responsible for the orientation and response to the task. Delta band phase-locked components are less in AD subjects with theta components being dominant in AD, mostly due to the use of drugs[26].

Alpha band in Alzheimer's presents weak induced oscillations in few subjects (3 out of 9) and it is presented in lower alpha band around 8Hz (Figure 6.2c). No higher alpha band activity is detected in any component in Alzheimer subjects, which is in line with other findings suggesting that alpha-band activations emerge at lower frequencies in Alzheimer subjects [29]. The most common alpha activity detected in AD subjects is evoked just after the event (as in Figure 6.2d), indicating increased effort for attention.

Figure 6.2c shows central alpha pre-stimulus oscillations that get desynchronized after the event. Notice the partial phase locking of delta and theta bands, which results in power increase while alpha is desynchronized. Figure 6.2d shows an alpha oscillation which is evoked after the event and then desynchronizes, while evoked theta also increases its power. Both effects contribute to increased efforts for attention in ADs, but this effort is not succeeded by strong alpha ERD to engage in cognition processes. Induced late alpha oscillations are nonexistent in AD subjects. Late alpha induced activity has been related to semantic memory processing and possibly indicates an alternative pattern of activation in order to replenish the reduced abilities due to the disease.

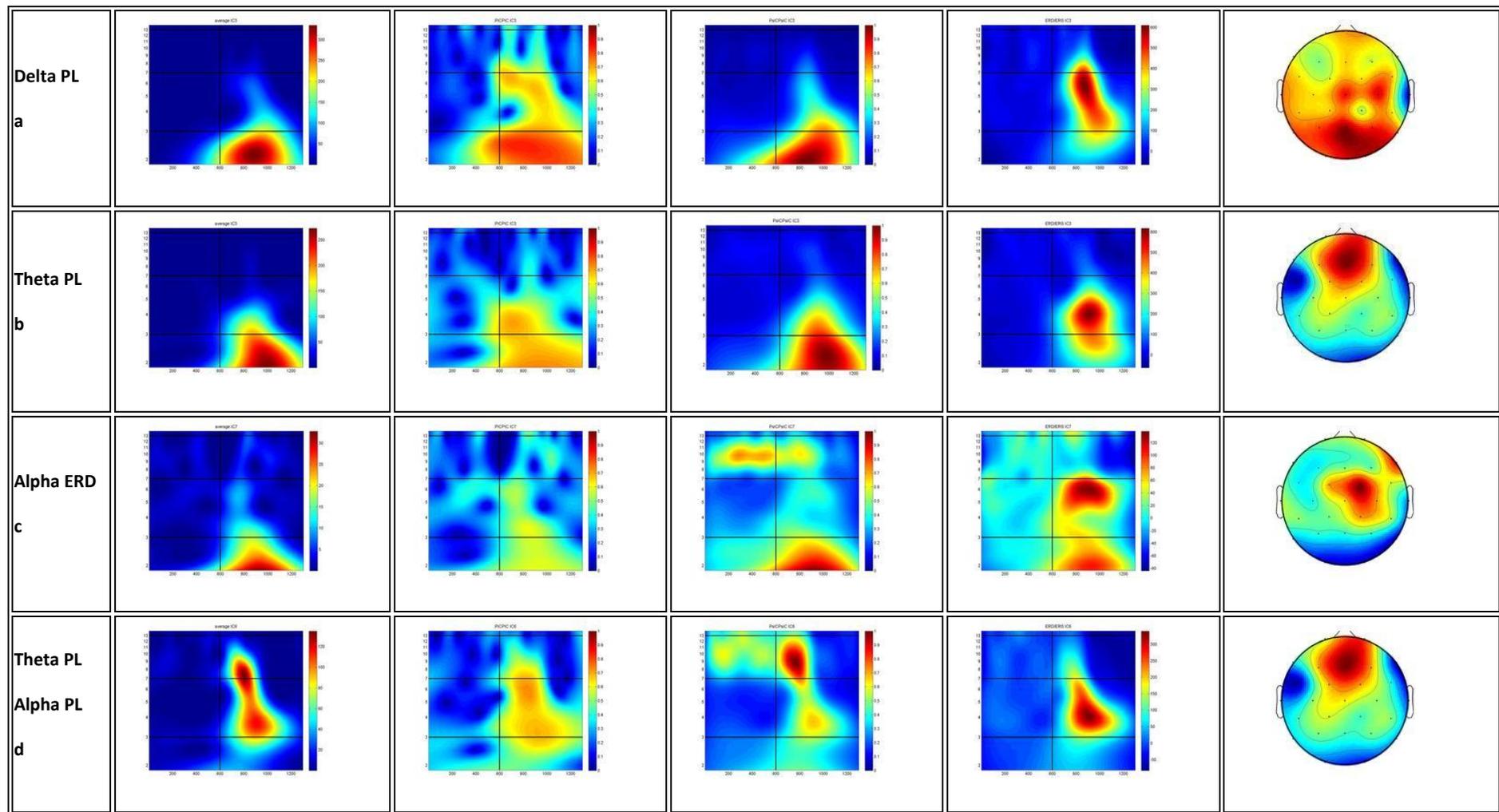


Figure 6.2: Row a) displays a typical phase locked delta activation found in AD subjects. Row b) displays a theta phase locked component with energy increase expressed by the ERD/S measure. An ongoing alpha oscillation is also captured in the same component. Rows c), d) display components that present alpha activations. Row d) presents an alpha evoked activity which desynchronizes before the P300 response.

Table 6-1: Group Statistics for Control and AD subjects. Left Table shows statistics for the phase locked activations. Right table displays statistics for the induced activations.

	Controls PIC	AD PIC	P-Value: a=0.05		Controls PsIC	AD PsIC	P-Value: a=0.05
Band	Statistics	Statistics		Band	Statistics	Statistics	
Delta	Central time= 873	Central time = 869.8	No signif diff. p=0.022(Normal)	Delta	Central time= 910	Central time = 902	No signif. diff.
	Central freq =3.4	Central freq =3.2			Central freq =2.8	Central freq =3.2	
	Mean PIC energy = 658	Mean PIC energy = 675			Mean PsIC energy = 662	Mean PsIC energy = 697	
	Mean PIC values= 0.58	Mean PIC values = 0.45					
Theta	Central time = 802	Central time = 764	p=0.038514(AD) p=0.026(Normal)	Theta	Central time = 811	Central time = 734	No signif diff.
	Central freq =5.5	Central freq = 5.6			Central freq =5.6	Central freq = 5.6	
	Mean PIC energy = 354	Mean AD energy = 398			Mean PsIC energy = 206	Mean PsIC energy = 196	
	Mean PIC values = 0.68	Mean Delta PIC = 0.56					
Alpha	Central time = 914	Central time = 895	No signif diff. No signif diff.	Early Alpha	Central time = 711	Central time = 734.5	No signif diff.
	Central freq =11.3	Central freq =10.55			Central freq =9	Central freq =9.18	
	Mean PIC energy = 239.8	Mean PIC energy = 215		Mean PsIC energy = 239	Mean PsIC energy = 231		
	Mean PIC values = 0.50	Mean PIC values = 0.47					
				Late Alpha	Central Time = 1195	No late induced alpha	
					Central Freq = 10		
					Mean PsIC energy = 302		

6.3.3 Discussion of activation patterns for the 1st experiment

The P300 experiment considered in this work focuses on the retrieval of events with the presentation of a repetitive acoustic probe. As a response to basic information processing mechanisms of attention allocation and immediate memory, phase-locked (evoked) activity is produced that causes an increase in the amplitude and energy of the average signal in the delta and theta bands. Delta evoked activity displays posterior topography and is closely tied to cognitive processing [19]. Theta evoked activations activate first, presenting anterior topography and are more closely related to orientation and stimulus processing [19, 20]. Early (in time) alpha synchronization has been attributed to phase-locked oscillations related to attention and orientation to stimulus. Furthermore, later desynchronization of the alpha band has been related to non-phase locked activity instigated by semantic processing [18]. In terms of topography, the phase-locked alpha activity appears frontally, whereas the non-phase locked alpha is mostly localized in the parietal areas [13, 17].

The Alzheimer's disease reduces the ability to perform functions related to complex memory, resulting in reduced delta activity and phase-locking, compared to control state [27, 28]. In turn, theta activity has been reported to increase in ADs due to increased attention requirements [16] and also due to the use of drugs [26]. Furthermore, the alpha-band activity is weaker or completely absent in AD subjects. In essence, AD respondents may attempt to initiate phase-locked (evoked) alpha processes related to 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 [69].

The patterns of activation identified here are in agreement with the findings of previous studies. Summarizing delta components preserve a posterior topography, closely related to cognitive processing, while theta components appear earlier in time and are localized on anterior sites related to acoustic stimulus. Both types of components appear with high PIC coherence measures, indicating their phase-locked nature. The theta band in ADs is reflected with slightly more energy than in controls, in agreement to previous studies [16, 70]. It is worth mentioning here that the AD patients considered were all under drug treatment (cholinesterase inhibitors), which has been reported to increase theta activity in AD subjects [26]. Thus, the effects on the theta band need to be associated with the medication and cannot be purely attributed to the disease itself.

The control subjects reflect early phase locked alpha components, that are mostly centrally and anterior localized, related to orientation to the acoustic stimulus. Early alpha activity is observed only in some AD subjects and appears as partial phase-locked activity. This pattern of activation could be attributed to the increased effort of ADs for alertness and attention and different progression of the disease [17]. Our

results on AD subjects indicate absence of late alpha modulated by stimulus-induced effects and semantic memory processing. This performance is also supported by the findings in [18] during the experimental condition of demanding mental tasks in ADs, where the attentional capacities are exceeded and “alternative” cognitive strategies are utilized, explaining the lack of (late) alpha band in the cognition process.

6.3.4 Evaluation of Unstable Components.

A similar procedure based on the synchronization measures was also followed for the remaining components that were found unstable. For each unstable component, we derived the back-projected channels and the measures of phase intertrial coherence and phase-shift intertrial coherence were used in order to identify channels (of components) that may capture interesting activity.

In average 5 components were found to contain interesting activity per subject in the group of control subjects and 6 components in the group of Alzheimer subjects. All control subjects had components capturing induced alpha oscillations starting at pre-stimulus and desynchronizing at post-stimulus. Many delta components were detected in unstable components with induced activity and some weak phase-locking.

Alzheimer subjects presented similar patterns in their unstable components. The main alpha activity in AD subjects is expressed as phase-locked evoked just after the event. Deviating from the pattern of stable components, the unstable ones showed a larger number of components with alpha pre-stimulus oscillations which desynchronize post-stimulus, but with weak power. Furthermore, induced delta components were found in AD subjects, as in controls. Theta band components with significant phase-locked activity were observed in the majority of the Alzheimer subjects resembling closely comparable stable components.

In general, patterns of unreliable components closely resemble the activity and topography of those found in the stable components. An indicative example is presented in Figure 6.3. In the first row Figure 6.3a we observe an unreliable component containing theta phase-locked energy. Examining the ERD/S map, there is no significant increase in the corresponding band post-stimulus. The second row presents a stable component from the same subject. It contains similar topography, but the theta phase-locked activation reflects energy increase post-stimulus and is also reflected in the spectral energy of the average. The above analysis verifies that ICA techniques often suffer from instabilities. Significant event-related activity is carried by stable components, but some energy is also “leaked” into other components as a result of imperfect unmixing. The increased presence of (weak) alpha activity in the unreliable set, as well as the presence of induced delta components, supports the argument that unstable components engage minor activity complementary to the major one appearing in stable components. Indeed,

ICA has been reported to decompose strong sources into single components, while combining or splitting smaller sources into several other components. Alpha activity presents the lower power than the other two bands, since EEG follows a 1/f distribution of power. Nevertheless, the study of unstable components does not seem to contribute much complementary information to our analysis and is not recommended for a general study.

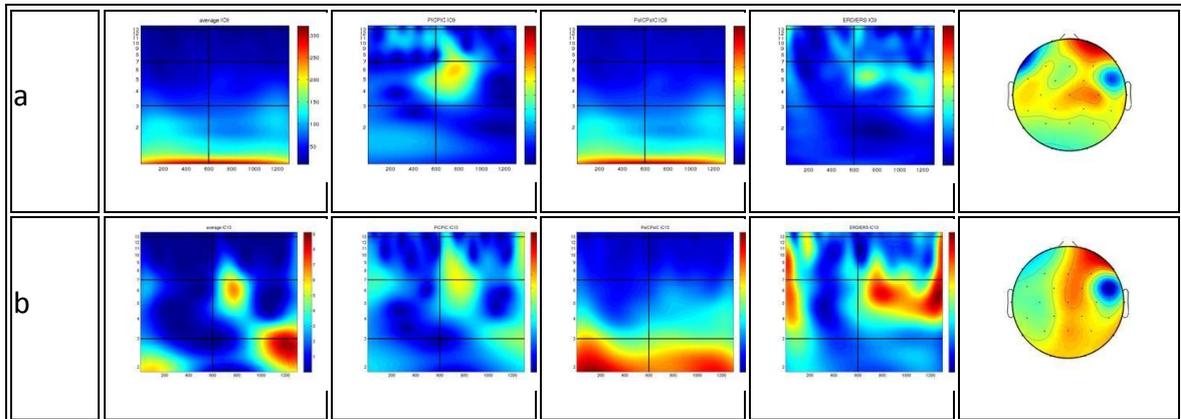


Figure 6.3: An example of an unstable component (a) and a stable one(b) from the same subject. Notice how the activation patterns are similar and also present similar topographies. The only difference is how phase locked theta is expressed in the average Spectral energy and the increase in power post-stimulus. The stable components present both these features, while the unstable component presents only phase locked activity. This means that some phase locked energy “leaked” to this component due to unsuccessful mixing.

6.4 Application on a working memory experiment

6.4.1 Dataset description

A group of 24 healthy, right-handed elderly volunteers (mean age: 77.5 ± 8.2 years) composed the elderly control (EC) group of the present study. All individuals were screened with different mental examinations like the Mini Mental State Examination (MMSE) the Lawton’s Instrumental Activities of Daily Living (IADL) and the Hospital Anxiety and Depression Scale (HAD). Furthermore, extensive neuropsychological testing was performed to confirm the absence of cognitive deficits. Subjects were also evaluated with the Clinical Dementia Rating scale, and only CDR 0 cases with test scores compatible with the age appropriate mean in all tests were included in the EC group[33]. The final elderly controls group consists of 12 subjects.

MCI cases were recruited in a large acute and intermediate care geriatric hospital. The same clinical screening and neuropsychological tests as for elderly controls were used. Individuals having a test score more than 1.5 standard deviation below the age appropriate mean in any of the tests and a Clinical Dementia Rating score of 0.5 but no dementia, were diagnosed as possible MCI. 16 (55%) of the original MCI cases demonstrated significant cognitive decline and constituted the progressive MCI

group (mean age: 82.8 ± 5.4 years)[33]. The final Progressive MCI group consisted of 14 subjects.

6.4.2 Procedure

The subjects were comfortably seated and watched a computer-controlled display screen at a distance of 57 cm. They viewed pseudo-random sequences of consonant and vowels common to the French alphabet, and pressed a computer-controlled button with their right index finger as soon as a target appeared (response trials). For non-target stimuli, no motor response was required (no-response trials). Stimuli consisted of white letters, Arial font ($2^\circ \times 2.5^\circ$ visual angle), with 10% grey noise, embedded in a 50% random noise grey rectangular background patch ($6^\circ \times 6.7^\circ$ visual angle). They were presented in the center of the screen for 0.5 s, separated by 5 s intervals (onset to onset) during which a dot helped the subjects to maintain fixation. Three working memory tasks were tested, in which one third of the stimuli were targets. In the detection task, sequential letters (non-target) or background patches without letters (target) were presented. Each task was tested in three blocks composed of 30 sequential stimuli each, adding up to 90 trials per task (21 response trials, 69 no-response trials).

We used all the available trials from the detection task in our analysis. We used a portion of each trial for our analysis, starting 1000ms before the stimulus presentation and lasts for 3000ms after the stimulus. The vertical axis in the TF illustrations spans frequencies from 1Hz to 14Hz in a logarithmic scale.

For the analysis of all subjects in this experiment we evaluated only the stable components, which are the most relevant as also discussed before along with the first experiment.

6.5 Evaluation of Components of 2nd Experiment

6.5.1 Control Subjects

Elderly control subjects' present phase-locked activity in delta, theta and alpha bands in the period 200 to 300ms after the event. Theta and alpha phase-locked activity precedes that in delta band in all components. In essence, theta band seems to be spread in a larger temporal interval following the event. More specifically alpha or delta phase-locked activity occurs around 200ms, whereas delta PIC components peak around 300ms after the event. All phase locked theta and delta components present power increase compared to prestimulus, which appears as an increase in the ERD/S measure. In many cases, delta and theta activations are coupled into a single component like in Figure 6.41a and 1d. These results are directly compared with other findings indicating power increase in delta and theta bands related to working memory functions. Theta ERS response has been associated in many studies with working memory processes[71].

Theta band induced activity was observed 1500ms after the event. Since the PsIC measure is normalized according to the maximum energy, we evaluated theta and alpha activity separately from delta. Delta band presents high energy which masks the theta and alpha activations when the component captures mixed activity, making the evaluation of theta band difficult. There was an increase in the power of the theta induced activity post-stimulus, as expressed in the ERD/S measure. An example can be reviewed in Figure 6.4c. As induced theta we treat the components that do not present significant PIC measure at the same region where strong PsIC is identified.

Alpha-activity components generally display a power decrease in the time period of 250ms to 350ms after the event. In fact, alpha activity can be distinguished into three groups. a) There is alpha oscillatory activity that synchronizes its phase right after the event and present high phase locking coherence at 200ms, right before desynchronizing its power. Examples are shown in Figures 1b (early alpha component), 1c, 1e and 1f. The phase locking effect in alpha oscillatory activity does not necessarily display significant increase in power compared to pre-stimulus. Phase locking without power increase appears in Figure 6.4b and 1c, revealing phase synchronization of ongoing oscillations after the event. Alternatively, the power increase in Figure 6.4 1e and 1f indicates a slightly different, additive nature of these components that contribute to an evoked part of ERP. b) Alpha oscillatory activity pre-stimulus, with power decrease after the event but no phase locking before or after the event, is present in control components as in Figure 6.4 1g. This kind of activity, which recovers after some time, forms the main contribution to the alpha desynchronization in the ERD effect. c) Finally there exist additive non phase locked (or partially phase-locked) alpha activations, such as in Figure 1a and 1b (late alpha activation), which also display an instantaneous increase in power in the ERD/S measure.

Control subjects display no alpha ERS in the 0-1500ms interval after the event, which is in line with other studies in 0-back experiments. Alternatively, alpha desynchronization (ERD) is known to correlate with semantic memory performance in normal subjects and is also correlated with memory performance[22]. It was reported that alpha phase locking is related to attention, perception, and memory process. Higher alpha phase locking seems to be related to good perception and memory performance in normal subjects [71]. Our findings support such findings, since control subjects present alpha phase locked activity and strong alpha ERD after the event. Finally, late alpha induced oscillations were found in control subjects after the desynchronization effect. This late alpha induced activity can be seen in Figure 6.4a and Figure 6.4b and is associated with the working memory activation as a response to the target presented.

In terms of the spatial localization of activations, components that present strong delta phase-locked activity exhibit posterior topography (Figure 6.4 1d and 1e), as in the corresponding components of the 1st experiment. Theta phase-locked components present parietal and occipital topography as can be observed in Figure 6.41b, 1d and 1e. Theta induced (non phase-locked) activity displayed mainly frontal topography as in Figure 6.41c. Alpha early phase-locked (Figure 6.41b, 1e and 1f) activity appears with posterior localization, whereas late additive (partially phase-locked) activation shows frontal localization (Figure 6.41a). Induced alpha oscillations displayed mainly a central topography (Figure 6.41g), similar to the previous experiment.

6.5.2 PMCI Subjects

Progressive MCI subjects display decreased phase locking in delta and theta bands. The complete statistics can be observed in **Error! Reference source not found.** and **Error! Reference source not found.**. Activations follow the same pattern as in controls, with delta band presenting later evoked oscillations. Theta phase-locked activations occur at 200ms post-stimulus and precede delta band activations that occur at 350ms post-stimulus on the average. Delta phase-locked activations are centered in higher frequencies at 2-3Hz than in controls. Recall that controls also present induced delta activity in lower delta, which does not appear in PMCI subjects. Delta and theta power found to be significantly lower in PMCI than in control subjects. This finding is in agreement with other studies that attribute theta as the main band related to memory functions. Reduced theta phase-locking and power in Progressive MCI subjects can be attributed to the diminished memory functionality due to MCI pathology[33].

A major difference between PMCI and control subjects is the absence of higher alpha phase locked activations in most subjects. This is in line with studies that report lower peak of alpha power in progressive MCI and AD subjects. Furthermore, induced alpha activations (Figure 6.52c, 2e, 2f and 2h) formulate the ERD with little or no power reorganization after the event in contrast to controls. This results in significantly higher PIC values and corresponding alpha phase locked energy in control subjects than PMCI. Another important finding is that ERD of alpha induced oscillations lasts longer compared to controls, about 300ms, compared. This can be explained by the fact that the duration of alpha ERD response has been found to increase with increasing memory load. Memory deficits in PMCI subjects would result in greater memory load, thus resulting in increased duration of alpha ERD. Also, the alpha decrease in PMCI subjects was found significantly higher than in controls. This is also in agreement with the literature, where increased ERD has been observed in higher load memory tasks compared to simpler ones. Another interesting finding is that PMCI subjects present reduced late alpha reactivity in contrast to control. This is in line with findings that suggest that higher frequencies

are sensitive to mental decline and associate late alpha activity with memory processes.

The topography of the different activations is similar with the controls. Theta phase locked activations present posterior topography while theta induced activity is much weaker than in controls and has frontal topography. Delta phase locked activity is also present in frontal locations. Alpha induced activity also presented a central and posterior topography.

6.5.3 Discussion of activation patterns for the 2nd experiment.

During a memory task different brain mechanisms activate that allow recognizing and processing of the incoming information. We analyzed a 0-back, visual detection, working memory experiment performed by control and PMCI subjects.

Phase locked activity that can be attributed to the stimulus perception and attention was detected in both groups. More specifically, phase locked delta theta and alpha activations were found in the period following the stimulus presentation in both groups and they all present parietal topography in both cases being affected by the nature of the stimulus (visual with parietal response). This is in line with reports suggesting that the superposition of phase-locked activity in delta to alpha frequency contribute to the generation of the average ERP[71]. Delta and theta phase locked activations were accompanied by a power increase, reflected in the ERD/S measure, in the specific bands supporting its stimulus related characterization, since such power increase has been reported in many other studies[18, 72].

PMCI subjects presented reduced phase-locking in delta and theta bands but with no significant power differences. This can be attributed to the limited capacities of PMCI subjects to attention, leading in response variations from trial to trial. Alpha band also presented reduced phase-locking with reduced energy compared to control subjects. This is related to stimulus processing and attention problems related to PMCI, since higher bands tend to be more sensitive to mental decline[27].

Theta induced oscillations (as expressed by the PsIC measure with no PIC values) were found with central latencies near 1200ms. Such late theta activity also presented power increase (ERS) and can be attributed to focused attention and working memory [33]. PMCI subjects presented reduced theta induced power compared to controls which is related to reduced attention and memory functionality.

Alpha presented significant power decrease following the stimulus, expressed as ERD, which was found to be significantly more in the AD subjects with longer duration. Alpha ERD is known to correlate with memory performance and mental activation. Increased alpha ERD in PMCI subjects suggests use of increased cognitive

resources for completing the task[73]. Increased ERD has been also reported in patients with dementia and AD in other studies[73]. Alpha induced oscillations presented a late ERS in control subjects which was reduced or completely absent in PMCI subjects. Such late alpha induced activity is related to working memory maintenance for further tasks [33]. The lack of alpha power increase in PMCI subjects could be related to memory deficits and can also be associated with the reduced energy at late induced theta: the later relates to reduced engagements of working memory[33].

Our results are in agreement and extend a previous study made in the same dataset[33]. This study used the technique described in section 3.3.5 to separate the induced from global energy and perform their analysis in the electrodes. The authors report that evoked theta activity is located at posterior locations while induced theta activity is located in frontal regions. Furthermore their findings report that induced theta activity emerges later in time than the theta evoked activation. Global theta energy did not provide any differentiation in energy between the two groups because of the mixed activity. In contrast, they find significant reduction in the power of induced theta activations. These findings are in agreement with our results where we found no significant difference in the evoked theta energy between the two groups but we found significant decrease in theta induced energy. The PsIC measure shows that the induced delta activation lags in time the evoked theta, expressed by the PIC measure. Furthermore, the derived topography in our analysis agrees with that of [33] for both induced and evoked theta.

Our methodology enables the separation of evoked and induced activations that can be observed and evaluated in parallel in terms of their activation content and in terms of their topography. Besides the study of theta activation, it enables the analysis of other bands, such as alpha, where we also detect evoked and induced activity of different nature and topography.

Another advantage is that we make no assumptions about the generation and manifestation of the evoked activity. Removing the mean activity in a specific band, as in [33], assumes that the evoked activation is generated in each trial at the same latency and with equal amplitude. Many studies suggest that phase locked activations emerge with some jitter in latency from trial to trial and are affected by fatigue and the level of attention[74]. This means that by removing the mean activation in a certain band leaves residual energy which cannot be anticipated beforehand.

6.6 Comparison between the two datasets.

Progressive MCI has been considered as a transitional stage in the pathogenesis of Alzheimer. Comparing the activations between the AD and the PMCI subjects in the

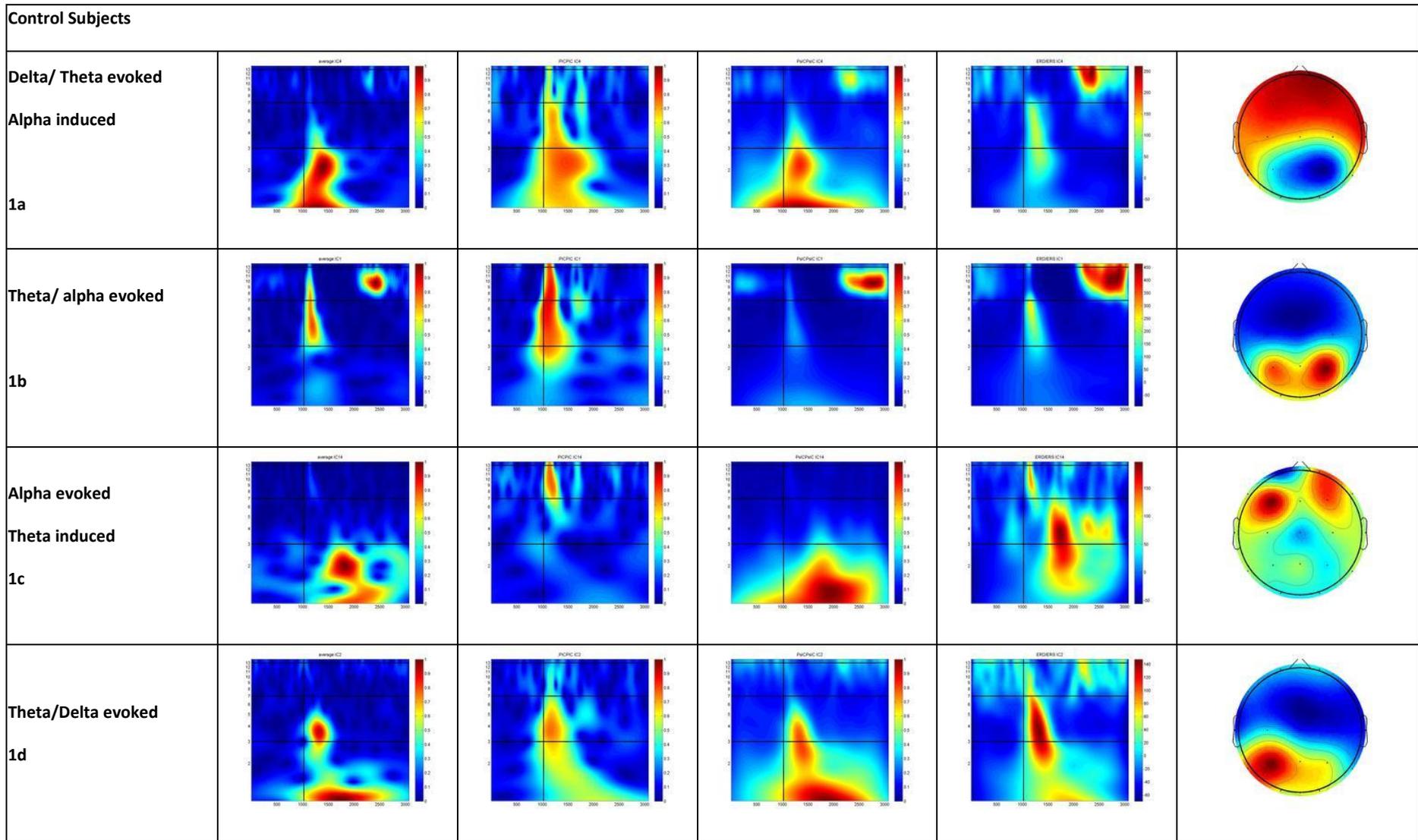
two datasets we considered could provide interesting insights. Although the two tasks are different, the former is an auditory oddball task and the latter a visual memory task, we can argue that there exist basic underlying mechanisms that take place during both experiments.

For example the pattern of activation for delta and theta bands displays the same behavior. In both cases theta phase locked activations precede delta band phase locked activations. In both cases theta and delta bands present lower phase locking values than the controls. These two types of activations are probably modulated by the stimulus presentation and are related to stimulus processing. In both datasets theta phase locked activity presented a post-stimulus power increase. The difference is that in the case of P300 data AD subjects presented increased theta reactivity, which should be attributed to the medication used (cholinesterase inhibitors) [26]. Another difference is the topography of the activations. Phase-locked theta band in the P300 experiment appears to emerge at frontal locations while in the 0-back experiment appears at posterior sites. This is a difference due to the stimulus presented in each case. The visual processing regions are located in the posterior regions of the brain so it is reasonable that visual stimulus will activate these sites.

Late induced theta appears mostly frontally, which is in accordance to working memory related to the activity [5]. Early evoked theta in the second experiment appears posterior in response to the visual stimulus. Recall that evoked theta in the first experiment appears frontal in response auditory stimulus. Similar topography in the two experiments is observed for evoked alpha which acts like theta in relation to stimulus. The induced alpha which is associated to ERD and semantic processing/cognition appears at the same parietal-central location in both experiments.

Alpha power is lower in both MCI and Alzheimer subjects with AD and PMCI subjects lacking late alpha activations and presenting reduced alpha phase locking. In the 0-back case where the experiment is related strongly to working memory PMCI subjects significant differences were found in the activity of early phase locked. The AD subjects do not present significant differentiations in the early alpha band, probably due to the nature of the experiment that requires less mental effort.

Overall, in both cases alpha band activity seems to be increasingly affected by mental decline and could be used as a marker for assessing the status of a subject. Theta band differentiates in both cases PMCI/ AD subjects from control, although the drug effect reverses the way this difference is expressed (increase for AD, decrease for PMCI).



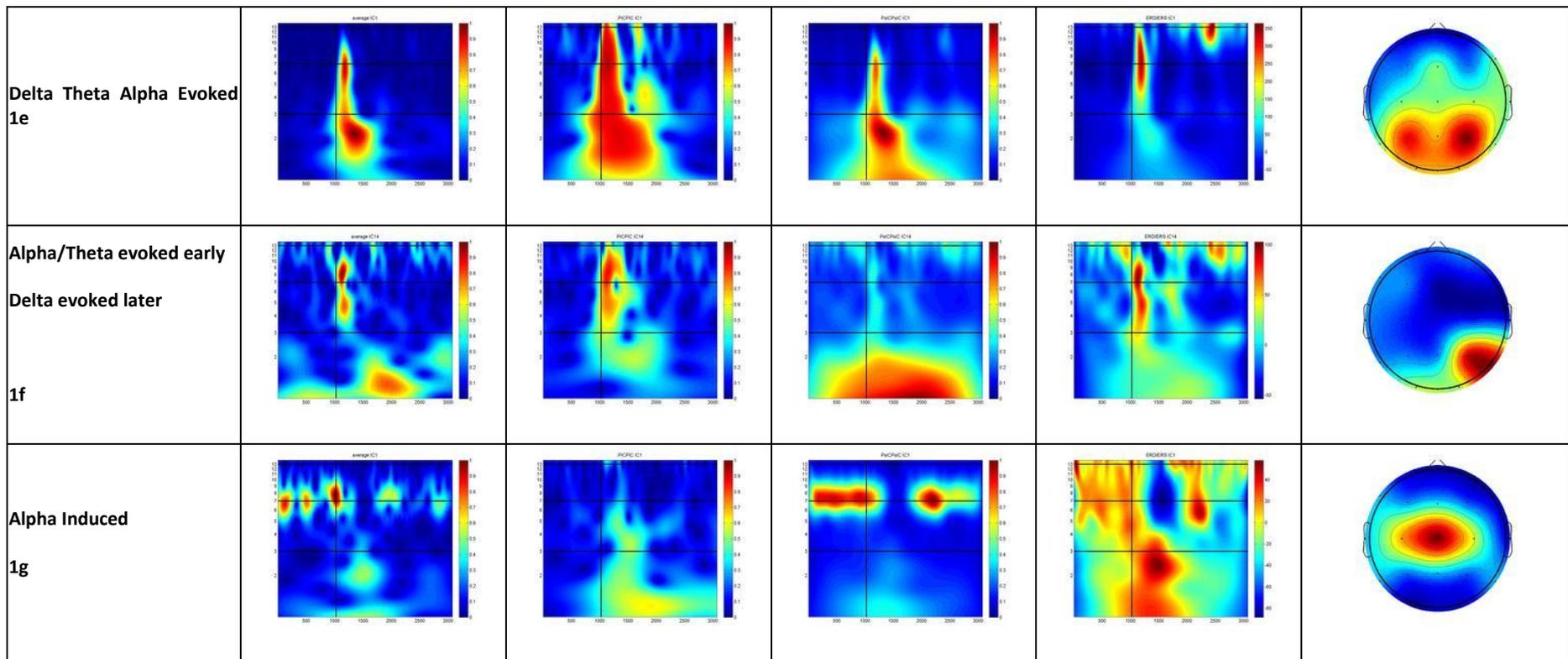
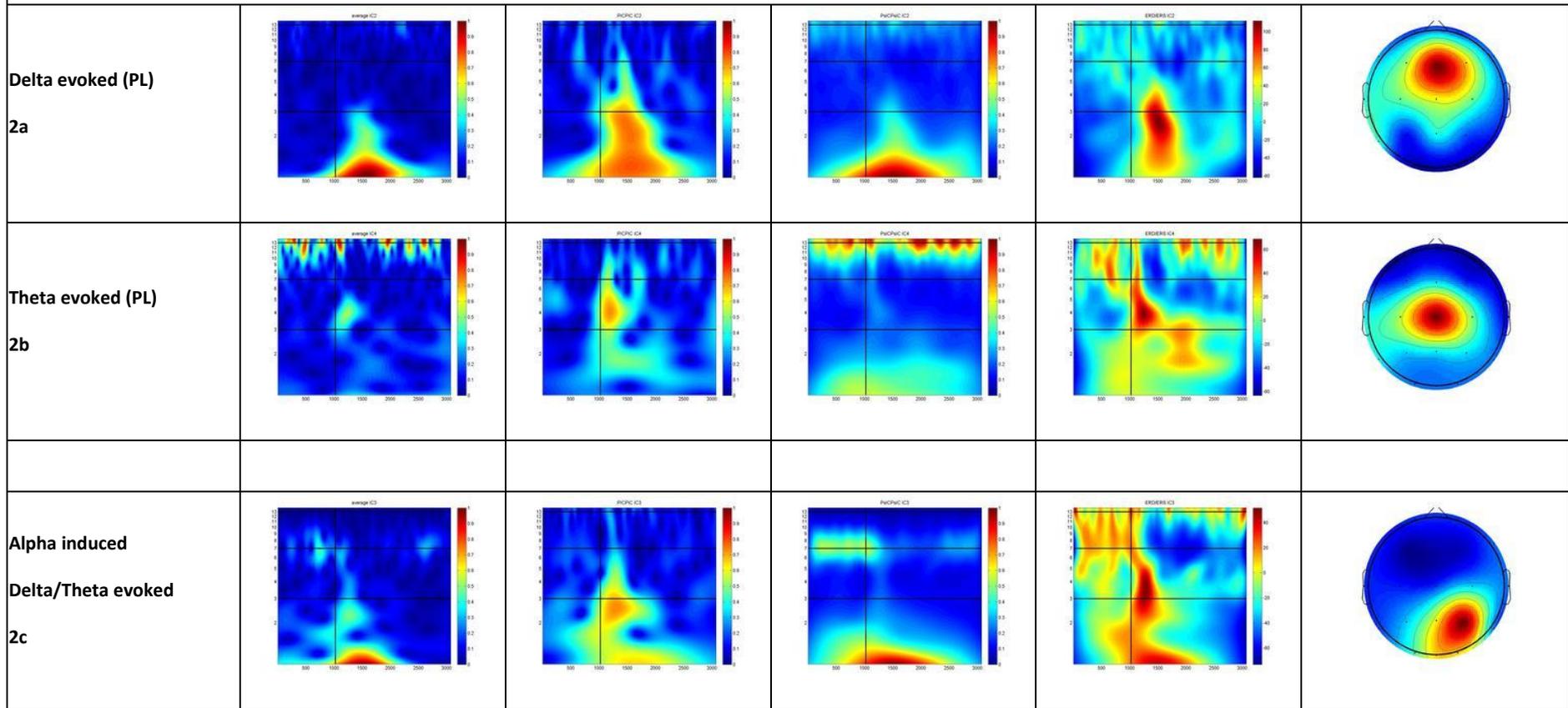


Figure 6.4: Results for control subjects in the visual memory task. The first column provides short description of the activations that take place.

PMCI results



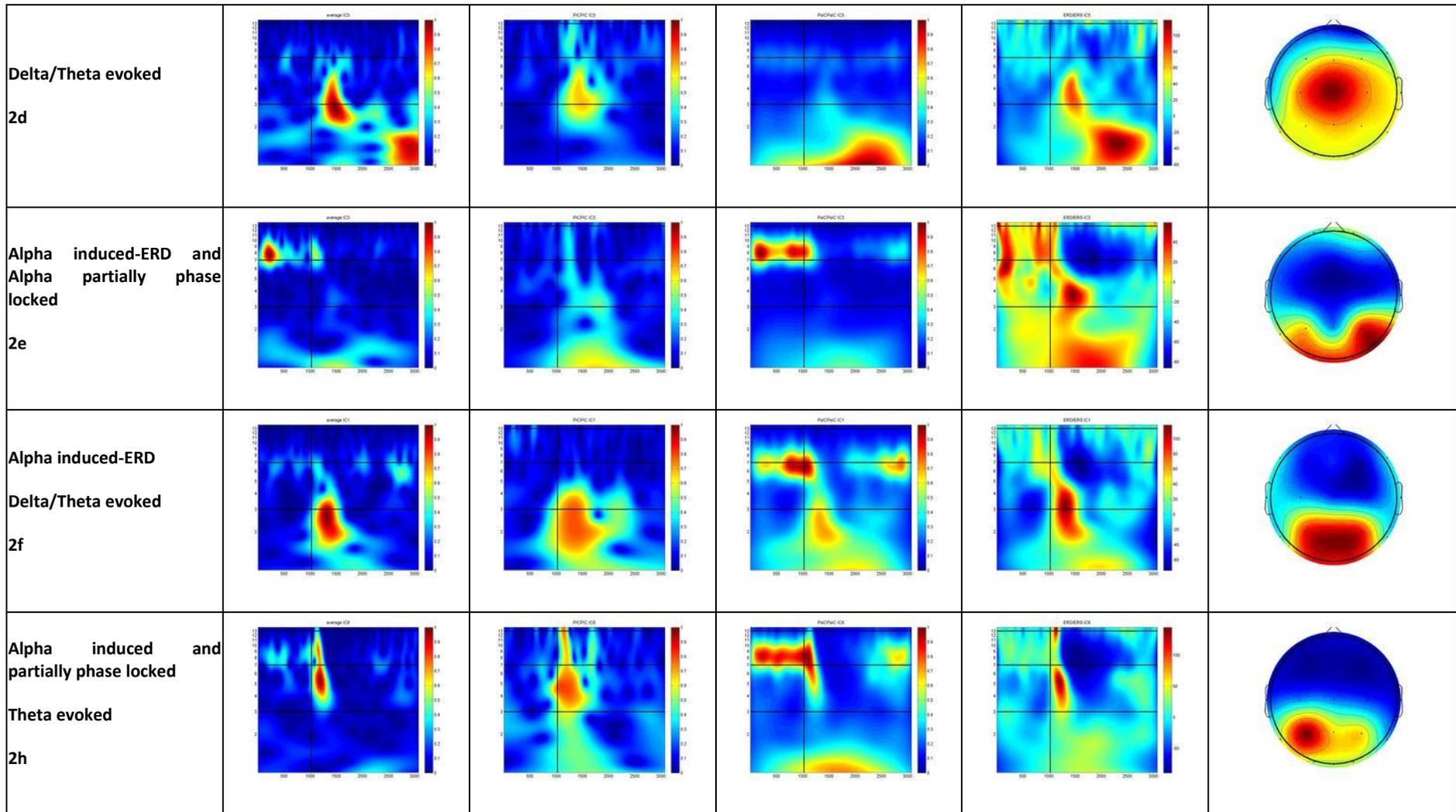


Figure 6.5: Results for PMCI subjects in the visual memory task. The first column provides short description of the activations that take place.

Table 6-2

Table 1a. Phase locked statistics			
	Controls PIC	PMCI PIC	P-Value: $\alpha=0.05$
Band	Statistics	Statistics	
Delta	Central time= 473 Central freq =3.4 Mean PIC energy = 349 Mean PIC values = 0.60	Central time = 339.5 Central freq =3.0 Mean PIC energy = 303 Mean PIC values = 0.51	No signif difference $p=0.0216(\text{Normal})$
Theta	Central time = 236 Central freq =6 Mean PIC energy = 510 Mean PIC values = 0.72	Central time = 269 Central freq = 5.2 Mean PIC energy = 490 Mean PIC values = 0.55	No signif difference $P = 0.004(\text{Normal})$
Alpha	Central time = 296 Central freq =11.3 Mean PIC energy = 205 Mean PIC values = 0.65	Central time = 193 Central freq =10.55 Mean PIC energy = 105 Mean PIC values = 0.47	$P = 0.0096(\text{normal})$ $P = 0.0312(\text{normal})$

Table 1b. Phase Shift Statistics			
	Controls PsIC	PMCI PsIC	P-Value: $\alpha=0.05$
Band	Statistics	Statistics	Alpha
Delta	Central time= 629 Central freq = 3 Mean PsIC energy = 550	Central time = 474 Central freq =3.2 Mean PsIC energy = 529	No signif difference
Theta	Central time = 1249 Central freq =6 Mean PsIC energy = 810	Central time = 1266 Central freq = 4.2 Mean PsIC energy = 470	$P = 0.008(\text{normal})$
Early Alpha	Central time = 229.70 Central freq =12.12 Mean PsIC energy = 322	Central time = 302 Central freq =9.8 Mean PsIC energy = 307	No signif difference
Late Alpha	Central time = 1036 Central Freq = 11.3 Mean PsIC energy = 267	Central time = 1300 Central Freq = 10.5 Mean PsIC energy = 185	$p=0.032(\text{Normal})$

6.7 Discussion

Our study was able to identify and characterize the intertrial coherence of independent components involved in the multiple trials of experiments. We explored a recent method for evaluating the stability of the ICA components. The use of derived stable components for the characterization of the ERP content, presents the advantage that the statistics produced by this analysis can be reproduced safely. On top of that, the number of reliable components is usually smaller than the number of total extracted independent components, so that the assessment of activity of stable components becomes faster and easier.

We considered a population of nine control and nine AD subjects performing an auditory oddball experiment and derived activations that displayed evoked and induced activity related to the stimulus. Based on the extracted information we calculated group statistics related to the energy of the significant components. Our analysis was able to identify phase-locked and induced oscillations related to the task and we were able to find differences between the two groups. More specifically, the AD subjects presented reduced phase-locking in delta and theta bands, which is related to the stimulus perception, as well as absence of late alpha induced activations in the AD subjects. Our findings are in agreement with previous research showing that AD subjects present reduced phase-locking in task-related activities and reduced energy in the alpha band[13, 69].

We proceeded with the analysis of a second experiment involving control and progressive MCI populations performing a visual 0-back (detection) memory experiment. Our methodology on control subjects revealed early phase-locked activity in delta, theta and alpha bands, with posterior topography that are directly related to stimulus processing. Theta induced activity was detected later in latency with frontal topography, which implies an association to working memory processes. Alpha-band induced oscillations displayed significant reduction in post-stimulus energy shortly after the stimulus (ERD), which re-emerged at later latencies. Alpha ERD is associated with memory load, increased attention and performance while late alpha ERS is associated with maintenance of working memory processes.

PMCI subjects presented significantly lower phase-locking in delta and theta bands, without differences in energy at the same regions. They also displayed a reduction in induced theta energy at frontal locations. These findings are consistent with previous studies in the same dataset that reported no significant differences in theta global power between control and PMCI subjects[33]. Instead, they reported significant difference only in theta induced energy at frontal locations, which is in accordance with our findings. In this study, the method described in section 3.3.5 was used in order to isolate and subtract induced from total activity. Our methodology was able

to separate induced and evoked activity into separate components and evaluate them using the proposed measures, without subtracting any portion of the signal as the evoked part. Our proposed method revealed differentiations in early evoked and late induced alpha band. Previous studies reported difficulties in evaluation of channel activations in alpha band, due to possible overlapping effects [33]. Using ICA we managed to decouple the different alpha activities and derive meaningful differentiations between the groups. Specifically, we found reduced phase-locking in early alpha-evoked activity and reduced late alpha-induced activity in PMCI subjects.

An attempt to compare activations from the two datasets revealed some common patterns. Despite the difference in pathology and type of event, our analysis showed similarities in evoked activations that are related to the stimulus. Activations of phase-locked delta and theta bands display a common pattern of activation, with theta band preceding delta in latency. Controls from both datasets displayed a common pattern in alpha induced activations, which were desynchronized after the event presentation and then re-emerging at later latencies. PMCI and AD subjects displayed reduced phase-locking in their evoked activations. These activations in theta and alpha bands emerged at different locations due to the different type of stimulus (acoustic versus visual). Alpha band was also affected by pathology with late alpha activity being absent or diminished in both pathologic groups. Overall, the phase-locking of evoked activations and the reduction of induced alpha/theta activity seem to be affected most by the mental decline and could be possibly explored further as markers to diagnose and differentiate aging pathologies.

The results of this study indicate that the proposed component analysis and framework is able to depict the synchronized activations during a certain mental task (like the working memory). As such, it can efficiently reveal and quantify group (as well as individualized) differences in pathology (AD) populations. Of course the methodology should be extended to a larger population sample and other pathology groups, as to validate the initial results presented.

Chapter 7. Conclusions

In this work we presented a novel technique and useful measures for the analysis of event-related potentials. The measures presented can be used to discriminate between different activations that take place during an ERP experiment. Used in conjunction with techniques as independent component analysis and time-frequency transforms, we were able to separate, characterize and evaluate multiple brain activations that sum up for the generation of the ERP.

We considered the stability of a popular technique in independent component analysis, namely Infomax. We presented results in simulated and real EEG data that demonstrate the effect of sample size and initial weights on the stability of ICA results. We explored a methodology to evaluate the stability of the independent components before used in analysis and developed an analysis framework that takes under consideration the stable components. The method uses split-half comparison to discriminate stable components and does not require extreme computational resources, thus becoming attractive for analysis of large datasets. Besides the robustness against spurious activity, the identification and use of stable components has the additional advantage of reducing the number of components to be evaluated.

We applied the proposed technique in two datasets with subjects suffering from pathologies related to aging. The first dataset consisted from control and Alzheimer subjects performing an auditory oddball experiment. The second dataset consisted from control and progressive MCI subjects performing a visual memory test. We were able to identify activations related to the given task for each group. AD subjects presented reduced phase-locking in delta and theta bands, related to the stimulus perception, as well as absence of late alpha induced activations. The latter effect is related to increased attention and retention of mental resources and its absence can be attributed to the Alzheimer pathology. Theta phase-locked activity in AD subjects presented increased energy in comparison to controls. This finding cannot be attributed solely to the pathology, since the AD subjects were under medication that has been reported to increase theta activity.

MCI subjects of the second dataset displayed reduced theta induced activity in frontal regions, which is related to memory deficits. They also displayed reduced phase-locking of the evoked activity in all bands, but showed no significant difference in the power levels. Alpha band was furthermore affected by the pathology displaying reduced late alpha, induced activity. Using our technique, we were able to derive interesting results that support and supplement previous studies. We were able to separate induced from evoked activity in independent components and reveal differences in the alpha and theta bands that are otherwise

masked due to overlapping activations in channels. The proposed methodology provides a compact analysis tool that operates on the entire EEG collection of many trials without the need of subtracting any portion of the signal (such as the average), which might affect the analyzed content.

Finally, our attempt to compare the results from the two pathologies provided some common patterns, which are expected despite the difference in the task performed. More specifically, we could identify a common pattern in the activations related to stimulus perception with evoked theta being activated first followed in latency by delta activations. Alpha band also displays some common patterns of induced activation, with power decrease (ERD de-synchronization) following the stimulus and a late power rebound related to working memory maintenance. The detected differences in the topography of evoked activations are attributed to the different type of task. In general, mental deterioration seems to increasingly affect the phase-locking of evoked activations and the power of induced oscillations in theta and alpha bands.

7.1 Future Work

Further work involves the application of our methodology in more difficult mental tasks performed by PMCI and control subjects. We can also extend our research in stable MCI subjects in order to find patterns that could describe the evolution of the pathology and could be used as biomarkers of AD. We expect that the differentiations will be more apparent in more difficult mental tasks that require more attention and heavier memory load. In particular, it would be interesting to consider the alpha band, for which contradictory results have been reported in relation to memory load.

Even though we were able to separate evoked from induced activity by means of independent component analysis and reassert previous studies, we can also compare with the method of separating induced from evoked activity as described in 3.3.5 and apply ICA on the resulting dataset without the evoked part. In short, it would be interesting to replicate the study in [33] and extend it by applying our techniques in the global and induced set, as to evaluate the degree of complementarity of the methods.

Another aspect that needs further investigation is the inter-trial behaviour of the different activations identified. Towards this direction, we can further evaluate the characterized components by means of intertrial entropy in order to assess activations that are indeed modulated by the stimulus or other mental functions.

References

1. Hendrikse, A.J., R.N.J. Veldhuis, and L.J. Spreeuwers, *Component ordering in independent component analysis based on data power*, in *Proceedings of the 28th Symposium on Information Theory in the Benelux*, R.N.J. Veldhuis and H.S. Cronie, Editors. 2007, Werkgemeenschap voor Informatie- en Communicatietechniek: Enschede, The Netherlands. p. 211-218.
2. Yeung, N., R. Bogacz, C.B. Holroyd, and J.D. Cohen, *Detection of synchronized oscillations in the electroencephalogram: An evaluation of methods*. *Psychophysiology*, 2004. **41**(6): p. 822-832.
3. Miwakeichi, F., et al., *Decomposing EEG data into space-time-frequency components using Parallel Factor Analysis*. *NeuroImage*, 2004. **22**(3): p. 1035-1045.
4. Heinze, H.J., T.F. Münte, and G.R. Mangun, *Cognitive electrophysiology*. 1994, Boston: Birkhäuser.
5. Tallon-Baudry, C. and O. Bertrand, *Oscillatory gamma activity in humans and its role in object representation*. *Trends in cognitive sciences*, 1999. **3**(4): p. 151-162.
6. Makeig, S., et al., *Dynamic Brain Sources of Visual Evoked Responses*. *Science*, 2002. **295**(5555): p. 690-694.
7. Delorme, A. and S. Makeig, *EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis*. *Journal of Neuroscience Methods*, 2004. **134**(1): p. 9-21.
8. Shannon, C.E., *A mathematical theory of communication*. *Bell System Technical Journal*, 1948. **27**(1): p. 379-423 and 623-656.
9. Yordanova, J., V. Kolev, and J. Polich, *P300 and alpha event-related desynchronization (ERD)*. *Psychophysiology*, 2001. **38**(1): p. 143-152.
10. Basar, E., *Brain function and oscillations. Brain oscillations. Principles and approaches*. Vol. I. 1998a, Berlin: Springer.
11. Jung, T.P., et al., *Imaging brain dynamics using independent component analysis*. *Proceedings of the IEEE*, 2001. **89**(7): p. 1107-1122.
12. Tallon-Baudry, C., O. Bertrand, C. Delpuech, and J. Pernier, *Stimulus Specificity of Phase-Locked and Non-Phase-Locked 40 Hz Visual Responses in Human*. *J. Neurosci.*, 1996. **16**(13): p. 4240-4249.
13. Klimesch, W., M. Doppelmayr, D. Röhlm, D. Pöllhuber, and W. Stadler, *Simultaneous desynchronization and synchronization of different alpha responses in the human electroencephalograph: a neglected paradox?* *Neurosci. Lett.*, 2000. **284**(1-2): p. 97-100.
14. Bernat, E.M., S.M. Malone, W.J. Williams, C.J. Patrick, and W.G. Iacono, *Decomposing delta, theta, and alpha time-frequency ERP activity from a visual oddball task using PCA*. *Int. J. Psychophysiol.*, 2007. **64**(1): p. 62-74.
15. Wyart, V. and C. Tallon-Baudry, *Neural Dissociation between Visual Awareness and Spatial Attention*. *J. Neurosci.*, 2008. **28**(10): p. 2667-2679.
16. Bennys, K., G. Rondouin, C. Vergnes, and J. Touchon, *Diagnostic value of quantitative EEG in Alzheimer's disease*. *Clin. Neurophysiol.*, 2001. **31**(3): p. 153-160.
17. Karrasch, M., M. Laine, P. Rapinoja, and C.M. Krause, *Effects of normal aging on event-related desynchronization/synchronization during a memory task in humans*. *Neurosci. Lett.*, 2004. **366**(1): p. 18-23.
18. Krause, C.M., et al., *The effects of memory load on event-related EEG desynchronization and synchronization*. *Clin. Neurophysiol.*, 2000. **111**(11): p. 2071-2078.
19. Demiralp, T., A. Ademoglu, M. Comerchero, and J. Polich, *Wavelet Analysis of P3a and P3b*. *Brain Topogr.*, 2001. **13**(4): p. 251-267.

20. Porjesz, B., et al., *The utility of neurophysiological markers in the study of alcoholism*. Clin. Neurophysiol., 2005. **116**(5): p. 993-1018.
21. Klimesch, W., H. Russegger, M. Doppelmayr, and T. Pachinger, *A method for the calculation of induced band power: implications for the significance of brain oscillations*. Electroencephalogr. Clin. Neurophysiol., 1998. **108**(2): p. 123-130.
22. Klimesch, W., *Memory processes, brain oscillations and EEG synchronization*. Int. J. Psychophysiol., 1996. **24**(1-2): p. 61-100.
23. Krause, C.M., A. Heikki Lang, M. Laine, M. Kuusisto, and B. Pörn, *Event-related. EEG desynchronization and synchronization during an auditory memory task*. Electroencephalogr. Clin. Neurophysiol., 1996. **98**(4): p. 319-326.
24. Polich, J., *Updating P300: An integrative theory of P3a and P3b*. Clin. Neurophysiol., 2007. **118**(10): p. 2128-2148.
25. Intriligator, J. and J. Polich, *On the relationship between background EEG and the P300 event-related potential*. Biol. Psychol., 1994. **37**(3): p. 207-218.
26. Yener, G.G., B. Güntekin, A. Öniz, and E. Basar, *Increased frontal phase-locking of event-related theta oscillations in Alzheimer patients treated with cholinesterase inhibitors*. Int. J. Psychophysiol., 2007. **64**(1): p. 46-52.
27. Karrasch, M., et al., *Brain oscillatory responses to an auditory-verbal working memory task in mild cognitive impairment and Alzheimer's disease*. Int. J. Psychophysiol., 2006. **59**(2): p. 168-178.
28. Jeong, J., *EEG dynamics in patients with Alzheimer's disease*. Clin. Neurophysiol., 2004. **115**(7): p. 1490-1505.
29. Berendse, H.W., J.P.A. Verbunt, P. Scheltens, B.W. van Dijk, and E.J. Jonkman, *Magnetoencephalographic analysis of cortical activity in Alzheimer's disease: a pilot study*. Clin. Neurophysiol., 2000. **111**(4): p. 604-612.
30. Jelic, V. and A. Nordberg, *Early diagnosis of Alzheimer disease with positron emission tomography*. Alzheimer Dis. Assoc. Disord., 2000. **14**(SUPPL. 1): p. S109-S113.
31. Zervakis, M., K. Michalopoulos, V. Iordanidou, and V. Sakkalis, *Intertrial coherence and causal interaction among independent EEG components*. Journal of Neuroscience Methods, 2011. **197**(2): p. 302-314.
32. Onton, J. and S. Makeig, *Information-based modeling of event-related brain dynamics*, in *Prog. Brain Res.* 2006, Elsevier. p. 99-120.
33. Deiber, M.-P., et al., *Abnormal-induced theta activity supports early directed-attention network deficits in progressive MCI*. Neurobiology of aging, 2009. **30**(9): p. 1444-1452.
34. Onton, J., A. Delorme, and S. Makeig, *Frontal midline EEG dynamics during working memory*. NeuroImage, 2005. **27**(2): p. 341-356.
35. Barrie, J., et al., *The independent components of auditory P300 and CNV evoked potentials derived from single-trial recordings*. Physiol. Meas., 2007. **28**(8): p. 745.
36. Graps, A., *An introduction to wavelets*. Computational Science & Engineering, IEEE, 1995. **2**(2): p. 50-61.
37. Daubechies, I., *The wavelet transform, time-frequency localization and signal analysis*. IEEE Transactions on Information Theory, 1990. **36**(5): p. 961-1005.
38. Cohen, L., *Time-frequency distributions-a review*. Proceedings of the IEEE, 1989. **77**(7): p. 941-981.
39. Groppe, D.M., S. Makeig, and M. Kutas, *Identifying reliable independent components via split-half comparisons*. NeuroImage, 2009. **45**(4): p. 1199-1211.

40. Klimesch, W., P. Sauseng, S. Hanslmayr, W. Gruber, and R. Freunberger, *Event-related phase reorganization may explain evoked neural dynamics*. *Neurosci. Biobehav. Rev.*, 2007. **31**(7): p. 1003-1016.
41. Martínez-Montes, E., et al., *Exploring event-related brain dynamics with tests on complex valued time–frequency representations*. *Stat. Med.*, 2008. **27**(15): p. 2922-2947.
42. Pfurtscheller, G. and F.H. Lopes da Silva, *Event-related EEG/MEG synchronization and desynchronization: basic principles*. *Clin. Neurophysiol.*, 1999. **110**(11): p. 1842-1857.
43. Torrence, C. and G.P. Compo, *A Practical Guide to Wavelet Analysis*. *Bulletin of the American Meteorological Society*, 1998. **79**(1): p. 61-78.
44. Cardoso, J.-F., *Dependence, correlation and Gaussianity in independent component analysis*. *J. Mach. Learn. Res.*, 2004. **4**(7-8): p. 1177-1203.
45. Jolliffe, I.T., *Principal Component Analysis*. 2nd ed. ed. 2002, New York, NY: Springer-Verlag.
46. Hyvärinen, A., J. Särelä, and R. Vigário. *Bumps and Spikes: Artefacts Generated by Independent Component Analysis with Insufficient Sample Size*. in *Int. Workshop on Independent Component Analysis and Blind Signal Separation (ICA'99)*. 1999.
47. Haykin, S.S., *Neural networks and learning machines*. 2009: Prentice Hall.
48. Linsker, R., *An application of the principle of maximum information preservation to linear systems*, in *Advances in neural information processing systems 1*. 1989, Morgan Kaufmann Publishers Inc. p. 186-194.
49. Laughlin, S., *A simple coding procedure enhances a neuron's information capacity*. *Z Naturforsch C*, 1981. **36**(9-10): p. 910-2.
50. Nadal, J. and N. Parga, *Non linear neurons in the low noise limit : a factorial code maximizes information transfer*. 1994.
51. Jung, T.-P., et al., *Extended ICA removes artifacts from electroencephalographic recordings*, in *Proceedings of the 1997 conference on Advances in neural information processing systems 10*. 1998, MIT Press: Denver, Colorado, United States. p. 894-900.
52. McKeown, M.J., et al., *Analysis of fMRI data by blind separation into independent spatial components*. *Human Brain Mapping*, 1998. **6**(3): p. 160-188.
53. Jutten, J.H.a.C., *Space or time processing by neural network models*. *Proc. AIP Conference: Neural Networks for Computing*, 1986. **151**.
54. Comon, P., *Independent component analysis, a new concept?* *Signal Process.*, 1994. **36**(3): p. 287-314.
55. Bell, A.J. and T.J. Sejnowski, *An information-maximization approach to blind separation and blind deconvolution*. *Neural Comput.*, 1995. **7**(6): p. 1129-1159.
56. James, C.J. and D. D'Alimonte. *Tracking multisource brain activity in multi-channel EEG with a probabilistic model through ICA*. in *Mediterranean Conf. on Medical and Biological Engineering - MEDICON2004*. 2004. Ischia, Italy.
57. Delorme, A., S. Makeig, M. Fabre-Thorpe, and T. Sejnowski, *From single-trial EEG to brain area dynamics*. *Neurocomputing*, 2002. **44-46**: p. 1057-1064.
58. Meinecke, F., A. Ziehe, M. Kawanabe, and K.R. Muller, *A resampling approach to estimate the stability of one-dimensional or multidimensional independent components*. *Biomedical Engineering, IEEE Transactions on*, 2002. **49**(12): p. 1514-1525.
59. Efron, B. and R. Tibshirani, *An introduction to the bootstrap*. 1993: Chapman & Hall.

60. Hyvarinen, A., *Fast and robust fixed-point algorithms for independent component analysis*. Neural Networks, IEEE Transactions on, 1999. **10**(3): p. 626-634.
61. Amari, S.-i., *Natural Gradient Learning for Over- and Under-Complete Bases in ICA*. Neural Computation, 1999. **11**(8): p. 1875-1883.
62. Shao, J. and D. Tu, *The jackknife and bootstrap*. 1995: Springer Verlag.
63. Himberg, J. and A. Hyvarinen. *Icasso: software for investigating the reliability of ICA estimates by clustering and visualization*. in *Neural Networks for Signal Processing, 2003. NNSP'03. 2003 IEEE 13th Workshop on*. 2003.
64. Maulik, U. and S. Bandyopadhyay, *Performance evaluation of some clustering algorithms and validity indices*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2002. **24**(12): p. 1650-1654.
65. Bezdek, J.C. and N.R. Pal, *Some new indexes of cluster validity*. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 1998. **28**(3): p. 301-315.
66. Ware, C., *Information Visualization: Perception for Design*. 2004: Morgan Kaufman.
67. Venna, J. and S. Kaski, *Neighborhood Preservation in Nonlinear Projection Methods: An Experimental Study*, in *Proceedings of the International Conference on Artificial Neural Networks*. 2001, Springer-Verlag. p. 485-491.
68. Zar, J.H., *Biostatistical Analysis*. 4th ed. Vol. 4. ed. 1999, New Jersey: Prentice Hall. 929.
69. Hogan, M.J., G.R.J. Swanwick, J. Kaiser, M. Rowan, and B. Lawlor, *Memory-related EEG power and coherence reductions in mild Alzheimer's disease*. Int. J. Psychophysiol., 2003. **49**(2): p. 147-163.
70. Czigler, B., et al., *Quantitative EEG in early Alzheimer's disease patients -- Power spectrum and complexity features*. Int. J. Psychophysiol., 2008. **68**(1): p. 75-80.
71. Klimesch, W., et al., *Phase-locked alpha and theta oscillations generate the P1-N1 complex and are related to memory performance*. Cognitive Brain Research, 2004. **19**(3): p. 302-316.
72. Tesche, C.D. and J. Karhu, *Theta oscillations index human hippocampal activation during a working memory task*. Proceedings of the National Academy of Sciences, 2000. **97**(2): p. 919-924.
73. Babiloni, C., et al., *Alpha rhythms in mild demented during visual delayed choice reaction time tasks: A MEG study*. Brain Res. Bull., 2005. **65**(6): p. 457-470.
74. Makeig, S., S. Debener, J. Onton, and A. Delorme, *Mining event-related brain dynamics*. Trends in Cognitive Sciences, 2004. **8**(5): p. 204-210.