Adaptive techniques for the detection and localization of event related potentials from EEGs using reference signals.

Thesis submitted to Cardiff University in candidature for the degree of Doctor of Philosophy.

Loukianos Spyrou

Centre of Digital Signal Processing
Cardiff University
2008

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Summary

In this thesis we show the methods we developed for the detection and localisation of P300 signals from the electroencephalogram. We utilised signal processing theory in order to enhance the current methodology. The work done can be applied both to EEG averages and single trial EEG data. We developed a variety of methods dealing with the extraction of the P300 and its subcomponents using independent component analysis and least squares. Moreover, we developed novel localisation methods that localise the desired P300 subcomponent from EEG data. Throughout the thesis the main idea was the use of reference signals, which describe the prior information we have about the sources of interest.

ABSTRACT

The main objective of this thesis is to utilize adaptive techniques, namely blind source separation (BSS), least squares (LS) and spatial filtering, in order to extract the P300 subcomponents from the electroencephalogram (EEG) with greater accuracy than the traditional methods.

The first topic of research, is the development of constrained BSS and blind signal extraction (BSE) algorithms, to enhance the estimation of the conventional BSS and BSE algorithms. In these methods we use reference signals as prior information, obtained from real EEG data, to aid BSS and BSE in the extraction of the P300 subcomponents. Although, this method exhibits very good behaviour in terms of EEG averaged data, its performance degrades when applied to single trial data, which is the response of the brain after one single stimulus.

The second topic deals with single trial EEG data and is based on least squares. Again, we use reference signals to describe the prior knowledge of the P300 subcomponents. In contrast to the first method, the reference signals are Gaussian spike templates with variable latency and width. The target of this algorithm is to measure the properties of the extracted P300 subcomponents and obtain features that can be used in the classification of schizophrenic patients and healthy subjects.

Finally, the idea of spatial filtering combined with the use of a refer-

ence signal for localisation is introduced for the first time. The designed algorithm localises our desired source from within a mixture of sources where the propagation model of the sources is available. It performs well in the presence of noise and correlated sources.

The research presented in this thesis paves the path in introducing adaptive techniques based on reference signals into ERP estimation. The results have been very promising and provide a big step in establishing a foundation for future research.

To all my friends.....

to my family.....

and to you.....

ACKNOWLEDGEMENTS

I would like to give my great thanks to my supervisor Saeid Sanei for his continuous help and support over the duration of my PhD.

Thanks go to the other faculty members, Prof. Jonathon Chambers and Dr. Sangarapillai Lambotharan, for their valuable suggestions and comments. The provision of real EEG data from Mr. Alex Sumich is greatly appreciated.

I express my thanks and appreciation to my family for their understanding, motivation and patience. Lastly, but in no sense the least, I am thankful to all colleagues and friends who made my stay at the university memorable and valuable.

PUBLICATIONS

Journal Papers

- **L. Spyrou**, S. Sanei, M. Jing and A. Sumich "Separation and localisation of P300 sources and the subcomponents using constrained blind source separation" *EURASIP Journal on Advances in Signal Processing*, Vol. 1, pp. 89-98, January 2006

- M. A. Latif, S. Sanei, J. Chambers and **L. Spyrou** "Partially constrained blind source separation for localization of unknown sources exploiting non-homogeneity of the head tissues" *Journal of VLSI Signal Processing Systems*, Vol. 49, No. 2, pp. 217 - 232, November 2007

- **L. Spyrou**, S. Sanei "Localisation of Event Related Potentials Incorporating Spatial Notch Filters" *IEEE Transactions on Biomedical Engineering*, accepted and ready for publication

Conference Papers

- **L. Spyrou**, S. Sanei " A new beamforming approach for the localisation of event related potentials." *Proc. EUSIPCO*, pp. 2489-2493, 15 September 2007

- **L. Spyrou**, S. Sanei, and C. Cheong-Took " Estimation and Location Tracking of the P300 Subcomponents from Single-Trial EEG" *Proc. IEEE, ICASSP 2007*, Vol. 2, pp. 1149-1152, 15-20 April 2007

- **L. Spyrou**, S. Sanei "A Robust Constrained Method for the Extraction of P300 Subcomponents" *Proc. IEEE, ICASSP 2006*, Vol. 2, pp. II-II, 14-19 May 2006

- **L. Spyrou**, S. Sanei "A Novel Approach For The Single-Trial EEG Estimation Of The P300 Subcomponents" *Proc. 5th European Symposium. on BioMedical Engineering 2006*, 7-9 July 2006

- **L. Spyrou**, S. Sanei, and A. Sumich "Separation and Localization of Auditory And Visual P300 Sources In Schizophrenia Patients Via Constrained BSS" *Proc. IEEE, ICASSP 2005*, Vol. 5, pp. 613-616, 18-23 March 2005

- S. Sanei, **L. Spyrou**, W. Wang and J. A. Chambers " Localization of P300 Sources in Schizophrenia Patients Using Constrained BSS" *Proc. The 5th Conf. in Independent Component Analysis (ICA2004)*, Vol. 3195, pp. 177-184, September 2004

CONTENTS

ANN—Artificial Neural Network

BEM—Boundary Element Method

BSE—Blind Signal Extraction

BSS—Blind Source Separation

CBSE—Constrained BSE

CBSS—Constrained BSS

CSP—Common Spatial Pattern

DL—Distributed Linear

DTF—Directed Transfer Function

ED—Equivalent Dipole

EEG—Electroencephalogram

ERP—Event Related Potential

FOCUSS–Focal Underdetermined System Solver

ICA—Independent Component Analysis

JADE—Joint Approximate Diagonalisation of Eigenmatrices

LCMV—Linear Constrained Minimum Variance

LORETA—Low Resolution Electromagnetic Topography

LS—Least Squares

MAP–Maximum a Posteriori

MEG—Magnetoencephalogram

MN—Minimum Norm

MUSIC—Multiple Signal Classification

MVAR—Multivariate Autoregressive

P300—ERP component that peaks at 300 milliseconds

PSF—Point Spread Function

RAP-MUSIC—Recursively Applied MUSIC

SNF—Spatial Notch Filter

SNR—Signal to Noise Ratio

SOBI—Second Order Blind Identification

STFT—Short Time Fourier Transform

SVD—Singular Value Decomposition

WMN—Weighted Minimum Norm

List of Figures

List of Tables

Chapter 1

INTRODUCTION

It has been established for a long time that the information processing within the brain is carried out in the form of electric signals. The comprehensive study of these signals has become possible through the advances of technology throughout the last century. Most notably, the invention of the electroencephalogram (EEG) has made it possible to study the electrical behaviour of the brain in great detail. EEG is a technique that records the electrical potentials created inside the brain from electrodes[1] placed on the scalp.

The study of the electrical properties of the brain has led to greater understanding of the functions of the mind. Based on EEG information the human brain is roughly divided into regions corresponding to various functions including moving, sensing and reactions[2].

There are many types of electric signals created inside the brain that serve different functions or that are used to explain different mental states of the brain (for more details see next chapter). In this thesis we concentrate on one subgroup of electric signals called event related potentials (ERPs). ERPs occur as the response of the brain to some specifically designed stimuli that aim to elicit the ERP signal. By

[1] small electric conductors

[2] There are other methods as well, such as magnetic resonance (MRI) imaging, but these are not in the scope of this thesis.

studying these signals clinicians can make deductions regarding many behavioural characteristics of the person under test. For example, the relative amplitude of some of the ERP subcomponents[3] can be indicative of the levels of suspiciousness of the person [1].

In this thesis we concentrate on the ERPs of schizophrenic[4] patients. It has been discovered that the ERPs of schizophrenic patients usually exhibit different properties than those of non-schizophrenics. Also, there are differences between subtypes of schizophrenia. In some cases studying the ERPs of patients can lead to greater accuracy and effectiveness of the medical treatment. The main aim of this thesis is to find ways to extract some specific ERP subcomponents or their spatial and temporal properties that are useful to clinicians. The main ERPs of interest here are the P300 subcomponents (for details see next chapter). The P300 subcomponents occur as positive spikes about 300 milliseconds after the application of the stimulus to the person. The stimulus can be visual, auditory or somatosensory. Here, our focus is on visual and auditory stimuli.

Traditonally, ERPs are studied by averaging a number of trials. This method inherently smoothes the waveform and filters out undesired components. A trial is defined as one particular application of a stimulus. Usually, tens of stimuli are presented to the subject. This method has many problems such as:

- Fails to characterise overlapping components. As it is widely regarded, the ERPs consist of a number of components that overlap in time.

[3]ERP is the collection of a number of different components, each possibly having a number of subcomponents.

[4]A psychiatric disease that impairs the perception of reality, can manifest in the form of hallucinations, delusions, disordered speech and disorganised thinking.

- Poor spatial resolution. As a result of the averaging process, the method hardly gives sufficient information regarding the locations of the components.

- Trial-to-trial variability is lost. It is observed that the ERP changes after repeated stimuli. This is known in P300; for example, as habituation occurs the amplitude is gradually reduced.

In the signal processing field a large number of methods have been applied to the problem of EEG and ERP processing and will be shown in Chapter 3.

1.1 Aims and Objectives

In this thesis we aim to develop signal processing methods that reliably extract ERP components and localise their sources within the brain, by using prior knowledge of their properties. As the use of prior knowledge is the main motivation behind this study, in all the methods we develop we use it in various ways. This thesis represents substantial and significant contribution in the following three ways:

- Extraction of the P300 subcomponents by developing a constrained independent component analysis (ICA) method.

- Extraction of the P300 subcomponents and estimation of their important parameters using a novel method based on introducing proper templates for the P300 subcomponents.

- Localisation of the P300 subcomponents. We have developed two methods to deal with the localisation problem. The first method is a simple way of categorising the locations of different signals. In

other words, it is used mainly to find out if two signals originate from the same location in general. The second and our most important contribution, is a method that locates exactly with high efficiency the locations of the P300 subcomponents.

1.2 Thesis Outline

The thesis is organised in 7 chapters that describe our work and the main references and resources. Chapter 2 shows some basic knowledge regarding EEG, ERP and the basic processes that govern the creation of electric signals in the brain. In Chapter 3 we describe the main signal processing methods used regarding EEG and ERP. Our work has been greatly inspired by some of these methods. In Chapter 4 we develop our ICA/BSS[5] method and we apply it to real EEG data. We also perform a simple localisation procedure to categorise schizophrenic patients and healthy subjects[6]. Chapter 5 shows our work on estimating the properties of the P300 subcomponents and we use a modified version of the localisation algorithm of Chapter 4. In Chapter 5 a technique based on the application of templates of the P300 signals is established. Our most important contribution, a new localisation algorithm, is described in Chapter 6. A new spatial filtering approach for the localisation of ERPs is developed in this chapter. Finally, we conclude this thesis in Chapter 7.

[5]Blind Source Separation. Term used interchangeably with ICA.

[6]In this thesis we refer to healthy subjects as people with no schizophrenia. We also use the terms control subject or normal subject.

Chapter 2

OVERVIEW OF THE ELECTROENCEPHALOGRAM AND EVENT-RELATED POTENTIALS

The electrical activity of the brain was first studied by Richard Caton in 1875 [2]. This study was performed on rabbits and monkeys. Similar experiments were carried out on rabbits and dogs in 1890 by Beck [2]. Hans Berger is widely credited for the invention of the human EEG in 1926 although similar experiments were performed by some other researchers.

EEG describes the electric activity recorded at the surface of the head using metal electrodes. These are specifically designed electric conductors that are placed on the scalp with a certain type of gel that reduces the contact impedance of the skin. The signal from each electrode is then fed to a separate differential amplifier[1] in the EEG machine to amplify the signal and reduce noise. The second input of the amplifier is fed from what is called the reference electrode, which

[1] A an amplifier circuit that amplifies the difference between two signals.

is usually an electrode placed on the earlobes or forehead. A typical adult human EEG signal has the amplitude of about $10 - 100\mu V$ when measured from the scalp. EEG is a non-invasive method in contrast with electrocorticogram which records the electrical activity from the surface of the brain.

Before explaining the source of electrical activity picked up by the EEG we first briefly describe the functions mainly responsible for the electrical activity inside the brain. This requires the study of a type of nerve cell, the neuron, which is provided in the next section.

2.1 Neuron

The neuron is the brain cell responsible for the information transfer within the brain. It has distinct features that allows the brain to perform a variety of functions that are used for intracellular and intercellular signalling. The signalling process is an electrochemical process and the information transfer is performed by the neuron while changing its chemical properties [3]. A neuron is shown in Figure 2.1; adapted from [4].

We now describe the main parts of a neuron responsible for information transfer [3]. A neuron consists mainly of the axon, the dendrites, the cell body and the synapse.

Axon. The axon is a thin tube structure that its length can vary from micrometers up to more than a meter, always terminating at a synapse (see below). The main function of an axon is to transfer the signal from the cell body (see below) to the synapse.

Dendrite. The dentrites are structures that are usually thinner than the axon and tend to form tree-like structures called dentritic tree.

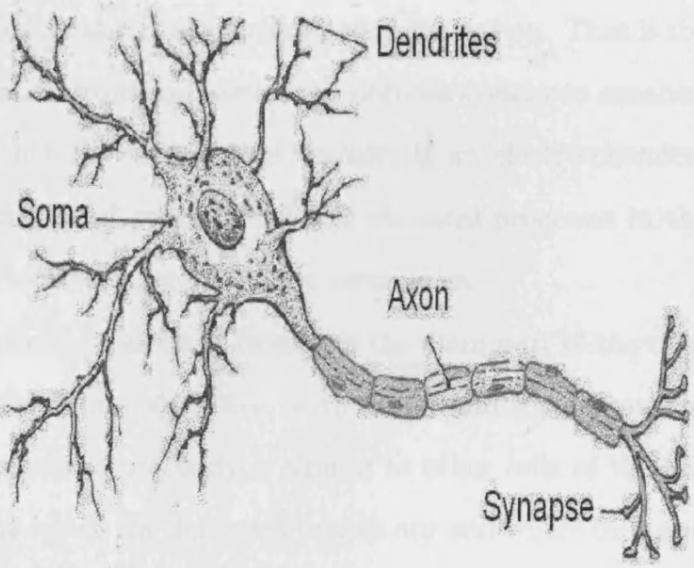

Figure 2.1. Schematic diagram of a neuron's basic parts.

The role of the dendrites is to receive information from the synapses of other neurons although that is not exclusive. Some dendrites share with axons the ability to transmit electric signals.

Synapse. The synapse is a unique and highly sophisticated structure that serves the function of intercellular communication. That is the transfer of information from one part of the nervous system to another. It is now known that the information transfer is an electro-chemical process and it is mediated by electrical and chemical processes in the synapses [3]. It is located where the axon terminates.

Cell body (soma). Cell body (soma) is the main part of the neuron and contains the neuron's nucleus (with DNA[2] and typical nuclear organelles). The neuron's cell body is similar to other cells of the human body and from which the dendrites branch out and where the axon starts.

As mentioned previously the information transfer in the brain is performed in the form of electric signals. These electric signals arise from potential differences in neurons. More specifically, when the cell is not conducting any nerve impulses (signals from other neurons or nerves), a potential difference is formed between the inside of the cell membrane and its outside. Hence, membranes are polarized or, in other words, exhibit a resting membrane potential. This means that there is an unequal distribution of ions (atoms with a positive or negative charge) on the two sides of the nerve cell membrane. This potential generally measures about $40-90mV$ (with the inside of the membrane negative with respect to its outside). So, the resting potential is expressed as -40 to -90 mV. The resting potential is due to unequal

[2]Deoxyribonucleic acid

distributions of potassium (K+), sodium (Na+) and chloride (Cl-).

When a cell membrane conducts nerve impulses the cell membrane is rapidly depolarised (the potential difference takes a positive jump). This jump in turn opens what are called sodium gates and (Na+) from inside the cell flows towards the outside. The system functions in such a way that when the potential difference reaches a certain level the sodium channels close (however, there are some left open still). The process then repeats in reverse and the potential difference goes back to the resting potential. That positive surge is called an action potential (AP) and is the signal that triggers the chemical processes at the synapse.

The action potential is transmitted through the axon to the synapses. A special chemical called neurotransmitter then either excites or inhibits the connecting cell. This depends on the type of the recipient cell (post-synaptic cell).

2.2 EEG generation

Scalp EEG measures summated activity of post-synaptic currents as a result of the neurotransmitters from a large number of neurons. The activity of many types of receptors results in a flow of ions into or out of the dendrite of the receiving cell. This results in a current flowing outwards which is responsible for the generation of EEG voltages. It has to be noted that the EEG is not sensitive enough to pick up the action potentials formed at the cell body and axon.

Surface EEG is the summation of the electrical activity of thousands of neurons with similar orientation. Scalp EEG is more sensitive to radially oriented neurons and it benefits from the parallel arrangement

of dendrites in the cortex.

EEG is highly described by rhythmic activity that has been divided into different frequency bands that generally relate to different functions and brain states such as sleep, awake, and mental disorders. These rhythmic activities, again, represent the synchronised activities of networks of neurons. The characteristics of the EEG frequency bands are briefly described here.

Delta is the frequency range up to 4 Hz. It is usually observed in adults during sleep and in infants. Also, it appears in adults as a result of brain damage.

Theta is the frequency range from 4 Hz to 8 Hz. It is seen normally in young children and in drowsy or emotionally aroused adults. A high amplitude theta activity may represent an abnormal activity in the brain.

Alpha is the frequency range from 8 Hz to 12 Hz. It is normally observed when the person is relaxed. It usually attenuates with mental activity and is usually larger for the non-dominant hemisphere (right hemisphere for right handed people). Mu rhythm is a special kind of alpha activity that is seen over the sensorimotor cortex.

Beta is the frequency range from 12 Hz to about 30 Hz and has usually a smaller amplitude than the other bands.

Gamma is the frequency range approximately 30 to 100 Hz. Because of the filtering properties of the skull and scalp, gamma rhythms can only be recorded by electrocorticography[3].

[3]Similar to EEG but electrodes are placed inside the head on the brain

2.3 Event Related Potentials

Event related potentials (ERP) belong to a special kind of brain electrical activity that arises as a result of some external or internal stimulus. It is an automatic response to either a perceived event or a thought. ERPs are the main subject of this thesis.

ERPs are used mainly by clinicians in order to assess a number of neurological disorders and cognitive processes. They are elicited by specifically designed experiments that provide different kinds of stimuli presented to the person's attention. For example, the classic oddball experiment[4] is usually performed to elicit the P300 ERP component (see below). Most of the neurological disorders or mental abnormalities can be diagnosed by analysing the temporal and spatial properties of the ERPs. Traditionally, due to the fact that ERPs are comparable in amplitude to non-ERP related activity and are time-locked to the stimulus, the most common method is to apply various repetitive stimuli many times and then average the responses. The response of each application is called a frame. This way non-ERP related activity is filtered out.

P300. This is the main ERP component of interest in this thesis. P300 wave is a positive ERP component which occurs with a latency of about 300ms after novel, or task relevant stimuli. It requires an effortful response on the part of the individual under test [5] [6] [7] [8] [9]. When a stimulus strikes a person's attention a number of ERP components are observed that are usually functionally independent, meaning that they represent different brain states. One of those components is the

[4]For example, the person under test is presented with two types of stimuli, one frequent and one rare. The person is asked to press different button on each case. Such a process elicits the P300 from the rare stimulus.

P300 that can be further decomposed into the P3a and P3b subcomponents. P3a reflects an automatic orientation of attention to novel or salient stimuli independent of task relevance [8] [10]. Prefrontal, frontal and anterior temporal brain regions play a major role in generating P3a giving it a frontocentral distribution [7] [8]. In contrast, P3b has a greater centro-parietal distribution due to its reliance on posterior temporal, parietal and posterior cingulate mechanisms [7] [9]. P3a is also characterised by a shorter latency and more rapid habituation[5] than P3b [8] [9].

The P300 wave represents cognitive functions involved in orientation of attention, contextual updating, response modulation and response resolution [5] [7], and consists of multiple overlapping subcomponents, two of which are identified as P3a and P3b [8] [9]. Figure 2.2 illustrates some typical P3a and P3b waveforms from temporal-basal and temporo-superior dipoles, adapted from [11].

Abnormalities in P300 are found in several psychiatric and neurological conditions [6] however, differences may exist in particular in the specific subcomponents [9]. Moreover, changes to certain P300 subcomponents may distinguish between relatives discordant for psychiatric illness, and between subdiagnosis of illness [12] [13]. That is, although reduced amplitude of the auditory P300 is reported in many studies of schizophrenia, some features of the P300 such as its topography vary with subdiagnosis and sex [14] [12]. Finally, certain subcomponents may be modality specific, whilst others may be independent of modality [9]. Thus, auditory and visual P300 appear to be differentially affected by illness and respond differently to treatment, suggesting differ-

[5]The process that while a stimulus is reapplied the amplitude of the resulting ERP decreases.

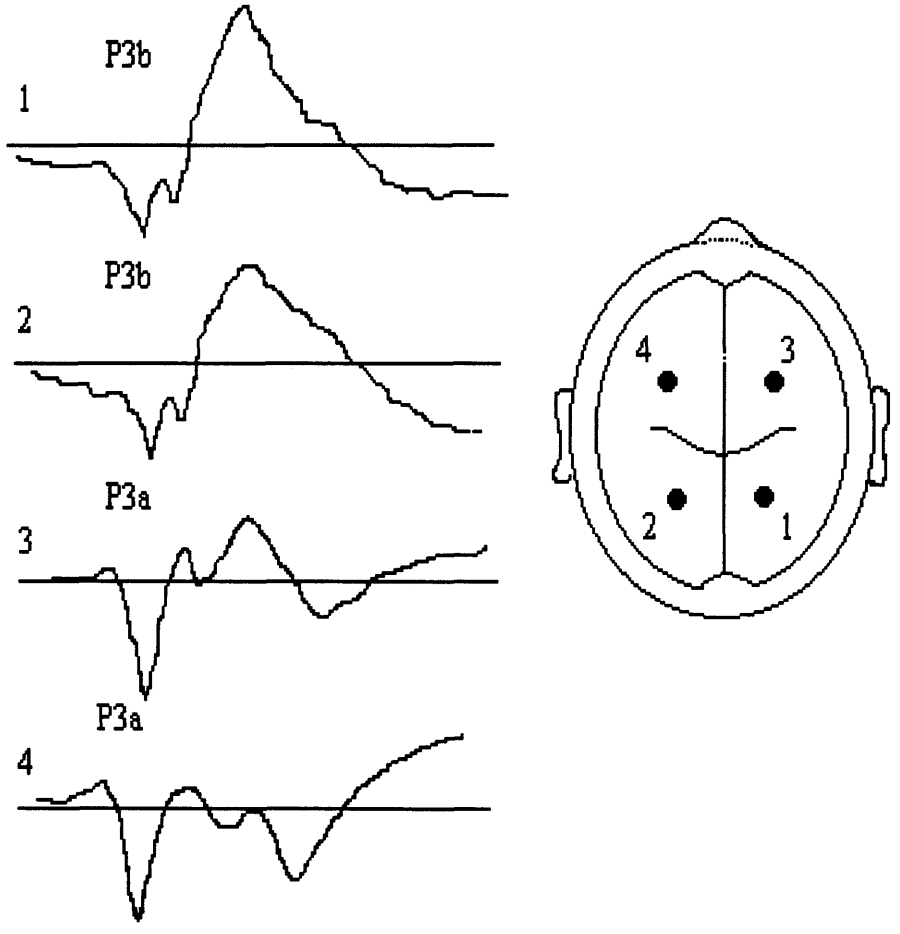

Figure 2.2. Some examples for P3b (1 and 2) and P3a (3 and 4) signals and their corresponding typical locations.

ences in underlying structures and neurotransmitter systems [9]. P300 has significant diagnostic and prognostic potential especially when it is combined with clinical evaluation, that is the procedure of evaluating the symptoms and condition of the patient [6] [9]. However, in order for this to be fully realised, efficient and reliable methods for separating P300 sources and its subcomponents must be established [6].

There are two main ways that any kind of data processing is carried out on ERP signals. The first, which is the traditional and most common method deals with applying some signal processing algorithm on the averaged response from a number of frames. The second way, is to apply some signal processing algorithm on one single frame. The latter approach is called single-trial analysis.

2.4 Conclusions

This chapter showed the grounding knowledge regarding some basic concepts exploited in this thesis. We described briefly about the nature of the EEG, its generation and its acquisition. The neuron structure, a brain cell responsible for the electrical activity in the brain, was briefly shown. The processes that initiate the electrical activities in neurons are also explained. Finally, we talked about ERPs including the P300 subcomponents, which are the ERP components of interest in this thesis.

Chapter 3

SIGNAL PROCESSING METHODS FOR BRAIN SIGNAL SOURCES

In this chapter we describe a number of methods that attempt to estimate the brain activity employing various signal processing methods. We discuss the main developments in EEG signal processing, and particularly the methods that influence the work of Chapter 3 to Chapter 5.

3.1 Dipole fitting methods

The dipole fitting methods are best described by the work done in [15] and [16] although these were not the first papers on the subject [17] [18]. These methods comprise of a current dipole model for each source and by a model of the propagation medium between the sources and the electrodes. The methods described in [15] [16] deal primarily with MEG[1] but the ideas can be easily extended to EEG, see section 3.3. The aim is to estimate the magnetic field or electric potential (in our

[1]Magnetoencephalography, similar to EEG but now it is magnetic fields that are measured.

case) by estimating the strength and locations of the dipole sources. For a dipole at location $\mathbf{L}$, the magnetic field $\mathbf{B}$ at sensor i located at $\mathbf{R}(i)$ is:

$$\mathbf{B}(i) = \mu_0/4\pi \frac{\mathbf{Q} \times (\mathbf{R}(i) - \mathbf{L})}{|\mathbf{R}(i) - \mathbf{L}|} \tag{3.1.1}$$

where $\mathbf{Q}$ (3×1) is the dipole moment including the orientation and strength of the dipole. The model of the data for one dipole source can be written as:

$$\begin{aligned}\mathbf{B} = [\mathbf{B}(1) \dots \mathbf{B}(m)]^T &= \mu_0/4\pi[(\frac{(\mathbf{R}(1) - \mathbf{L}) \times \mathbf{Q}}{|\mathbf{R}(1) - \mathbf{L}|})^T \dots (\frac{(\mathbf{R}(m) - \mathbf{L}) \times \mathbf{Q}}{|\mathbf{R}(m) - \mathbf{L}|})^T]^T \\ &= \mu_0/4\pi[\mathbf{g}^T(1) \dots \mathbf{g}^T(m)]^T \mathbf{Q} = \mu_0/4\pi \, \mathbf{G}\,\mathbf{Q}\end{aligned} \tag{3.1.2}$$

where $\mathbf{g}(i)$ is a 1×3 vector ($\mathbf{G}$ is $m \times 3$) which models the propagation of the source to the magnetometer i, the dipole moment $\mathbf{Q}$ is a 3×1 vector, and $\mathbf{B}$ is an $m \times 1$ vector. For p dipoles the data model can be written as:

$$\mathbf{B} = [\mathbf{G}_1 \dots \mathbf{G}_p][\mathbf{Q}_1^T \dots \mathbf{Q}_1^T]^T \tag{3.1.3}$$

The dipole moments can be factorised to the product of their unit orientation moments and strengths:

$$\mathbf{B} = \mathbf{GMS} \tag{3.1.4}$$

where $\mathbf{M}$ and $\mathbf{S}$ are the dipole moments and strengths respectively. This can be written as a function of the location and orientation such as:

$$\mathbf{B} = \mathbf{H}(\mathbf{L}, \mathbf{M})\mathbf{S} \tag{3.1.5}$$

where $\mathbf{H}$ denotes the function of location and moments. Then, the least squares solution between the model and the actual data has to be obtained:

$$J_{ls} = ||\mathbf{X} - \mathbf{H}(\mathbf{L}, \mathbf{M})\mathbf{S}||_F^2 \tag{3.1.6}$$

The parameters to be estimated are the location $\mathbf{L}$, dipole orientation $\mathbf{M}$ and magnitude $\mathbf{S}$. The major drawback of this approach is that the number of dipoles has to be known beforehand. If too few dipoles are selected then the obtained dipole parameters are influenced by the missing dipoles. If too many dipoles are selected, some of them will not account for valid brain activity. Also, the computation cost to minimise such a cost function is too high. A way to go round it is to consider it as a projection minimisation:

$$J_{ls} = ||\mathbf{X} - \mathbf{H}(\mathbf{L}, \mathbf{M})\mathbf{S}||_F^2 = ||\mathbf{P}_H^{\perp}\mathbf{X}||_F^2 \tag{3.1.7}$$

The matrix $\mathbf{P}_H^{\perp}$ projects the data onto the orthogonal complement[2] of the column space[3] of $\mathbf{H}$. This function has to be iteratively minimised with respect to the desired parameters. By forming the Singular Value Decomposition[4] (SVD) of $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^T$ the cost function can be rewritten as:

$$J_{ls} = ||\mathbf{P}_H^{\perp}\mathbf{U}\Sigma\mathbf{V}^T||_F^2 = ||\mathbf{P}_H^{\perp}\mathbf{U}\Sigma||_F^2 \tag{3.1.8}$$

[2]The orthogonal complement of vector space, Q, is the set of vectors that are orthogonal to all vectors of Q.

[3]The vector space defined by the columns of a matrix.

[4]Method that extracts the basis vectors and their magnitudes from a data matrix.

since orthogonal matrices preserve the F-norm (Frobenius-norm). The matrix $\mathbf{Z} = \mathbf{U}\Sigma$ is now $m \times m$ instead of $\mathbf{X}$ which is $m \times T$, T is the number of samples. So, a great reduction in computation cost has been accomplished. Note that Σ can have only m non-zero singular values because $rank(\mathbf{X}) \leq m$ and generally $T \gg m$. The SVD can also be used to reduce the computations regarding the projection matrix $\mathbf{P}_H^{\perp} = (\mathbf{I} - \mathbf{H}\mathbf{H}^{\dagger})$. The pseudoinverse $\mathbf{H}^{\dagger}$ can be decomposed as $\mathbf{V}_H\Sigma_H^{\dagger}\mathbf{U}_H^T$ or $\mathbf{H} = \mathbf{U}_H\Sigma_H\mathbf{V}_H^T$. The cost function becomes:

$$J_{ls} = ||\mathbf{U}_{m-r}^T\mathbf{Z}||_F^2 = ||\mathbf{Z}||_F^2 - ||\mathbf{U}_r^T\mathbf{Z}||_F^2 \tag{3.1.9}$$

where $\mathbf{U}_r$ is the matrix corresponding to the r non-zero singular values of $\mathbf{H}\epsilon\Re^{m\times r}$ and it spans the column space of $\mathbf{H}$ whereas the rest of the columns correspond to the left nullspace [5] of $\mathbf{H}$. So, for the true dipole locations the F-norm $||\mathbf{U}_{m-r}^T\mathbf{Z}||_F^2$ is zero since the two matrices are orthogonal because the first one belongs to the left nullspace of $\mathbf{H}$ and the second one to the column space and these two spaces are orthogonal. This result was used to develop the MUSIC[6] algorithm for EEG/MEG localisation. It employs a grid search over a number of locations and the quantity $\lambda_{min}[\mathbf{U}_{Gi}^T\Phi_n\Phi_n^T\mathbf{U}_{Gi}]$ is zero for any true dipole location, where $\mathbf{G}_i = \mathbf{U}_{Gi}\Sigma_{Gi}\mathbf{V}_{Gi}^T$ which is the forward matrix for a true dipole source. The matrix Φ_n is called the noise subspace and is estimated by selecting the p less significant eigenvectors[7] of $\mathbf{R}_X = \mathbf{X}\mathbf{X}^T/T$. An extension of the MUSIC algorithm is the R-MUSIC which searches for two dipoles simultaneously to solve the problem of wrong localisation for synchronous sources.

[5] The set of all vectors $\mathbf{a}$ for which $\mathbf{a}^T\mathbf{H}^T = \mathbf{0}$.
[6] Multiple Signal Classification.
[7] Basis vectors of a data matrix.

More recently studies like [19] and [20] have enhanced the dipole fitting model for focal and line sources as well. They consider constant and varying parameters such as the orientation and intensity.

3.2 Inverse Solutions

The inverse solution to the EEG problem consists of estimating the electrical activity at every predefined location in the brain. In such methods the propagation model is available beforehand and consists of a matrix, termed the lead-field matrix, which contains all the propagation information from every location to all electrodes. This model can be described as:

$$\mathbf{X} = \mathbf{HS} \tag{3.2.1}$$

where $\mathbf{X}$ is the $n \times T$ data matrix (EEG or MEG), n is the number of electrodes, T the number of samples, $\mathbf{H}$ is the lead-field matrix which is $n \times 3p$ where p is the number of locations, and $\mathbf{S}$ is a $3p \times T$ matrix which contains the strength and orientation of each dipole (each dipole is represented by a $3 \times T$ vector; from now on we will refer to a component of $\mathbf{S}$ as one $3 \times T$ vector). These methods differ from the dipole fitting methods since we estimate the electrical activity in every location. The solution to that is:

$$\mathbf{S} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{X} \tag{3.2.2}$$

The problem which these methods are inherently faced with is that they are generally underdetermined. Matrix $\mathbf{H}$ is of rank much smaller than p, since the number of electrodes n is much smaller than the number

of locations (usually $n = 16$ or $n = 64$ and $p \simeq 3000$). So, the rank of $\mathbf{H}$ can be at maximum n. This means that there are infinite solutions that satisfy Equation (3.2.2). In other words there exist an infinite number of possible electrical activations that produce the same EEG recording. This can be verified by noting that the columns of $\mathbf{H}$ are not linearly independent and hence the actual activity in a specific location could be estimated as two or more separate activities in different locations. Another point to consider is that the selection of the locations to compute the lead-field matrix $\mathbf{H}$ can have an effect on the result; consider the case where $\mathbf{H}$ does not include a location where there is some electrical activity. Methods that compute the lead-field matrix are shown in the next section. Methods of solving the inverse problem consider the lead-field matrix known and presumably correct and deal with finding a physiologically sound solution to Equation (3.2.2).

As mentioned previously, electrical activity in the brain is synchronised and concentrated in some specific regions. This means that the rows of $\mathbf{S}$ should evolve smoothly and should be nearly zero in most places, meaning that contiguous components of a row of $\mathbf{S}$ should have similar strengths. This desired property of $\mathbf{S}$ has lead to a number of different algorithms that produce a unique physiologically sound solution.

One approach is the minimum norm solution which minimises the norm of $\mathbf{S}$ under the constraint of the forward problem [21]:

$$\min ||\mathbf{S}||_2^2, \quad \text{subject to } \mathbf{X} = \mathbf{HS} \tag{3.2.3}$$

with solution:

$$\mathbf{S} = \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^\dagger\mathbf{X} \tag{3.2.4}$$

The motivation for the minimum norm solution is to create a sparse solution with zero contribution from most sources. This method has the serious drawback of poor localisation performance in 3-D space. An extension to this method is the weighted-minimum norm (WMN) method which compensates for deep sources and hence performs better in 3-D space [21]. In this case the norms of the columns of $\mathbf{H}$ are normalised. For this method the constraint problem is:

$$\min ||\mathbf{WS}||_2^2, \quad \text{subject to } \mathbf{X} = \mathbf{HS} \tag{3.2.5}$$

with solution:

$$\mathbf{S} = \mathbf{W}^{-1}\mathbf{H}^T(\mathbf{H}\mathbf{W}^{-1}\mathbf{H}^T)^\dagger\mathbf{X} \tag{3.2.6}$$

where $\mathbf{W}$ is a diagonal $3p \times 3p$ weighting matrix which compensates for deep sources in the following way:

$$\mathbf{W} = diag[||\mathbf{H}_i|| \ldots ||\mathbf{H}_{3p}||] \tag{3.2.7}$$

Another minimum norm method, known as LORETA (low-resolution brain electromagnetic tomography) [22] employs a spatial smoothing Laplacian operator. This operator produces a spatially smooth solution agreeing with the physiological assumption mentioned earlier. The desired optimisation formulation is then:

$$\min ||\mathbf{BWS}||_2^2, \quad \text{subject to } \mathbf{X} = \mathbf{HS} \tag{3.2.8}$$

where $\mathbf{W}$ is the same as in equation (3.2.7) and $\mathbf{B}$ is the Laplacian operator and its definition can be found in [22].

LORETA produces a blurred image of the true sources but the maxima are correctly located for the noiseless case [22] [23]. FOCUSS is a high resolution iterative WMN method that uses information from previous iterations [23] [24]. The cost function is:

$$\min ||\mathbf{CS}||_2^2, \quad \text{subject to } \mathbf{X} = \mathbf{HS} \tag{3.2.9}$$

In this case $\mathbf{C} = (\mathbf{Q}^{-1})^T\mathbf{Q}^{-1}$ and $\mathbf{Q}_i = \mathbf{W}\mathbf{Q}_{i-1}[diag(\mathbf{S}_{i-1}(1) \ldots \mathbf{S}_{i-1}(3p)]$ and the solution at iteration i is [24]:

$$\mathbf{S}_i = \mathbf{Q}_i\mathbf{Q}_i^T\mathbf{H}^T(\mathbf{H}\mathbf{Q}_i\mathbf{Q}_I^T\mathbf{H}^T)^\dagger\mathbf{X} \tag{3.2.10}$$

The iterations stop when there is no significant difference in the estimation. $\mathbf{W}$ is the same weighting matrix as before. The initialisation to the FOCUSS algorithm plays an important part in a correct estimation. In practise, the algorithm converges to a solution close to the initialisation. It has been suggested that the LORETA algorithm can be used for the initialisation of FOCUSS [25] [26]. Standardised LORETA (sLORETA) is another method that achieves a unique solution to the inverse problem [27]. It uses a different cost function such as:

$$J(\mathbf{S}) = \min\{||\mathbf{X} - \mathbf{HS}||_2^2 + a||\mathbf{S}||_2^2\} \tag{3.2.11}$$

This is the zero-order Tikhonov regularisation which provides a solution to ill-posed inverse problems. The regularisation parameter a can be estimated by the discrepancy principle or by the L-curve method [23]. The solution is then obtained as:

$$\mathbf{S} = \mathbf{H}^T(\mathbf{H}\mathbf{H}^T + a\mathbf{I})^{-1} \tag{3.2.12}$$

sLORETA normalises each source estimate using the resolution matrix. The resolution matrix $\mathbf{R}$ describes a mapping from the actual to the estimated source activity and is given by:

$$\begin{aligned} &\mathbf{R} = \mathbf{H}^T(\mathbf{H}\mathbf{H}^T + a\mathbf{I})^{-1}\mathbf{H} \\ &\text{and} \\ &\mathbf{S}_{est} = \mathbf{R}\mathbf{S}_{act} \end{aligned} \tag{3.2.13}$$

where $\mathbf{S}_{est}$ is the estimated source activity and $\mathbf{S}_{act}$ is the actual source activity. The resolution matrix can provide useful information as for the so-called point-spread-functions (PSF) of the original sources. PSFs show the spatial dispersion of each actual source and depend on the method used. The normalisation is performed as follows:

$$\mathbf{S}_l^T(\mathbf{R}_{ll})^{-1}\mathbf{S}_l \tag{3.2.14}$$

where $\mathbf{S}_l$ corresponds to the l_{th} source and $\mathbf{R}_{ll}$ corresponds to the l_{th} diagonal block of $\mathbf{R}$. In [22] it is shown that the resolution matrix is equal to the source variance. So, by standardising the estimated sources they are normalised by their power.

In [23] and [26] the LORETA and the FOCUSS solutions were combined to improve the results. It has been claimed that the localisation was successful even for closely spaced and correlated sources. It uses sLORETA as an initialisation to a standardised FOCUSS process. Also, at each iteration only the prominent nodes are kept (along with their neighbouring nodes), which means that the columns of matrix $\mathbf{H}$ are

removed for elements of $\mathbf{S}(i)$ less affecting $\mathbf{X}$. This creates a much better solution because at each iteration "wrong" columns do not interfere with the solution.

It has been also attempted to use the temporal properties of brain signals to improve the localisation performance. This was performed in [28] by adding an additional constraint motivated by the assumption that for each location the change in the source amplitude with time is minimal. The constraint to be added is $min||S(t) - S(t-1)||_F^2$ where t denotes time.

The inverse problem has also been tackled under a Bayesian approach [29]. A method to estimate the sources has been proposed, based on Bayesian theory that introduces some information about the spatio-temporal properties of the sources [30]. This information is incorporated by using the prior probabilities.

The model of the EEG generation is again [30]:

$$\mathbf{X}_t = \mathbf{H}\mathbf{S}_t + \mathbf{b}_t \tag{3.2.15}$$

where $\mathbf{X}_t$ is an $m \times 1$ vector containing the EEG measurements at time t, $\mathbf{H}$ is the $m \times n$ matrix containing the information of the head model, the source set pattern and measurement positions on the surface of the skull. $\mathbf{S}_t$ is a $n \times 1$ vector containing the t_{th} time sample of the source magnitudes, and $\mathbf{b}_t$ is the total additive noise. The a priori information regarding the sources imposes some constraints on their location and their temporal properties. The estimation is carried out using a maximum a posteriori (MAP) estimator. That is, the estimator tries to find $\mathbf{S}_t$ that maximises the probability distribution of $\mathbf{S}_t$ given the measurements $\mathbf{X}_t$. The estimator is:

$$\hat{\mathbf{S}}_t = \max[p(\mathbf{S}_t|\mathbf{X}_t)] \tag{3.2.16}$$

and the posterior probability according to Bayes law is:

$$p(\mathbf{S}_t|\mathbf{X}_t) \propto p(\mathbf{X}_t|\mathbf{S}_t)p(\mathbf{S}_t) \tag{3.2.17}$$

where $p(\mathbf{X}_t|\mathbf{S}_t)$ is the likelihood of $\mathbf{S}_t$ given $\mathbf{X}_t$. In other words it is the forward model between the sources and the sensors. $p(\mathbf{S}_t)$ is the prior probability. The posterior can be written in terms of energy functions:

$$p(\mathbf{S}_t|\mathbf{z}_n) = \frac{1}{Z} exp\,[-U(\mathbf{S}_t)] \tag{3.2.18}$$

and $U(\mathbf{S}_t) = U_1(\mathbf{S}_t) + \lambda U_2(\mathbf{S}_t)$ where U_1 and U_2 correspond to the likelihood and the prior respectively (λ is the regularisation parameter between the two functions). The prior is separated into two functions, one for the spatial priors U_s and one for the temporal priors U_t. The spatial prior function can take into account the smoothness of the spatial variation of the sources. A cost function that determines the spatial smoothness is:

$$\Phi(u) = \frac{u^2}{1 + (\frac{u}{K})^2} \tag{3.2.19}$$

where K is the scaling factor, which determines the required smoothness. Then the prior function for the spatial constraints can be written as:

$$U_s(\mathbf{S}_t) = \sum_{k=1}^{n} [\Phi_k^x(\nabla_x \mathbf{S}_t|_k) + \Phi_k^y(\nabla_y \mathbf{S}_t|_k)] \tag{3.2.20}$$

where the index x or y corresponds to the horizontal or vertical gradi-

ents respectively.

The temporal constraints are imposed by assuming that the source magnitudes are evolving slowly compared to the sampling frequency. It is implemented by minimising the projection of $\mathbf{S}_t$ to the space perpendicular to $\mathbf{S}_{t-1}$. Thus, the temporal prior function could be written as:

$$U_t(\mathbf{S}_t) = \beta||\mathbf{P}_{t-1}^{\perp}\mathbf{S}_t||^2 \tag{3.2.21}$$

where $\mathbf{P}_{t-1}^{\perp}$ is the projector onto the perpendicular space to $\mathbf{S}_{t-1}$. Hence, the total energy function to be minimised can be written as:

$$U(\mathbf{S}_t) = ||\mathbf{X}_t - \mathbf{H}\mathbf{S}_t||^2 + \lambda\{\sum_{k=1}^{m}[\Phi_k^x(\nabla_x\mathbf{S}_t|_k) + \Phi_k^y(\nabla_y\mathbf{S}_t|_k)] + \beta||\mathbf{P}_{t-1}^{\perp}\mathbf{S}_t||^2\} \tag{3.2.22}$$

The estimator in this formulation is:

$$\hat{\mathbf{S}}_t = \min[U(\mathbf{S}_t)] \tag{3.2.23}$$

In [31] an interesting work was carried out that combines the so called equivalent dipole (ED) and distributed linear (DL) models. ED refer to the model of section 3.1, of the dipole fitting methods, where the EEG is explained as the summation of the activity of various dipoles. DL refers to the imaging techniques of section 3.2 where all possible sources are estimated at the same time. Temporal and spatial priors are introduced in a general framework and the parameters are estimated with a maximum a-posteriori estimator (MAP).

There are methods that solve the inverse problem but do not use the actual lead-field vectors for the propagation of the sources to the

electrodes. In [32] a regularised least squares method is exploited and the spatial correlation between the channels is used as additional information in the estimation procedure. The estimation is carried out in a single trial basis. A model has to be established for the generation of the sensor signals as:

$$\mathbf{x}_i = \mathbf{s}_i + \mathbf{e}_i \tag{3.2.24}$$

where x_i is the sensor signal after trial i, s_i denotes the ERP and e_i is the noise signal, which can be the background EEG. The evoked potential s_i is modelled as a linear combination of some basis vectors ψ_j:

$$\mathbf{s}_i = \mathbf{H}\boldsymbol{\theta}_i \tag{3.2.25}$$

where H is the matrix having the basis vectors ψ_j in its columns and $\boldsymbol{\theta}_i$ is the vector of the parameters. The difficulty in this model is the selection of the basis vectors. The best basis would be the true physical model which could describe the spatial and temporal properties of the evoked potential. Some simple Gaussian or sigmoid functions are used that can describe the evoked potentials as consisting of positive and negative humps. As mentioned before the method used is based on regularised least squares (similar to sLORETA). The problem to solve is:

$$\boldsymbol{\theta}_i = \arg\min_{\boldsymbol{\theta}_i}\{||\mathbf{x}_i - \mathbf{H}\boldsymbol{\theta}_i||^2 + a^2||\mathbf{L}\boldsymbol{\theta}_i||^2\} \tag{3.2.26}$$

where a is the regularisation constant. The second part of the right hand side of equation (3.2.26) is called the side constraint. Equation

(3.2.26) is a modification of the ordinary least squares to the direction in which the side constraint norm is minimised. A regularisation matrix L is proposed as the second derivative approximation. That should smooth any sharp peaks in the estimated vector. The solution to (3.2.26) is:

$$\boldsymbol{\theta}_i = (\mathbf{H}^T\mathbf{H} + a^2\mathbf{L}^T\mathbf{L})^{-1}\mathbf{H}^T\mathbf{x}_i \tag{3.2.27}$$

The previous method can be combined with the principal component regression approach. If the first p eigenvectors of the data correlation matrix Rx are placed in a matrix $\mathbf{H}_S$, then that matrix will be an orthonormal basis of the space S of the measurements. The evoked potentials $\mathbf{s}_i$ are desired to be close to that space. The projection of $\mathbf{s}_i$ to S is $(\mathbf{H}_S\mathbf{H}_S^T)\mathbf{H}\boldsymbol{\theta}_i$ and their distance is $||(I - \mathbf{H}_S\mathbf{H}_S^T)\mathbf{H}\boldsymbol{\theta}_i||$. $\mathbf{L}$ is then written as $(I - \mathbf{H}_S\mathbf{H}_S^T)\mathbf{H}$ which gives a solution as:

$$\boldsymbol{\theta}_i = (\mathbf{H}^T\mathbf{H} + a^2\mathbf{H}^T(\mathbf{I} - \mathbf{H}_S\mathbf{H}_S^T)\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}_i \tag{3.2.28}$$

The estimator for the source signals is then $\mathbf{s}_i = \mathbf{H}\boldsymbol{\theta}_i$. The information about the spatial correlation of sources is modelled in the eigenvectors of the data correlation matrix.

3.3 Forward-Problem

Up to now the lead-field matrix $\mathbf{H}$ has been considered known. Here, we show how the electric and magnetic fields arise from a current density[8] $\mathbf{J}^i$ [33]. This is called the *impressed current* and assume that it lies in a conductor with conductivity σ. To compute the electric field $\mathbf{E}$ and

[8] i denotes the source

the magnetic induction **B** Maxwell's quasi-static equations are used:

$$\begin{aligned} \mathbf{E} &= -\nabla V \\ \nabla \times \mathbf{B} &= \mu_0 \mathbf{J}^i \\ \mathbf{J} &= \mathbf{J}^i + \sigma \mathbf{E} \\ \nabla \cdot \mathbf{B} &= 0 \end{aligned} \tag{3.3.1}$$

where V is the electric potential and **J** is the total current. For a homogeneous medium the equation for the electric potential is:

$$V(\mathbf{q}) = -\frac{1}{4\pi\sigma} \int \mathbf{J}^i(\mathbf{q}') \frac{\mathbf{q} - \mathbf{q}'}{|\mathbf{q} - \mathbf{q}'|^3} d\mathbf{q}' \tag{3.3.2}$$

for a point source the electric potential reduces to:

$$V(\mathbf{q}) = -\frac{1}{4\pi\sigma} \mathbf{Q} \frac{\mathbf{q} - \mathbf{q}'}{|\mathbf{q} - \mathbf{q}'|^3} \tag{3.3.3}$$

where **Q** is the dipole moment. When the medium is inhomogeneous (it comprises of regions of different conductivities) then the expressions are more complex. The potential then consists of a number of integrals for each region of different conductivity:

$$\sigma(\mathbf{q})V(\mathbf{q}) = \sigma_n V_0(\mathbf{q}) - \Sigma_{j=1}^{a} \frac{\sigma_j' - \sigma_j''}{4\pi} \int_{S_j} V(\mathbf{q}') \mathbf{n}(\mathbf{q}') \cdot \frac{\mathbf{q} - \mathbf{q}'}{|\mathbf{q} - \mathbf{q}'|^3} dS_j \tag{3.3.4}$$

and S_j denotes the corresponding surface, a is the number of surfaces and σ_j' and σ_j'' are the conductivities of the inner and outer sides of S_j.

The forward model can be simplified by assuming a spherically symmetric conductor or horizontally layered conductor. Detailed derivations can be found in [33].

Another way to compute the electric potential or magnetic field at the scalp is using the so-called boundary element method (BEM). This method approximates segments of the conductor with triangular meshes. Each mesh has a constant conductivity which provides a solution to the mathematical modelling of the electric field. BEMs formulations are such that enable the differentiation of the electric field without any problem. Some studies on forward solutions can be found in [34] [35] [36] [33] [37].

3.4 Spatial-Filtering

Spatial Filtering is analogous to frequency filtering, where the signals are considered disjoint in space. It is also similar to beamforming widely used in telecommunications but the angle of arrival cannot be estimated since the low sampling rate provides the presence of brain signals on all electrodes simultaneously. Linear Constrained Minimum Variance (LCMV) is described below [38] [39]. The spatial filter designed by this method produces a signal which is a linear combination of the data. The weights are selected in order to minimise the signal output while forcing it to pass any activity from a specific location. The design of such a filter will be shown here. Let the filter be denoted as $\mathbf{W}_q$ which is an $n \times 3$ vector, where n is the number of electrodes and q denotes the location. The output signal is $\mathbf{y} = \mathbf{W}^T\mathbf{X}$, where $\mathbf{X}$ is $n \times T$ and T is the number of samples. Also, $\mathbf{H}$ is the lead-field matrix $(n \times 3p)$ and $\mathbf{X} = \mathbf{HMS}$, where $\mathbf{S}$ has the usual meaning of the source matrix and $\mathbf{M}$ is $3p \times p$ containing the dipole orientations. An ideal spatial filter has the following properties: $\mathbf{W}_{q_0}\mathbf{H}_{q_0} = \mathbf{I}$ and $\mathbf{W}_{q_0}\mathbf{H}_q = \mathbf{0}$ for every $q \neq q_0$ and $\mathbf{H}_q$ is the lead-field matrix $(n \times 3)$ for the q location and q_0 is the location

that the filter is centered. So, the signals from all other locations other than q_0 will be cut off. The major drawback of such a filter is that the lead-field matrices are nearly dependent (similar to each other), so any linear filter designed to pass a signal from location q_0 is bound to pass some energy of signals from most nearby locations, though they are mostly attenuated. This effect does not exist in the frequency domain where the basis vectors are linearly independent. Minimisation of the variance attempts to overcome this problem. The filter is defined as the solution to the following constrained optimisation problem:

$$\min_{\mathbf{W}_{q_0}} \text{trace}\,\mathbf{C}(\mathbf{y}) \text{ subject to } \mathbf{W}_{q_0}^T \mathbf{H}_{q_0} = \mathbf{I} \tag{3.4.1}$$

where $\mathbf{C}(y)$ is the covariance matrix of the output signal. The solution to the above optimisation problem at location q_0 has the form:

$$\mathbf{W}_{q_0} = [\mathbf{H}_{q_0}^T \mathbf{C}^{-1}(\mathbf{X}) \mathbf{H}_{q_0}]^{-1} \mathbf{H}_{q_0}^T \mathbf{C}^{-1}(\mathbf{X}) \tag{3.4.2}$$

The energy of the signal at location q_0 is:

$$var_{q_0} = \text{trace}[\mathbf{H}_{q_0}^T \mathbf{C}^{-1}(\mathbf{X}) \mathbf{H}_{q_0}]^{-1} \tag{3.4.3}$$

The estimation can be improved if the noise covariance matrix is known. Also, it has to be noted that superposition does not hold in the spatial spectra. The spectrum of a sum of signals is not necessarily the sum of the individual signals' spectra. This is because of the inverse operation in calculating the energy at each location. Also, the number of electrodes, the number of sources, closeness of the electrodes and the sources as well as the correlation between the sources affect the results.

Another approach to perform spatial filtering is the method of com-

mon spatial pattern (CSP) [40] [41]. This method has been used in classification of single-trial EEG and also hand movement-related patterns (distinguishing between left and right hand movement from the EEG). A CSP algorithm distinguishes between classes and finds a filter that maximises the variance for one class while minimising the variance for the others.

The method works by utilising the covariance matrix for each class. In a two-class scenario suppose $\mathbf{X}_1$ corresponds to class 1 and $\mathbf{X}_2$ corresponds to class 2. Then, a matrix $\mathbf{W}$ is calculated so that:

$$\mathbf{W}\Sigma_1\mathbf{W}^T = \mathbf{D} \text{ and } \mathbf{W}\Sigma_2\mathbf{W}^T = \mathbf{I} - \mathbf{D} \tag{3.4.4}$$

where Σ_j is the covariance matrix for class j and $\mathbf{D}$ is a diagonal matrix consisting only of ones and zeros. Hence, for different classes the result is orthogonal. The matrix $\mathbf{W}$ can be calculated in the following way: first whiten $\Sigma_1 + \Sigma_2$, i.e. determine $\mathbf{P}$ such that:

$$\mathbf{P}(\Sigma_1 + \Sigma_2)\mathbf{P}^T = \mathbf{I} \tag{3.4.5}$$

Then define $\mathbf{S}_1 = \mathbf{P}\Sigma_1\mathbf{P}$, $\mathbf{S}_2 = \mathbf{P}\Sigma_2\mathbf{P}$ and compute the *RD* decomposition of $\mathbf{S}_1$ and $\mathbf{S}_2$. After some algebra it follows that $\mathbf{S}_1 = \mathbf{R}\mathbf{D}\mathbf{R}^T$ and $\mathbf{S}_2 = \mathbf{R}(\mathbf{I} - \mathbf{D})\mathbf{R}^T$. The matrix $\mathbf{W}$ is then:

$$\mathbf{W} = \mathbf{R}^T\mathbf{P} \tag{3.4.6}$$

Hence, by obtaining $\mathbf{W}$ we get the required matrix to distinguish between the classes.

3.5 ICA methods

One of the most useful tools in signal processing with numerous applications is Independent Component Analysis (ICA). This method refers to the recovery of the independent original signals from their combinations. The applications are widespread and range from telecommunications to EEG. The general ICA problem can be formulated as follows: There is a number n of sensor signals $x_i(t)$ which are mixtures of m source signals $s_j(t)$, where usually $n \geq m$. The objective is to find an unmixing system which produces m outputs $y_j(t)$ which are as close as possible to the original sources. This procedure is usually performed in an adaptive manner based on an optimisation procedure. The original sources and the mixing system are assumed unknown, but some information is placed upon the statistical properties of the sources. The sources are assumed to be independent and in the ERP case this is not an unrealistic assumption. There are two main BSS mixing models based on ICA. The first regards the mixing as linear and instantaneous and the second as linear and convolutive. There are also nonlinear models where the medium changes with respect to its inputs. The case of interest is the linear and instantaneous mixing since the electrical activity that reaches the electrodes is a function of distance and the electrical properties of the brain. It is regarded instantaneous mainly because the sampling rates of EEG machines are too low compared to the propagation speed of electromagnetic waves. Thus, at every sampling period the electrodes pick up activity of the same temporal activation. In the EEG case the sensor signals correspond to electrodes and the sources correspond to the sources of brain activity. The mixing system can then be written as:

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ . \\ . \\ x_n(t) \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & . & . & h_{1m} \\ h_{21} & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ h_{n1} & . & . & . & h_{nm} \end{pmatrix} \begin{pmatrix} s_1(t) \\ s_2(t) \\ . \\ . \\ s_m(t) \end{pmatrix} \tag{3.5.1}$$

where h_{ij} is the propagation coefficient from the j_{th} source to the i_{th} electrode[9]. The separation system can be modelled similarly as:

$$\begin{pmatrix} y_1(t) \\ y_2(t) \\ . \\ . \\ y_m(t) \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} & . & . & w_{1n} \\ w_{21} & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ w_{m1} & . & . & . & w_{mn} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ . \\ . \\ x_n(t) \end{pmatrix} \tag{3.5.2}$$

It is convenient to describe the mixing and unmixing models in a noiseless environment in matrix form. A general model for the mixing system is defined as:

$$\mathbf{X} = \mathbf{HS} \tag{3.5.3}$$

where $\mathbf{H}$ is the matrix of the propagation coefficients, $\mathbf{S}$ is a matrix whose rows contain the source signals and $\mathbf{X}$ contains the sensor signals. During the unmixing process the sources are estimated using the following model:

[9] noise is not considered for simplicity

$$\mathbf{Y} = \mathbf{WX} \tag{3.5.4}$$

where $\mathbf{W}$ is the unmixing matrix and is proportional to $\mathbf{H}^{-1}$ subject to the scaling and permutation ambiguities of the estimated sources i.e $\mathbf{W} = \mathbf{DPH}^{-1}$ ($\mathbf{D}$ and $\mathbf{P}$ are the scaling and permuation matrices respectively) and $\mathbf{Y}$ is the estimate of the sources. There are two main ambiguities in the ICA solution. Firstly, the scaling of the original sources cannot be determined and secondly the ordering of the sources is unknown. As mentioned before in ICA the sources are assumed independent, which means that the joint pdf can be expressed as the product of the marginal pdfs of each source. This can be expressed mathematically as:

$$q(\mathbf{s}) = q_1(s_1) \times \ldots \times q_m(s_m) = \prod_{i=1}^{m} q_i(s_i) \tag{3.5.5}$$

From Equation 3.5.5 we can see one way to explain the ICA ambiguities. Since, $q(\mathbf{s})$ is the product of the individual pdfs any arbitrary scaling gives the same solution. For example, $q(\mathbf{s}) = -2q_1(s_1) \times -\frac{1}{2}q_2(s_2) \ldots \times q_m(s_m) = q_1(s_1) \times \ldots \times q_m(s_m)$. For the same reason we get the permutation ambiguity, any arbitrary ordering of the pdfs creates the same solution. In most ICA algorithms the pdfs of the sources are modelled a priori. This way one can minimise the Kullback-Leibler divergence [42] between the hypothesized pdfs and the estimated ones or can minimise the information between the sources having known pdfs. In the Infomax algorithm [43] it is attempted to maximise the information flow between the inputs and outputs of a neural network. The inputs correspond to the electrode signals and the outputs to a non-linear transformation

(which should correspond to the cumulative density function of the input) of the estimated sources. The Infomax algorithm is very similar to the maximum likelihood and the minimisation of mutual information algorithms. Other algorithms are the JADE[10] , SOBI[11], which perform joint diagonalisation of the eigenmatrices of fourth order cumulants and covariance matrices respectively, and the multichannel deconvolution class of algorithms. For an overview and analysis of ICA algorithms see [44] [45] [46] [47] [42].

The Infomax algorithm attempts to maximise the information flow between the inputs and the outputs of an artificial neural network (ANN). It is not nessesary to explain the theory behind ANNs; although it has to be noted that an ANN has the same properties as the model of Equation (3.5.4) with the difference that the output is now a non-linear transformation of the estimated sources, i.e. $\mathbf{Z} = f(\mathbf{Y})$. The inputs are the electrode signals. It is shown that if the nonlinear functions are selected appropriately [43], then the information maximisation will correspond to the minimisation of the dependence between the estimated sources. The Infomax cost function is:

$$J_m(\mathbf{W}) = I(\mathbf{z}, \mathbf{x}) = H(\mathbf{z}) - H(\mathbf{z}|\mathbf{x}) \tag{3.5.6}$$

where $\mathbf{z} \epsilon \Re^{n \times 1}$ is the output of the neural network ($\mathbf{z} = f(\mathbf{y})$, $f()$ is the non-linear activation function applied element wise to $\mathbf{y}$ which is the estimated source vector), $\mathbf{x}$ is the input to the neural network, $I(\mathbf{z}, \mathbf{x})$ is the information between the inputs and the outputs of the ANN, $H(\mathbf{z})$ is the entropy of the output and $H(\mathbf{z}|\mathbf{x})$ is the conditional

[10] Joint Approximate Diagonalisation of Eigenmatrices
[11] Second Order Blind Identification

entropy of the output assuming a known input; note, for convenience the time index is dropped. The natural gradient of (3.5.6) is

$$\nabla_{\mathbf{W}} I(\mathbf{z},\mathbf{x})\mathbf{W}^T\mathbf{W} = \nabla_{\mathbf{W}} H(\mathbf{z})\mathbf{W}^T\mathbf{W} \tag{3.5.7}$$

since $H(\mathbf{z}|\mathbf{x})$ is independent of $\mathbf{W}$. Minimisation based on the natural gradient is used to achieve good convergence [48]. The adaptation rule for the unmixing matrix $\mathbf{W}$ becomes:

$$\mathbf{W}_{t+1} = \mathbf{W}_t + \mu(\mathbf{I} + (\mathbf{1} - 2f(\mathbf{y}))\mathbf{y}^T)\mathbf{W}_t \tag{3.5.8}$$

where $f(\mathbf{y})=(\mathbf{1}+exp(-\mathbf{y}))^{-1}$ is called logistic function and it assumes super-Gaussian outputs, and μ is the learning rate. The adaptation for an individual weight can be described by the equation (using the gradient descent method):

$$\Delta w_{ij} = \frac{cof\, w_{ij}}{det\, \mathbf{W}} + x_j(1 - 2y_i) \tag{3.5.9}$$

where cof represents the cofactor and det the determinant. Thus, each individual weight is adapted in a way that the rows and columns differ from each other, as prescribed by the first term in the right hand side of the equation. When two rows or columns become similar, the matrix becomes singular, and then $det\,\mathbf{W}$ will tend to zero forcing the weight element to change dramatically. This change will be affected by $cof\, w_{ij}$ which expresses the relative singularity of the remainder of the matrix, regardless of the row and column this element belongs to, compared to the whole matrix.

An interesting variation of the classic multichannel ICA approach is Blind Signal Extraction (BSE). In BSE, instead of estimating all the

sources simultaneously they are estimated sequentially. Typical cost functions include the kurtosis estimate and maximum likelihood.

ICA has been applied to EEG data extensively by [49] [50]. The Infomax algorithm has been used to separate the independent components from multichannel EEG data. An extension of the Infomax has been developed where the algorithm is modified to account for super-Gaussian and sub-Gaussian sources. More recently, the same research group developed a convolutive Infomax algorithm [51]. The reason behind it was to take into account the spatio-temporal dynamics of the brain sources (i.e they move across the brain). Also, the convolutive Infomax algorithm can offer analysis in different spectral bands, which is justified based on the well established fact that EEG signals have distinct characteristics in different spectral bands. In this method the time-frequency transform of the electrode signals is computed:

$$x_i(T,f) = \sum_{\tau} x_i(T+\tau) b_f(\tau) \tag{3.5.10}$$

where x_i is the i_{th} electrode ($i = 1, ..., M$) and b_f is the basis function which extracts the f spectral band from the time-domain signal. The short-time Fourier transform used has a basis function:

$$b_f(\tau) = h(\tau) exp(-j\, 2\pi f \tau / 2K) \tag{3.5.11}$$

where $h(\tau)$ is a window function. Hence, each electrode signal (1 channel$\times$ T samples) is transformed to data of size (1 channel $\times$ T samples $\times$ f frequency bands). The mixing model is then:

$$\mathbf{x}(T,f) = \mathbf{A}(f)\mathbf{s}(T,f) \tag{3.5.12}$$

where $\mathbf{A}(f)$ is the mixing matrix at frequency band f and $\mathbf{s}(T, f)$ is the source vector at time T and frequency f. Similarly to the instantaneous case the estimates of the sources are obtained by multiplying the sensor signals with a frequency-domain unmixing matrix $\mathbf{W}(f)$:

$$\mathbf{u}(T, f) = \mathbf{W}(f)\mathbf{x}(T, f) \tag{3.5.13}$$

The algorithm is then a complex extension of the Infomax algorithm (using the natural gradient):

$$\nabla\mathbf{W}(f) = (\mathbf{I} - < \mathbf{v}(T, f)\mathbf{u}(T, f)^H >_T)\mathbf{W}(f) \tag{3.5.14}$$

where $< x(.)y(.) >$ is the inner product between $x(.)$ and $y(.)$, also $\mathbf{v}(T, f) = [v_1(T, f) \dots v_N(T, f)]$ and $v_i(T, f) = sign(u_i(T, f))\frac{\acute{g}(|u_i(T,f)|)}{g(|u_i(T,f)|)}$ where $g'(.)$ is the first derivative of $g(.)$. Also:

$$\frac{\acute{g}(x)}{g(x)} = \frac{1 - e^{-x}}{1 + e^{-x}} \tag{3.5.15}$$

which gives the complex generalisation of the logistic function used in the Infomax algorithm. The steps of the algorithm are:

- perform short-time Fourier transform (STFT)
- perform ICA using the complex Infomax algorithm obtaining independent components (ICs) for each band
- obtain the complex scalp maps for each band which are the columns of $\mathbf{A}(f) = \mathbf{PDW}^{-1}(f)$ (where $\mathbf{P}$ and $\mathbf{D}$ are the permutation and scaling matrices respectively).

The complex scalp maps describe the amplitude and phase differences between different ICs. A complex map $\mathbf{a}_i(f)\ \epsilon\ \mathbf{A}(f)$, whose elements

have a significant imaginary part suggests spatio-temporal variations of the source (spatial-propagation of the source).

The algorithm produces independent components (ICs) for each frequency band. Sources, however may exhibit activity in a number of frequency bands. In order to solve the permutation problem, i.e to find out which ICs in different bands correspond to the same source the Euclidean distances between the scalp maps is utilised. Components with small Euclidean distances between their maps are assumed to come from the same source. Also, the distance between ICs in different bands can be used. The correlation between ICs is used to obtain these components corresponding to the same source. In that paper these methods are described in detail and ways to assign the best matching ICs. Measurements of second and fourth-order statistical dependencies were also described to evaluate the performance of the algorithm.

Another ICA algorithm for separation and localisation was developed in [52]. In that algorithm the Infomax algorithm was used to separate the sources from an EEG. Then, the ICs which exhibit a significant degree of spikiness were selected from the IC group. To improve the selection process the subset of ICs which could be localised as neuronal sources was selected. This was done using the RAP-MUSIC[12] algoritm [16]. Finally, the ICs left are clustered automatically according to their location and time courses. Also, in [53] a joint EEG/MEG modality is used to enhance the localisation process. The combination is done by using mutual information as the criterion. It basically is a solution to the inverse problem where the gain matrices are estimated in such a way as to minimise their mutual information.

[12]Recursively Applied-MUSIC

The ICA process may be constrained by some a priori information to aid the extraction of desired sources. An overview of this is given in [54]. The applications can be found in [55] and in [56] [57].

Another recent development within the field of BSS/ICA that has had application in EEG is non-negative matrix factorisation. It is a technique that decomposes a non-negative matrix $\mathbf{X}$ into non-negative matrices $\mathbf{A}$ and $\mathbf{S}$ [58] [59] [60]. In other words:

$$\mathbf{X} = \mathbf{AS} \tag{3.5.16}$$

In order to perform such a factorisation a cost function is needed to quantify its quality. There are two main cost functions [60]. The first is the Euclidean distance:

$$d_e = ||\mathbf{X} - \mathbf{AS}|| \tag{3.5.17}$$

The other cost function is the Kullback-Leibler divergence:

$$d_{kl} = \sum_{ij} \mathbf{X}_{ij} log \frac{\mathbf{X}_{ij}}{(\mathbf{AS})_{ij}} - \mathbf{X}_{ij} + (\mathbf{AS})_{ij} \tag{3.5.18}$$

Both of these algorithms are implemented using multiplicative update rules. The positiveness is ensured by two main points. Firstly, $\mathbf{A}$ and $\mathbf{S}$ are initialised to positive values. Secondly, the update rule only involves multiplications between positive quantities so the result remains positive. For the Euclidean cost function the update rule is:

$$\mathbf{A}_{ij} \leftarrow \mathbf{A}_{ij} \frac{(\mathbf{XS}^T)_{ij}}{(\mathbf{ASS}^T)_{ij}} \tag{3.5.19}$$

and

$$\mathbf{S}_{ij} \leftarrow \mathbf{S}_{ij} \frac{(\mathbf{A}^T\mathbf{X})_{ij}}{(\mathbf{A}^T\mathbf{A}\mathbf{S})_{ij}} \tag{3.5.20}$$

In [58] and application of NMF onto EEG regarding Alzheimer's disease was performed. The motivation was to include constraints that will enhance the factorisation quality, specialising in the EEG signal. Namely, there is a temporal smoothness constraint that forces the sources to have the short-term variance small compared to the long-term variance. Also, there is a spatial constraint that forces the sources to be spatially as decorrelated as possible.

3.5.1 Optimisation

Optimisation is the process of finding a minimum (or maximum) of a function $f(x)$. We are interested in the case where $\mathbf{x}$ is a vector or even a matrix. Although, here the theory is presented for vectors it can be easily generalised for matrices. Only the theory for minimisation problems will be described for convenience, and the generality will not be lost since any maximisation problem can be converted to a minimisation problem by inverting the cost function.

We are interested in the conditions for which the following problem has a solution:

$$\text{minimise } f(\mathbf{x}) \,,\, \mathbf{x} \epsilon E^n \tag{3.5.21}$$

where E^n is the space of vectors with n elements. In order to have $\mathbf{x}^*$ as a minimum of (3.5.21) a set of first-order conditions can be established as:

$$\nabla f(\mathbf{x}^*) = \mathbf{0} \,,\, \mathbf{x}^* \epsilon E^n \tag{3.5.22}$$

Hence, we have n equations and we seek for n unknowns. The second-order conditions are of great importance since they determine if a solution point $\mathbf{x}^*$ achieves a minimum or a maximum. The second-order conditions are:

$$\begin{aligned} &i)\nabla f(\mathbf{x}^*) = \mathbf{0}\,,\ \mathbf{x}^* \epsilon E^n \\ &ii)\mathbf{d}^T \nabla^2 f(\mathbf{x}^*)\mathbf{d} \geq 0, \forall\ \mathbf{d} \epsilon E^n \end{aligned} \tag{3.5.23}$$

Condition (ii) means that in order for $\mathbf{x}^*$ to achieve a minimum the matrix $\nabla^2 f(\mathbf{x}^*)$ should be positive semidefinite. If $\mathbf{d}^T \nabla^2 f(\mathbf{x}^*)\mathbf{d} \leq 0$ then $\mathbf{x}^*$ would achieve a maximum.

3.6 Wavelet-based ERP detection

Apart from conventional time-frequency methods (such as short-time Fourier transform) wavelet decomposition has been also used in the study of ERPs [61]. The wavelet transform expands the ERP into different scales and latencies and allows for a more in-depth look at the brain activity. The variable scale allows us to choose at what detail we want to view the data. In comparison to time-frequency methods, wavelets can have theoretically unlimited resolution. Also, the transform is based on the choice of wavelets, which are waveforms that the decomposition is based on. Generally, wavelet functions are categorised into wavelet families comprising of similar characteristics. The wavelet transform has been used mainly in the following areas:

- Estimation of the overlapping signals in the time domain
- Extraction of single trial ERPs from background EEG. Remember that ERPs exhibit spiky characteristics suitable to most wavelet

families.

- Exploitation of the wavelet coefficients for the extraction of information regarding human behaviour.
- EEG denoising
- Jointly use with ICA.

ERPs are suitable for wavelet analysis since they contain a number of different frequencies (delta, alpha, theta, beta) and each ERP component (P3a, P3b, etc...) usually comprises of a number of subcomponents. Thus, a multiscale (frequency) and translation (latency) method such as the wavelet transform can extract different components in different frequencies and latencies. The method works by scaling and shifting a function $g(t)$ such as $g_{a,b}(t) = \frac{1}{\sqrt{(a)}} g[(t-b)/a]$. This function is called mother wavelet since it defines all the scaled and translated wavelets. The wavelet transform is defined as:

$$W(a,b) = \int_{-x}^{x} s(t)\, g_{a,b}(t)\, dt \tag{3.6.1}$$

In discrete time the scaled and translated wavelets are chosen to be orthogonal. This allows the signal to be partitioned into distinct scales and translations. The dyadic decomposition allows this: $a = 2^j, b = k2^j, [k, j \, \epsilon \, Z]$.

In [61] some applications of wavelets to ERPs are shown. In [62] a spike detection method is introduced in which the wavelet transform is applied and then a detection algorithm extracts the peaks. This method utilises what is known as wavelet denoising in which the wavelet coefficients are compared with a threshold level to restore the signal

from noise [63]. In [64] a mother wavelet is defined based on a real action potential waveform and its hypothesised time distribution. In this way different wavelets can be designed depending on the action potential. In [3] it is shown that a neuron's action potential can have different shapes depending on its electro-chemical state and behaviour. In [65] ICA is applied in the wavelet domain, similar to the method in Section 3.5 where the STFT was used. It has been shown that the proposed ICA algorithm converges faster in the wavelet domain than in the time domain.

3.7 Other separation/localisation methods

There are some other methods that do not fall into the above categories. A number of other approaches to brain source estimation and localisation have been studied. A brief review of them is given here. In [66] the directed transfer function (DTF) was developed. DTF measures the flow of a signal from a channel to another channel and it is based on a multivariate autoregressive (MVAR) model for the data (the data samples are created by a weighted sum of the previous data samples). Another interesting paper that compares different methods in estimating the number of sources of the EEG is [67]. A number of different approaches are compared and some are extended to take into account the spatial properties of the sources. In [68] the performance of inverse solutions for misspecified models has been analysed. In [69] [70] Prony's method is used for event related potential estimation. Prony's method was developed in 1795 by Gaspard Riche de Prony but it was not until recently that the application became practical. In Prony's method a function of a sum of damped sinusoids is fit to a data se-

ries. Regarding ERP, the parameters of the function are estimated to produce different signals corresponding to various ERP components. In [71] the differences in the response of the brain to various stimuli is quantified by information theoretic measures. Then, the resulting differences are used for the classification of the responses of two groups of people with different sensitivities to the stimuli.

3.8 Conclusions

In this chapter we described the major developments in EEG signal processing with emphasis on source localisation and extraction. In Section 3.1 we described the basic theory of the equivalent dipole (ED) model and how it is used to estimate the dipole parameters. This family of methods suffers from the main problem of having to know the number of sources beforehand. An extension of those methods, the MUSIC algorithm, estimates the electrical activity at any point in the brain and does not require any knowledge about the number of sources. However, as with all (ED) methods, it is required that the number of electrodes is greater or equal to the number of dipole sources. The main advantage of the ED methods is the ease of use and implementation and they produce good results for relatively simple cases, i.e. where there is not too high levels of noise or there are not too many sources. In Section 3.2 we showed the theory behind various Distributed Linear (DL) propagation model algorithms, that try to solve the inverse problem. DL methods have the advantage over the ED methods that they estimate the activity at every predefined location in the brain. However, they suffer from some indeterminacy in the solution which needs additional information. Inverse methods have been widely

implemented and there has been research that combines different DL methods which are successful. Section 3.3 dealt with the forward matrix, also known as lead-field matrix, and we showed its most common form. The form presented is the most commonly used in the literature and the one used by Brainstorm [72], a software package that we used to perform our simulations. In Section 3.4 the developments in spatial filtering for EEGs were demonstrated. Spatial filtering is analogous to the inverse methods in the sense that it tries to estimate the electrical activity at every point in the brain. It produces good results although it is sensitive to noise. Similarly to the ED methods, it requires more electrodes than sources. Section 3.5 covered the new ICA methods for EEG which are used primarily for extracting various sources from the EEG. ICA can have many applications including the estimation of the mixing matrix; the way various sources are mixed into the electrode signals. It is easy to implement and it potentially allows analysis that was not possible before. Wavelet methods were shown in Section 3.6, which have good single-trial performance because they behave well at high noise levels. Finally, in Section 3.7 other methods which do not exactly fit to the above categories were described. In general, there has not been any method specifically designed for the estimation of the ERP components, let alone for the P300 subcomponents. In this thesis we aimed to develop methods that use prior knowledge of the shape of those signals, and incorporate it in various signal processing methods.

Chapter 4

CONSTRAINED BSS

The main goal in ERP signal processing is to extract as an accurate description of the desired ERP components as possible. In this section the use of constrained ICA in separating the P3a and P3b subcomponents from the P300 composite wave is described. There are two novel methods introduced. The first one makes use of BSS where we extract a number of sources at the same time, while the second method involves BSE where we extract sources sequentially. Both methods employ a constrained optimisation process in which the primary cost function is based on ICA and the constraint function forces the ICA output towards a certain type of signal. Constrained optimisation is used in situations where we desire our optimisation to be limited or constrained by some other function. This chapter is supported by the following publications [56] [73] [74] [75].

4.1 CBSS

In this section we describe the constrained BSS algorithm we have developed. As mentioned before, we aim to apply prior information about our source signals to enhance the performance of the BSS algorithm. The prior information we have is the latency and the approximate shape of these components, which can be observed from many studies [49] [50],

and also from the data. The motivation is to use a constraint function together with Infomax to utilise the prior knowledge we have about the P3a and P3b subcomponents. We use Infomax because of the extensive use it has undergone in EEG and ERP signal processing (see section 3.5). There are two main constraint functions that can be used which incorporate the prior information of the P3a and P3b. Both involve the idea of a desired, or reference, signal. A reference signal is used to model the ERP source we desire to extract. The first possible constraint function is the correlation (zero-lag) between our reference signal and the output of BSS. The second one, and the one used in this study, is the Euclidean distance between the reference signal and the output of BSS. Both methods are valid and need the corresponding signals to be normalised to the same variance; here, we use the Euclidean distance as our distance measuring function. The next issue is how to obtain the reference signal.

Experiments can be specifically designed in order to elicit specific ERP components. For example, a task can be designed such that when a random stimuli of novel value (something that captures the attention) is applied, P3a is the main component and when a task-relevant stimuli (e.g pressing a button when a specific letter is shown) is applied P3b is the main component. These stimuli are applied a number of times, and the response of each is called a frame. By averaging the frames corresponding to each of the specific stimuli the ERP will have either the P3a or P3b dominant. In the case where P3b is dominant P3a is always present so we can observe both signals. Then the time period corresponding to P3a or P3b is chosen as a reference signal (windowed or zeroed). Another way to obtain a reference signal, used in the next

chapter, is to use some predefined spike functions (such as Gaussian, sinc or wavelet spikes) with latency and shape width according to our prior information. For example, by observing a dataset we can estimate approximately the duration of the P3a or P3b; then this information can be used to set the latency and width of a Gaussian spike. In the next section we show exactly the modelling procedure of the ERP components.

Provided that the reference signal is acquired we can proceed on describing the constraint function. While ICA is producing independent components we are interested in one (or more) of them to be as similar as possible to our reference signal(s). Here, we use the Euclidean distance as the constraint function and we illustrate the effect on one of the ICA outputs. In other words we desire to minimise the following:

$$||\mathbf{r} - \mathbf{y}||_2^2 \tag{4.1.1}$$

where $\mathbf{y} = \mathbf{wX}$ is one independent component of the current iteration of the ICA algorithm and $\mathbf{r}$ is the reference signal. So, by applying this constraint function, CBSS will direct that output towards the reference signal. ICA does not necessarily produce unique outputs, however this constraint function guarantees that the desired source is one of the outputs. The constrained problem in this study [73] [56] is formulated as:

$$\max \quad J_m(\mathbf{W}) \quad \text{subject to} \quad J_C(\mathbf{W}) = 0 \tag{4.1.2}$$

J_m and J_C are the Infomax and the constrained cost functions respec-

tively. The cost function of the CBSS algorithm is:

$$J(\mathbf{W}, \mathbf{\Lambda}) = J_m(\mathbf{W}) - \mathbf{\Lambda} J_C(\mathbf{W}) \tag{4.1.3}$$

where $\mathbf{\Lambda}$ is the matrix of Lagrange multipliers. The constraint function specialised for each row of $\mathbf{W}$ is defined as:

$$J_C(\mathbf{w}_i) = ||\mathbf{r} - \mathbf{w}_i\mathbf{X}||_2^2 \quad for\ i = 1, \ldots, m \tag{4.1.4}$$

where $\mathbf{r}$ is the reference signal, $\mathbf{y}_i = \mathbf{w}_i\mathbf{X}$) is the i_{th} output and $\mathbf{w}_i$ is the i_{th} row of $\mathbf{W}$. The unknown parameters in the problem are now two: the matrix $\mathbf{W}$ and the matrix $\mathbf{\Lambda}$. The matrix $\mathbf{W}$ is found adaptively via the following relation [73] [76]:

$$\begin{aligned} \mathbf{W}_{t+1} &= \mathbf{W}_t + \mu\big(\nabla_{\mathbf{W}_t} J(\mathbf{W}_t, \mathbf{\Lambda})\big)\mathbf{W}_t{}^T\mathbf{W}_t \\ &= \mathbf{W}_t + \mu\big(\mathbf{I} + (1 - (1 + exp(\mathbf{W}_t\mathbf{x}))^{-1})(\mathbf{W}_t\mathbf{x})^T \\ &- 2\mathbf{\Lambda}\big(\mathbf{x}(\mathbf{W}_t\mathbf{x} - \mathbf{P})^T\big)\mathbf{W}_t{}^T\big)\mathbf{W}_t \end{aligned} \tag{4.1.5}$$

and

$$\mathbf{\Lambda}_{t+1} = \mathbf{\Lambda}_t + \rho\ \mathrm{diag}\big((\mathbf{W}\mathbf{x} - \mathbf{P})(\mathbf{W}\mathbf{x} - \mathbf{P})^T\big) \tag{4.1.6}$$

where μ is the learning rate of the adaptation of the unmixing matrix, ρ is the learning rate for the Lagrange multiplier matrix and $\mathbf{P}$ is a matrix whose rows contain the corresponding reference P300 signal. If a block algorithm is required then the data vector $\mathbf{x}$ becomes a matrix and it should be scaled accordingly.

The basic form of the constrained algorithm can be modified to

mitigate some inherent problems with this approach. Firstly, in the present form of the algorithm attempts to produce n outputs that are as close as possible to the P300 reference signal. Although this effect is alleviated partly by the Infomax algorithm which tries to produce different outputs the constraint part of the algorithm will influence more of those outputs that are further away (in Euclidean distance terms) from the reference signal. Hence, it would be a good idea to try to enforce the constraint in one or a small number of the outputs. This comes from the fact that usually the P300 signal consists of a number of subcomponents in different regions of the brain. Secondly, the scaling ambiguity of every ICA algorithm can be a problem since one output could have exactly the same shape as the reference signal but it could be a scaled version of it. The algorithm would change that output (since it violates the constraint) which could damage its shape. So, a scaling procedure is used in which the reference signal matches the amplitude of the estimated sources. Finally, the problem of finding good initial conditions for $\mathbf{W}$, $\mathbf{\Lambda}$, μ and ρ can be overcome partly by using a variable which determines the contribution of the two separate cost functions (i.e main and constraint) to the adaptation of $\mathbf{W}$. This way, the algorithm can be made to work (by avoiding the rapid divergence of the Frobenius norm of $\mathbf{W}$) in a variety of situations. This way, the stability of the algorithm is ensured because the learning is kept bounded especially when $J_c(w_i) >> 0$. It also functions as a safety point to make sure that the algorithm converges to a solution, which produces outputs close to the reference signals. The convergence of the algorithm is stable to the optimum point since both parts of the CBSS function have a negative definite Hessian matrix (easy to prove by checking the sign

and the non-singularity of the Hessian). The constrained cost function can take any form that would be suitable for a specific application. A cost function which maximises the inner product between the estimated sources and the reference signals was used but its performance was not as satisfactory as the Euclidean distance function. Following the theory of constrained optimisation, in those cases where the separation needs to be improved over the traditional ICA methods, a number of new BSS algorithms can be developed. Other suggested cost functions for the present purpose can be maximising the spikiness of the output sources around the time of interest (300ms), estimating the pdf of the P300 sources and forcing the pdfs of the output sources to have a similar form[1] or even applying a spatial constraint using prior knowledge of the possible P300 positions.

A variation of this algorithm which was used to separate the P3a and P3b subcomponents was implemented by using the method of least squares. If the reference signals for P3a and P3b are known then we can model the EEG system as:

$$\mathbf{r} = \mathbf{w}_{opt}\mathbf{X} \tag{4.1.7}$$

where $\mathbf{r}$ is the reference signal, $\mathbf{X}$ is the data matrix and $\mathbf{w}_{opt}$ (row vector $1 \times n$) is the vector that should produce $\mathbf{r}$. Then, the constraint cost function will be:

$$J_C(\mathbf{w}_i) = ||\mathbf{w}_i - \mathbf{w}_{opt}||_2^2 \tag{4.1.8}$$

where $\mathbf{w}_i$ is the i_{th} row vector of the unmixing matrix. This vector

[1]this can be facilitated as part of the original Infomax algorithm where the activation function should ideally be derived from the pdfs of the sources

corresponds to the i_{th} output $\mathbf{y}_i$ expected to be the separated P3a or P3b. The selection of the appropriate $\mathbf{y}_i$ to enforce the constraint, is achieved in terms of which one is closer in terms of the Euclidean distance to the reference signal. $\mathbf{w}_{opt}$ is found using the common least squares (LS) solution:

$$\mathbf{w}_{opt}^T = (\mathbf{X}\mathbf{X}^T)^{-1}\mathbf{X}\mathbf{r}^T \tag{4.1.9}$$

Also, the gradient of Equation (4.1.8) is:

$$\nabla_{\mathbf{w}_i} J_C(\mathbf{w}_i) = 2(\mathbf{w}_i - \mathbf{w}_{opt}) \tag{4.1.10}$$

Then, this gradient is incorporated within the main Infomax update equation in a similar manner to Equation (4.1.5). This constraint is different from those used in [77] [78].

4.1.1 Acquiring the reference signal

Here we illustrate the method that we use to obtain the P3a and P3b reference signals. As mentioned before, in an EEG recording, various stimuli are applied a number of times and each one elicits specific ERP components. Our data contain 40 frames that elicit the P3a and P3b components. Traditionally, ERP estimation would include averaging these frames and using the resulting waveform as the ERP response. Our data are recorded using 18 electrodes (inluding 3 reference electrodes) so by averaging those 40 frames we get 18 waveforms, one for each electrode. In Figure 4.1 we show 3 of those waveforms for a schizophrenic patient. The reference signal has to be a single waveform so the next step is to average all those waveforms. Hence, in effect we

get a space-time average of the ERP components. Note, although the P3a and P3b are dominant there are other ERP components present as well. In Figure 4.2 we show the space-time average for the same patient.

After we obtain the space-time average we need to identify the latencies in which the P3a and P3b occur. This has to be done in order to obtain different reference signals for each one of them. By observing the waveform we can deduce that the P3a (first peak) is located at about 275 milliseconds while the P3b (second peak) is located at about 330 milliseconds. Then we create two windowing functions around those two latencies and we apply them to the space-time average to get the two reference signals. However, it is possible that for different subjects both the P3a and P3b occur with different latencies. The distinctive feature is then that P3a occurs before the P3b. P3a is hence selected by the space-time averaging of all the electrodes and selecting the first peak that occurs near the time of interest (250ms-350ms) and P3b by selecting the second peak.

4.1.2 Results using simulated data

In this section we show the performance of the CBSS algorithm for synthetic data. We mix 3 sources on 3 electrodes and investigate the produced signals for various cases. We use realistic head model obtained through the Brainstorm software [72].

$$\mathbf{X} = \mathbf{HMS} + \mathbf{N} = \sum_{j=1}^{3} \mathbf{H}_j \mathbf{m}_j \mathbf{s}_j + \mathbf{N} \tag{4.1.11}$$

In Figure 4.3 we show the three original sources. An example mixture is shown in Figure 4.4.

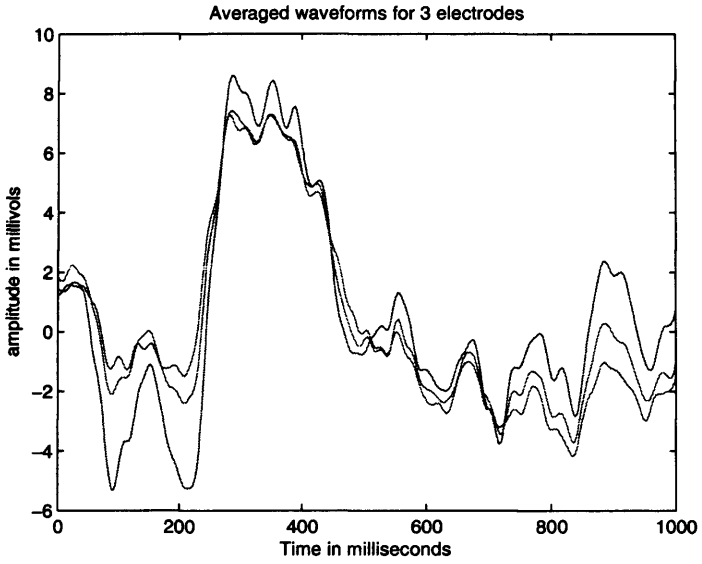

Figure 4.1. Example of a 40 frame average for 3 random electrodes. The common features are evident. In total there are 15 electrodes so the next step is to average those 15 signals.

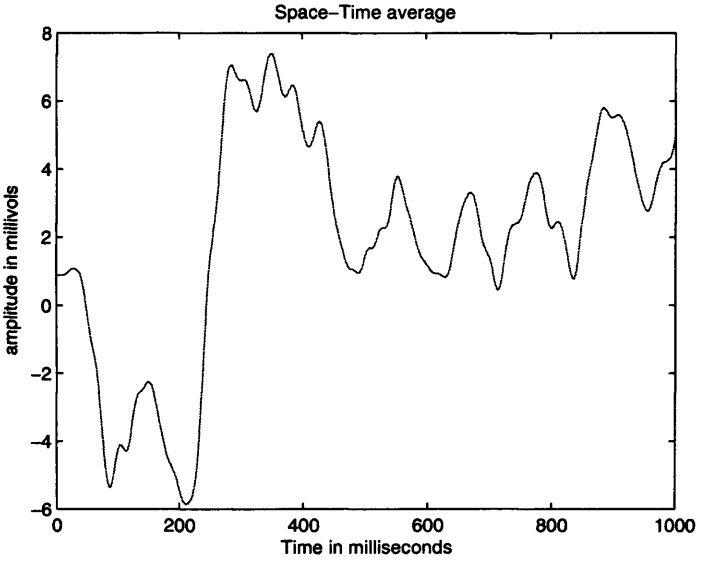

Figure 4.2. Space-time average of the 15 time-averaged(40 frames) electrode signals. This waveform is the basis from where the specific reference signals for an ERP component are obtained.

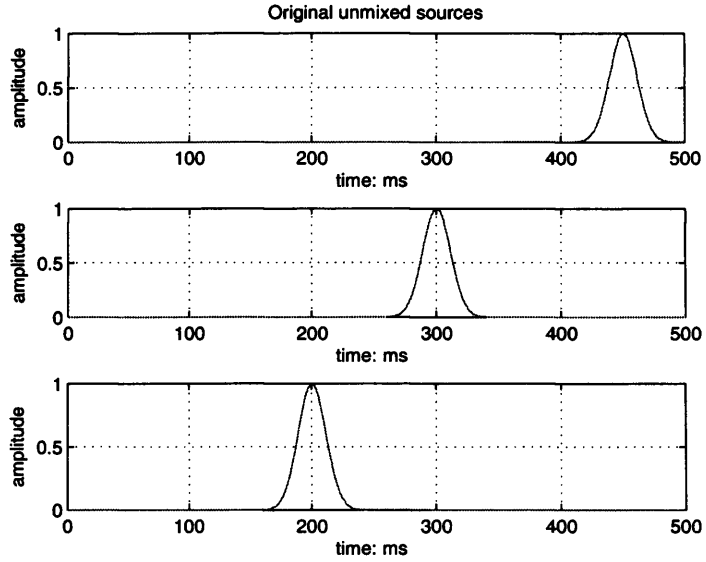

Figure 4.3. Three Gaussian-spike sources peaking at 200ms, 300ms and 450ms.

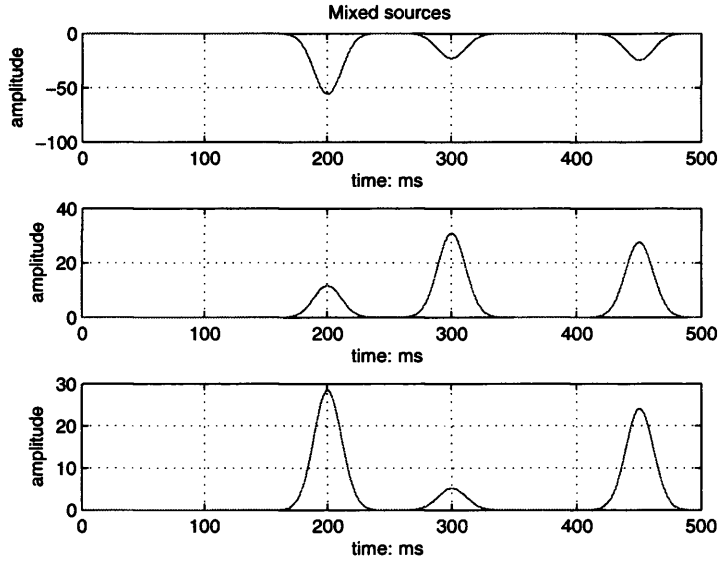

Figure 4.4. Three mixtures from the sources of Figure 4.3.

The first experiment is to constrain the first row of $\mathbf{W}$. Applying the CBSS algorithm has a few significant advantages over unconstrained Infomax. Firstly, it produces a better, in terms of error, description of the original source. Infomax sometimes fails to extract an acceptable source signal. In Figure 4.5 we show an example of this, where

none of the 3 sources are accurately extracted. By applying the CBSS algorithm, using the same initial conditions, we extract the sources of Figure 4.6. Note that we only applied the constraint on the first output source for the signal peaking at 450 ms. Hence, the second advantage of CBSS is that it guarantees our desired signal to be one of the outputs. Also, from Figure 4.6 we observe that the other two sources are better extracted as well, without applying the constraint on them.

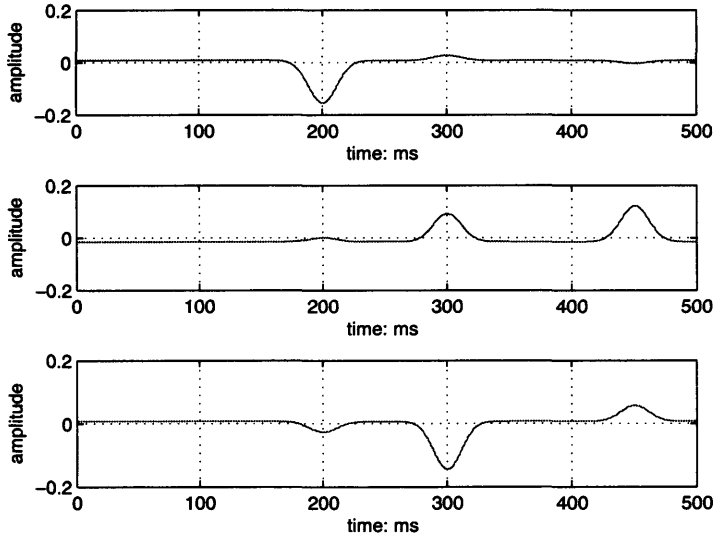

Figure 4.5. Obtained output sources from Infomax. The second and third outputs are still mixtures of the three sources. Only the first output resembles source 3, although it is inverted.

In Figure 4.7 we show a simple case where Infomax is successful but as depicted in Figure 4.8 CBSS extracts the sources with smaller error. We have to make clear at this point, that for all these cases we use the same sources and locations. The only difference is the dipole orientation $\mathbf{m}_j$. It has to be pointed that for the same initial conditions, Infomax (and hence CBSS) produce always the same output, the optimisation process is deterministic. As a result, we try to observe the behaviour of both algorithms for different initial conditions (dipole orientation being

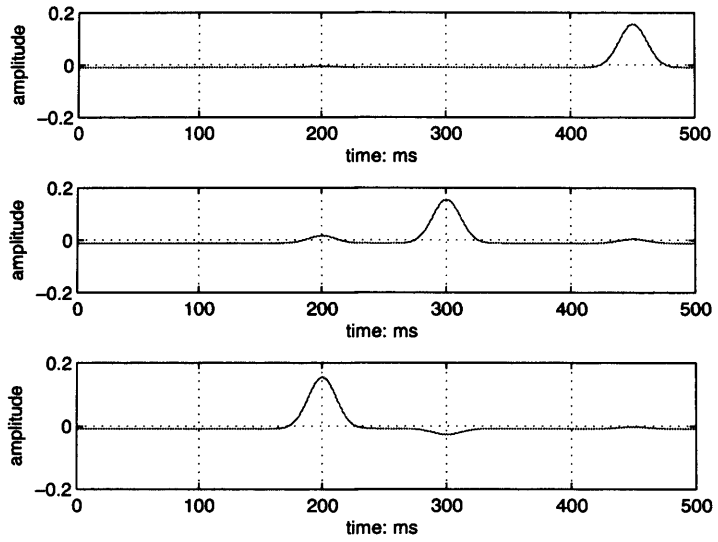

Figure 4.6. By applying CBSS we obtain all 3 sources quite accurately. Note that we applied the constraint only on the first output. It is significant to observe then, that using the constraint only one output, the other outputs converged to the true sources as well.

one of them), simulating real EEG analysis.

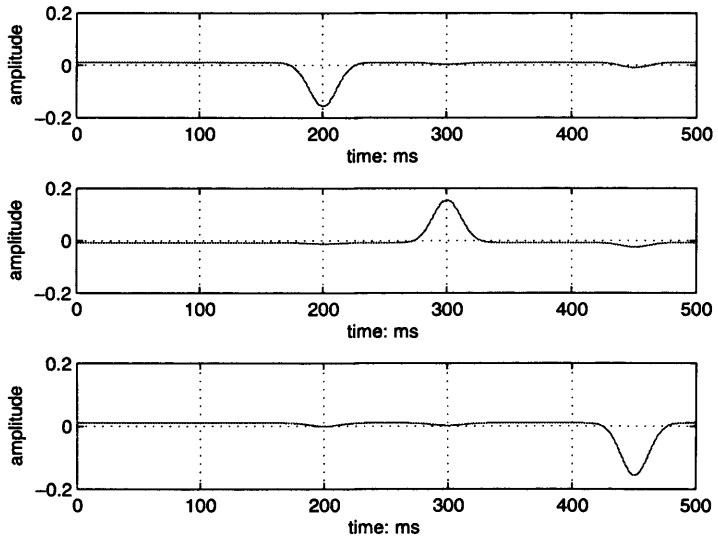

Figure 4.7. Here, Infomax succesfully obtains accurate descriptions off all 3 sources.

For the next case we apply some noise on the mixtures. The SNR

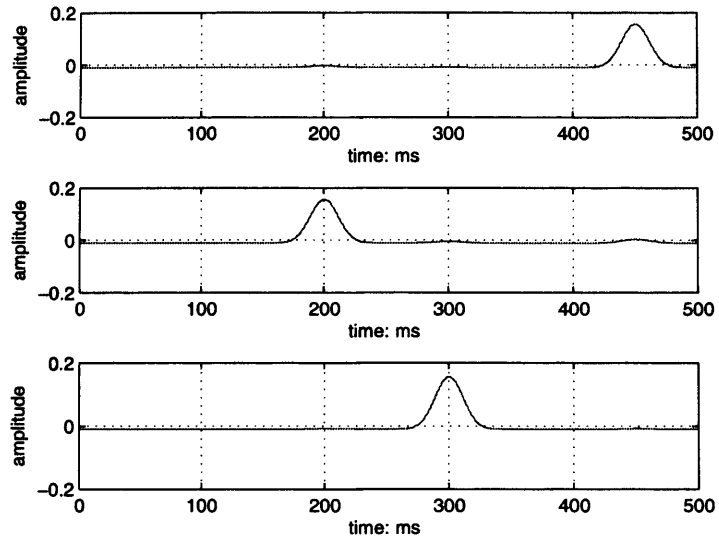

Figure 4.8. With CBSS we get more accurate descriptions of the sources.

is 8dB. In Figure 4.9 we see that Infomax fails to produce the correct output sources. After applying CBSS we get the sources of Figure 4.10.

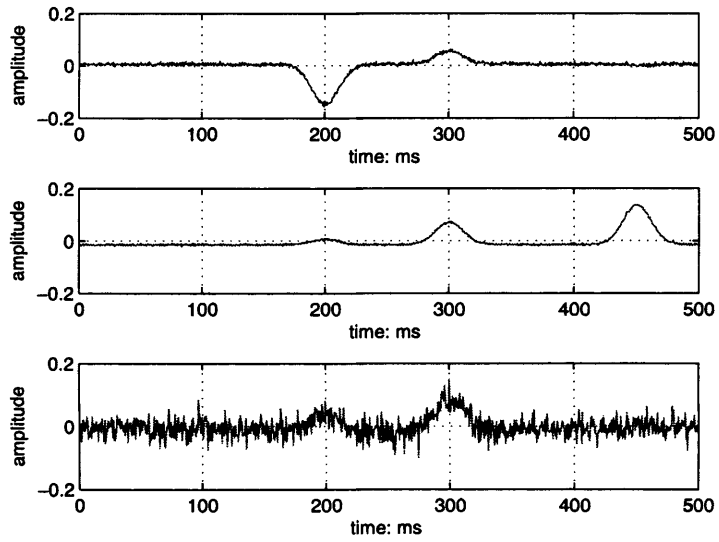

Figure 4.9. We see here the extracted output sources from Infomax. None of the sources are accurately extracted.

We carried out a statistical test to evaluate how the algorithm performs for various dipole orientations. We run the algorithm for 1000

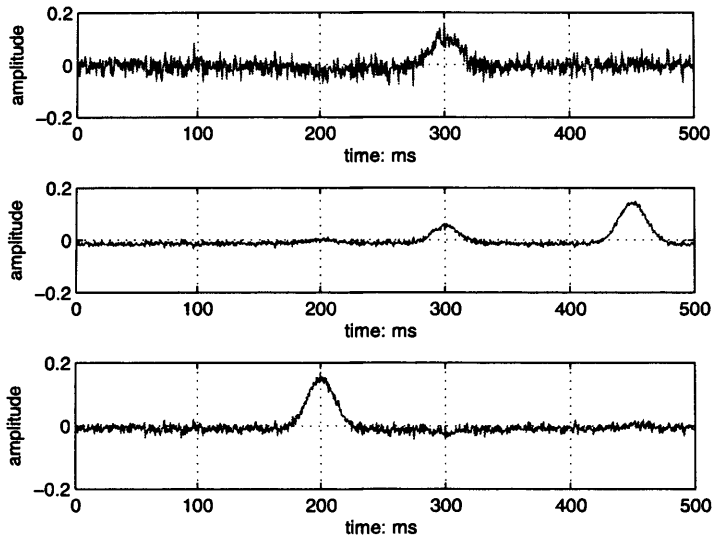

Figure 4.10. Applying CBSS on the same mixtures, still constraining the first output, we get a much better description of all 3 sources.

times for 1000 different dipole orientation, $\mathbf{m}_j$ in Equation (4.1.11). In Table 4.1 we see the results.

SNR	CBSS error	Infomax error
10dB	0.2824	0.8758
3dB	0.2853	1.0927
0dB	0.4032	1.0860

Table 4.1. Simulation results comparing CBSS and Infomax. The error is the Euclidean distance between the original source and the estimated one. It is evident that CBSS extracts the desired source with significantly less error.

In the next case, Table 4.2, we try to enforce the constraint on two rows, so we have two desired outputs. Again, we run the algorithm for 1000 different dipole orientations for each source and calculate the mean error between the original and estimated sources.

SNR	CBSS error1	Infomax error1	CBSS error2	Infomax error2
10dB	0.1673	0.8705	0.0607	0.7329
3dB	0.3421	1.0333	0.0519	0.8847
0dB	0.3970	1.0861	0.0667	0.9537

Table 4.2. Simulation results comparing CBSS and Infomax. The error is the Euclidean distance between the original source and the estimated one. It is evident that CBSS extracts the desired source with significantly less error. In this table the constraint was enforced on two outputs.

4.1.3 Results using Real EEG

In this section we show the results of applying CBSS on real EEG data. The EEG data were recorded using a Nihon Kohden model EEG-F/G amplifier and Neuroscan Acquire 4.0 software. EEG activity was recorded following the international 10-20 system from 15 electrodes. The reference electrodes were linked to the earlobes. The impedance for all the electrodes was below 5kΩ, sampling frequency Fs=2kHz and the data were subsequently bandpass filtered (0.1-70Hz). This frequency range was chosen to be compatible with [79].

Subjects were required to sit alert and still with their eyes closed to avoid any interference. Also, to avoid any muscle artefact the neck was firmly supported by the back of the chair. The feet were rested on a footstep. The stimuli were presented through ear plugs inserted in the ear. Forty rare tones (1kHz) were randomly distributed amongst 160 frequent tones (2kHz). Their intensity was 65dB with 10ms and 50ms duration for rare and frequent tones respectively. The subject was asked to press a button as soon as they heard a low tone (1kHz). The ability to distinguish between low and high tones was confirmed before the start of the experiment. The task is designed to assess basic memory processes. ERP components measured in this task included N100, P200, N200 and P3a and P3b.

Firstly, from the raw-EEG we temporally average event-related data (40 events), each event producing an EEG of size $n \times T$, where n is the number of electrode signals and T is the number of samples of the event. That averaged ERP is also of dimensions $n \times T$. Here, $n = 15$ and $T = 1000$. The advantage of averaging event-related data is not only to enhance the signal, but also to remove non-event-related noise.

Secondly, the reference sub-component signal is selected according to the method described in section (4.1.1). Thirdly, CBSS is applied to the ERP ($n \times T$) in order to separate the P300 and its sub-components. Filtering (at the delta range) is applied to the separated sources, based on the knowledge that the main power of the P300 component is in the delta range [80]. Figure 4.11 shows three electrodes of the averaged ERP and Figure 4.12 shows the estimated P3a and P3b sources for a schizophrenic patient. Figure 4.13 shows the estimated P3a and P3b using the unconstrained version of Infomax. We can see that CBSS produces outputs with more distinct peaks than unconstrained Infomax.

4.1.4 Localisation

Localisation of electrical sources inside the brain has been investigated by a number of people [15] [16] [81] [34] [82]. Unlike the work done in [15] which assumes that the electrical sources are magnetic dipoles, in the paper [73] we assumed that they are sources of isotropic propagation. Hence, the head simply mixes and attenuates the signals. Therefore based on Figure 4.14 we have

$$||\mathbf{f}_k - \mathbf{a}_j||_2 = d_j \quad for\, j = 1, 2, 3 \tag{4.1.12}$$

where $\mathbf{f}_k$ is the position of the source k, $\mathbf{a}_j$ are the positions of the electrodes and d_j are the distances between the source and the j_{th} electrode. The distances d_j are nonlinearly proportional to the inverse of the correlation between the estimated source and the electrode signals. This is because a source is attenuated non-linearly with the distance. Hence, the correlation of the electrodes with a source is non-linearly

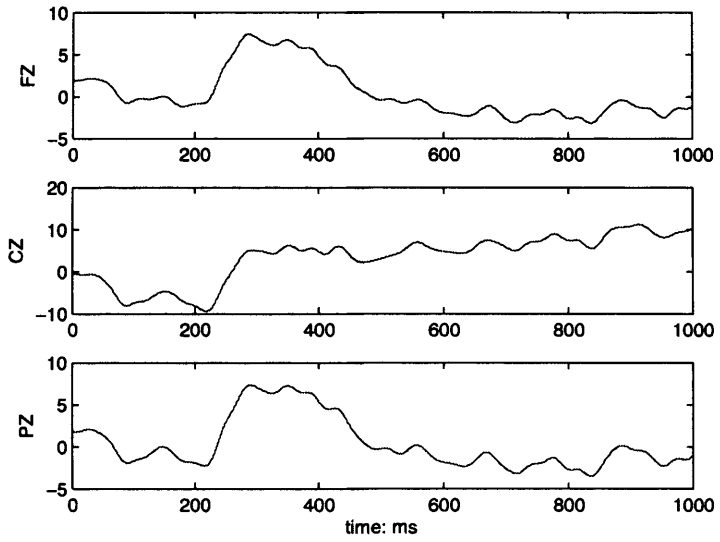

Figure 4.11. Three channel ERP of a schizophrenic patient obtained by averaging 40 related events.

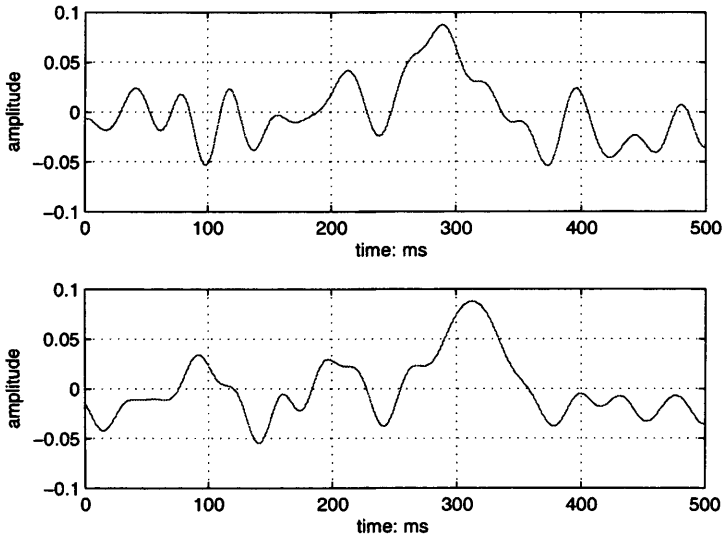

Figure 4.12. The separated P3a and P3b from the signals of Figure 4.11 using the proposed CBSS algorithms.

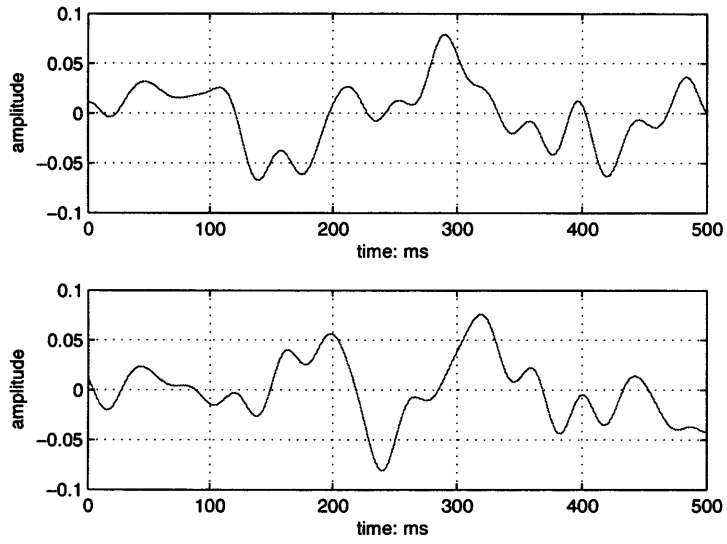

Figure 4.13. The separated P3a and P3b from the signals of Figure 4.11 using the original Infomax algorithm.

proportional to the distance [33]:

$$cor(\mathbf{X}, \mathbf{s}_k) =< \mathbf{X} \cdot \mathbf{s}_k >= \mathbf{H}\mathbf{S}\mathbf{s}_k^T \tag{4.1.13}$$

where $\mathbf{X} = \mathbf{HS}$, $\mathbf{H}$ describes the forward model for which the magnitude of a source attenuates with $\frac{1}{d_j^2}$ and $\mathbf{s}_k$ is the vector of all sample values of the k_{th} source. It has to be noted that the sources must be uncorrelated for the method to be efficient. After computing the correlation the values are normalised and converted to distances by the following:

$$d_j = \frac{1}{\sqrt{cor}} \tag{4.1.14}$$

It has to be noted that this approach does not provide a valid source reconstruction since it ignores the conductivity properties of the brain but it can be used to distinguish between sources in relatively different locations. Index j represents the three electrodes that have maximum correlation coefficients with source k, $k = 1, 2, \ldots, n$, shows the source number. In this equation all the variables except $\mathbf{f}_k$ are known.

The next step is to convert the above scenario to a mathematical problem, which requires to calculate the coordinates of an unknown point when both of the coordinates of d points and the distances of the unknown point from the given points are known. This problem is clearly equivalent to finding the intersection point(s) of d spheres in R^d. The points are the solution to the following least squares problem which can be obtained from [83]:

$$\min \quad S(\mathbf{f}_k), \quad \text{with } \mathbf{f}_k \in \Re^d, \tag{4.1.15}$$

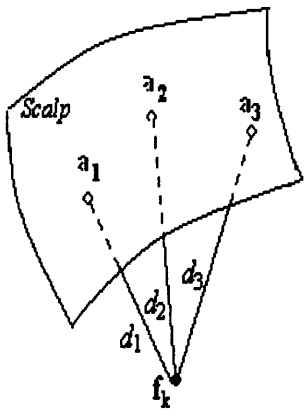

Figure 4.14. Part of the scalp including the electrode locations, a1, a2 and a3, and the location of the source k, $\mathbf{f}_k$, to be identified.

where

$$S(\mathbf{f}_k) = \sum_{j=1}^{3} \{\|\mathbf{f}_k - \mathbf{a}_j\|_2 - d_j\}^2 \tag{4.1.16}$$

In this study we applied CBSS to the data from 10 different people, 5 patients and 5 control subjects. The result of the localisation of the P3a and P3b components is shown in Figure 4.15 for five schizophrenic patients and in Figure 4.16 for five control subjects. It is evident that the P3a and P3b for a schizophrenic patient are closely and irregularly located, whereas for a control subject the P3a and P3b are located in distinct regions.

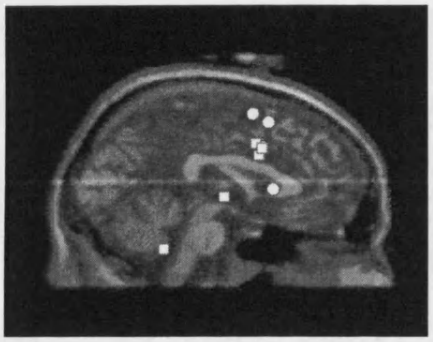

Figure 4.15. Localisation result for schizophrenic patients. The circles, *o*, correspond to the P3a and the squares , □, to P3b. The P3a and P3b are closely and irregularly located following no specific pattern.

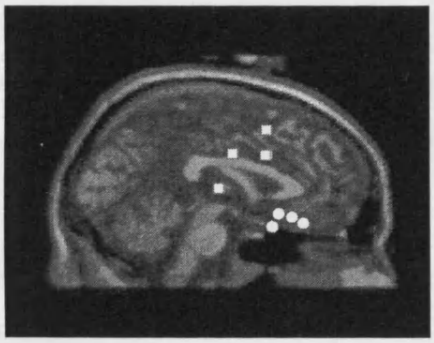

Figure 4.16. Localisation result for normal subjects. The circles, *o*, correspond to P3a and the squares, □, to P3b. The P3a and P3b sources are located in distinct regions in the brain.

4.2 CBSE

In Blind Signal Extraction (BSE) the signal sources are extracted sequentially. The main advantages of BSE over BSS can be the following:

- The algorithm concentrates only on the signals of interest based on their properties
- In general BSE algorithms are much simpler.
- They can be easily modified for a number of situations.

However, it does not give an estimate of the mixing matrix. Again we use prior knowledge of the shape and latency of P300 signals to obtain the signals. This is done by using a constraint function, which is imposed on the original BSE cost function.

We use the normalised kurtosis cost function, which estimates the deviation of a random variable from Gaussianity. For kurtosis equal to zero the signal is Gaussian, for positive values it is super-Gaussian and for negative values it is sub-Gaussian. From the Central Limit Theorem it is known that a signal consisting of a mixture of different signals tends to have a Gaussian distribution. So, by maximising the absolute value of the kurtosis of the output signal, we create a more unmixed output signal.

In the BSE model the output of the algorithm is described as :

$$\mathbf{y} = \mathbf{w}^T\mathbf{X} \tag{4.2.1}$$

where $\mathbf{y}$ is the output vector of size $1 \times T$ (T is the number of samples), $\mathbf{w}$ is the unmixing vector ($n \times 1$, where n is the number of electrode signals) obtained by the algorithm and $\mathbf{X}$ is the data matrix ($n \times T$)

consisting of the electrode signals.

The cost function is:

$$J_C(\mathbf{w}) = -\frac{1}{4}|kurt(\mathbf{y})| \tag{4.2.2}$$

where $kurt(\mathbf{y})$ is the normalised kurtosis and is given by:

$$kurt(\mathbf{y}) = \frac{E(|\mathbf{y}|^4)}{E(|\mathbf{y}|^2)} - 3 \tag{4.2.3}$$

where $E()$ denotes the statistical expectation. This leads to the following online adaptation rule:

$$\begin{aligned} \mathbf{w}(k+1) &= \mathbf{w}(k) + n(k)\varphi(y(k))\mathbf{x}(k) \\ \varphi(y(k)) &= b(\frac{m_2(y(k))}{m_4(y)(k)}y(k)^3 - y(k))\frac{m_4(y(k))}{m_2^3(y)(k)} \\ m_q(k) &= (1-n_0)m_q(k-1) + n_0|y(k)|^q \end{aligned} \tag{4.2.4}$$

where k is the iteration index (sample number), m_q is the q_{th} moment and $n_0 \epsilon (0, 1]$ adjusts the influence of the previous estimate of the moment and the current estimate.

The constraint is imposed upon the normalised kurtosis cost function and utilises prior knowledge of the P300 shape and latency. The procedure is to obtain a reference P300 signal and in turn a vector $\mathbf{w}_{opt}$ which minimises the Euclidean distance between that reference and the data in the following way:

$$d_{e1} = ||\mathbf{y}_{spike} - \mathbf{w}_{opt}\mathbf{X}||_2^2 \tag{4.2.5}$$

Then, we want to minimise the distance of the obtained $\mathbf{w}$ from (4.2.4) and $\mathbf{w}_{opt}$ from (6.9). So, we aim to minimise:

$$d_{e2} = ||\mathbf{w}_{opt} - \mathbf{w}||_2^2 \tag{4.2.6}$$

This is the constraint cost function incorporated into the original cost function (4.2.4) according to the theory of penalty parameters. Hence, the adaptation rule becomes:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + n(k)\varphi(y(k))\mathbf{x}(k) + K(\mathbf{w}(k) - \mathbf{w}_{opt}) \tag{4.2.7}$$

where K is the penalty parameter. It should not be too high so it does not overcome the effect of the main cost function or too low for the constraint to be ignored [56].

4.2.1 Results on P3a and P3b

The recording was the same as in section (4.1.2). After obtaining the data, they were temporally averaged for all event related trials (40 events). The algorithm automatically obtains the reference signal, $\mathbf{w}_{opt}$ and hence $\mathbf{y}_{LS}$. The algorithm was applied to many sets of data and it extracted the desired components successfully. The obtained P3a and P3b were more highlighted compared to the reference signal and their shape was more in agreement with typical P3a and P3b shapes. Some typical obtained signals can be seen in Figures 4.17 and 4.19 while the corresponding $\mathbf{y}_{LS}$ signals are shown in Figures 4.18 and 4.20 and the spatio-temporal averaged reference signal[2] is shown in Figure 4.21. As can be seen from the figures the constrained method obtains good representations of P3a and P3b. Their respective latencies are

[2]This signal is used to obtain the $\mathbf{y}_{ref}$ signals for P3a and P3b.

in agreement with prior physiological research and their shapes are more smooth than those of the $\mathbf{y}_{LS}$ signals. This is expected since the algorithm tries to obtain an output close to the $\mathbf{y}_{LS}$ but also as less mixed as possible.

By comparing the resulting signals with the unconstrained case useful insights can be obtained as to the selection of the appropriate parameters for the algorithm (such as the K penalty parameter, the learning rate n and the n_0 paratemter). If K is set to zero the algorithm is unconstrained and the resulting output is not the desired one. Gradually increasing the K parameter starting from a small value (about 10^{-5}) the algorithm's behaviour can be observed. At very small values the influence of the constraint is minimal and does not produce valuable results. High values tend to make the algorithm crash. A practical value that produces good results while the signal is not very close to the reference is 10^{-4}. In fact, the value of K can be adapted and updated iteratively according to the changes in the gradients of J_c and J_m. The learning rate was set to 10^{-3} and it was reduced every iteration by 1%. The parameter n_0 was set to 0.5.

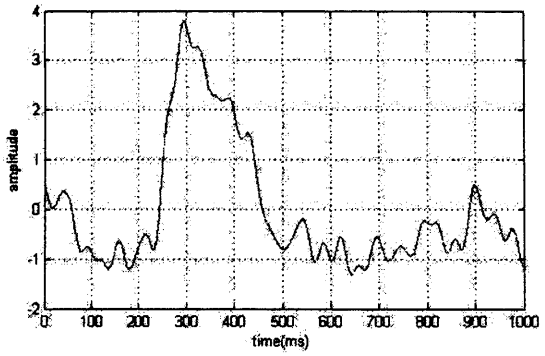

Figure 4.17. The P3a signal obtained by using the constrained BSE algorithm.

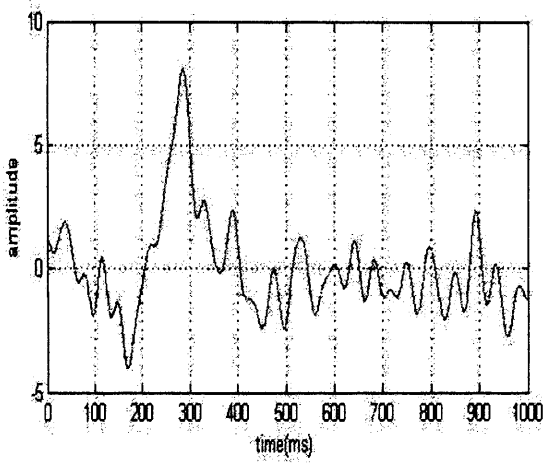

Figure 4.18. The $\mathbf{y}_{LS}$ signal used to obtain the signal of Figure 4.17.

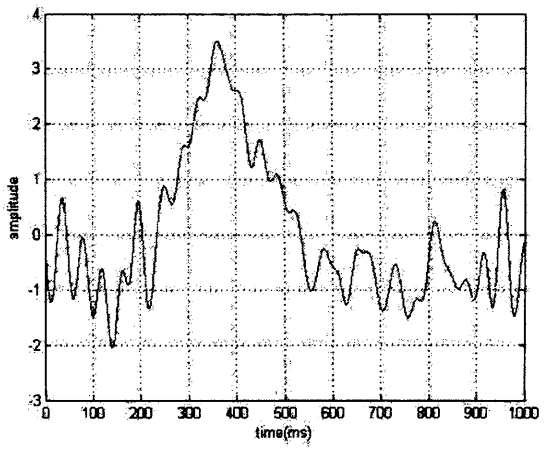

Figure 4.19. The P3b signal obtained by using the constrained BSE algorithm.

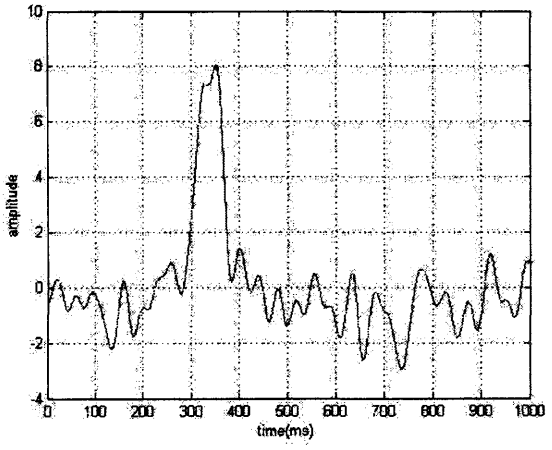

Figure 4.20. The $\mathbf{y}_{LS}$ signal used to obtain the signal of Figure 4.19.

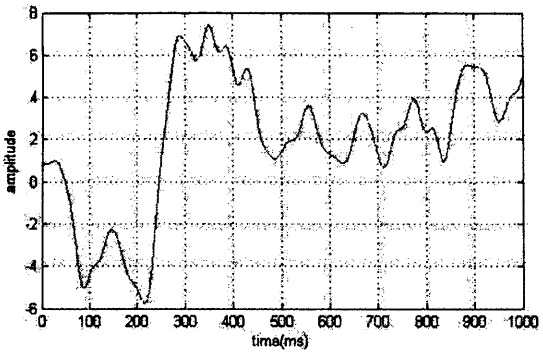

Figure 4.21. Spatio-temporal averaged EEG reference signal.

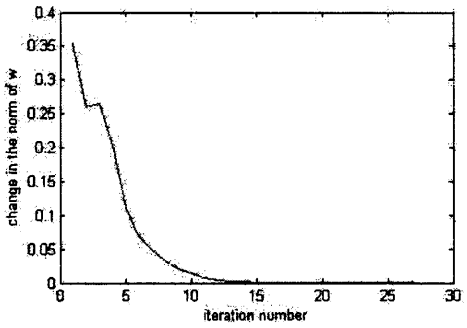

Figure 4.22. Convergence of the algorithm for the signal of Figure 4.17.

4.3 Conclusions

In this chapter we described the two methods we have developed that utilise ICA. Our first method, CBSS, extracts a number of sources at the same time, while requiring that one or more of them are close to our reference signal(s). The results were satisfactory and produced better representations of the original sources in our simulated experiments, having significantly less error than the sources produced by the original unconstrained Infomax algorithm. It is also superior to the original Infomax algorithm, when our reference signal is quite different from the original source. This experiment simulated the real-data case where we do not know exactly the waveform of the original sources. We

applied our CBSS algorithm to real EEG data, containing both the P3a and P3b ERP subcomponents. It created more realistic looking output sources. Also, we applied the algorithm on various datasets from different subjects and, as mentioned before, CBSS always produces outputs that strongly exhibit features of P3a and P3b, while Infomax may fail on a number of occasions. The main downside of CBSS.

The second method we developed based on ICA is the CBSE algorithm. In contrast with CBSS which extracts all sources at the same time, CBSE extracts only one, again requiring it to be close to our reference signal. The main advantage is that it is easier to use, with the only output being the desired one. Because there is only one output it is not necessary to search through the produced outputs to find which one is the desired one. In other words, the user can select, for example, to extract the P3a and the output will exhibit the properties of a P3a source.

Although the two methods were succesful in extracting the ERP components from the frame averaged ERP, they did not perform well for single trial EEG analysis. That is because the ERP components are not accentuated enough in individual frames for ICA to produce reasonable outputs. Hence, in the next chapter we develop a method for single trial EEG.

Chapter 5

SINGLE-TRIAL ERP COMPONENT ESTIMATION BASED ON ERP TEMPLATES

In the previous chapter we developed a method that extracts the P3a and P3b subcomponents from ERP averaged data. In this chapter we develop an algorithm that is especially designed for single-trial data. This method uses prior information of the shapes of the ERP components. As it is widely observed, ERP signals are transient waves time-locked at approximate latencies after an ERP eliciting event. A good approximation for the ERP components is to model them as Gaussian spikes (with certain latencies and variances) [49] [84]. The main motivation is that we need to avoid using any information that comes from ERP averages. Recall that in the previous chapter we used ERP averages to obtain the reference signals. One advantage of using a parametric approach, i.e. Gaussian-shaped spikes where the parameters are the latency and variance, is that the algorithm can be used in real-time or in cases where there are only a few number of trials, and hence we cannot use ERP averages. The Gaussian spikes then serve as reference signals onto which the EEG data are projected. Thus, we use the

spatio-temporal information which exist in the data to find the closest representation of the reference in the data. We use least squares to extract the ERP components. Therefore, our algorithm can provide a description of all ERP components for every trial.

We also aim to acquire an estimate of the locations of these components. By estimating all the existing ERP components in the data it becomes possible to obtain the implied mixing matrix. The locations of the sources are computed using an extension of the localisation method used in Chapter 4.

Given the approximate shape of the components as a priori, the method developed here has a superior performance compared to other methods. It is fast and robust in terms of noise since an estimation of the reference is expected to be in the data. Hence, it is suitable for single-trial data. Moreover, it is robust to model errors in the sense that it tries to find the closest match to the reference signal. No information about the number of sources is needed beforehand since it is estimated by the algorithm. Finally, it is consistent since it provides unique results which do not depend on any initial conditions. However, its application is limited since it can only work with spike-like signals. Also, it estimates the sources and the forward or mixing matrix with a scaling indeterminacy (similar to the BSS/ICA indeterminacy). This, however, is a minor problem considering that all the necessary information lie within the relative magnitudes of the elements of the forward vectors.

In the next section the procedure to obtain the sources and forward matrix is explained. Then, we show the performance and the details of the algorithm by illustrating its performance on simulated data. The

single-trial performance is also shown by applying the algorithm to a number of data of normal and schizophrenic patients and evaluating their differences. Finally, the paper concludes with a summary of the findings and the performance of the algorithm. The work in this chapter is supported by the following publications [85] [86] [87].

5.1 Proposed Method

Here we consider a more general model for the EEG signal by ignoring the dipole orientations; we model it as an $n \times T$ matrix (n is the number of electrodes and T is the number of time samples):

$$\mathbf{X} = \mathbf{HS} = \sum_i \mathbf{h}_i \mathbf{s}_i \tag{5.1.1}$$

where $\mathbf{H}$ is the $n \times m$ forward matrix of the m sources $\mathbf{s}_i$ and $\mathbf{h}_i$ are the forward vectors of the sources. The sources $\mathbf{s}_i$ are considered as the ERP components that are directly relevant and time-locked to the stimulus; we treat every other activity as noise. They are thought to have a transient spiky shape. The perfect source reconstruction would occur if we could design m filters $\{\mathbf{w}_i\}$ (note that we do not know the sources or their number beforehand) that satisfies:

$$\mathbf{s}_i = \mathbf{w}_i^T \mathbf{X} \tag{5.1.2}$$

This can be accomplished if each filter $\{\mathbf{w}_i\}$ is designed to minimise:

$$\mathbf{w}_{opt} = \arg\min_{\mathbf{w}_i} ||\mathbf{s}_i - \mathbf{w}_i^T \mathbf{X}||_2^2 \tag{5.1.3}$$

which would, however, require prior knowledge about the sources $\mathbf{s}_i$, which is not feasible in practise. Nonetheless we have a reasonable idea of how the sources look like:

$$\mathbf{s}_i = exp(-(t - \tau_i)^2/\sigma_i^2) \tag{5.1.4}$$

where τ_i is the latency of the ith source and σ_i its width. The width is chosen as the average width of the P3a and P3b subcomponents and it does not have to be accurately estimated since the LS solution (6.3.1) will find the closest match.

5.1.1 Estimating the source signals

In contrast with the CBSS method which obtains an estimate of the source that is close to our reference signal, here the aim is to find a good description of the source directly. For example, when we used a reference for the P3a, the obtained signal usually had different latency than the reference. Hence, the idea here is to vary the latency of the reference. By using a Gaussian shaped reference spike, with variable parameters, we produce the same signal but with different latency. By varying the latency of the reference we can create many reference signals and hence produce many estimates of the sources. Some of those estimates will have similar properties, most notably their latency. That comes from the fact, as mentioned previously, that the extracted output source will be as close to the reference signal as possible. Therefore, since many reference signals will be close to a source, we expect to obtain many estimates representing the same source. These estimates will be similar to each other and we expect to observe groups of estimates

belonging to separate sources. From those estimates we can do two things; firstly, we can average all the estimates that belong to the same group, each group consisting of all the estimates with similar latencies. Secondly, within each group we find the source with the minimum error $e(\mathbf{w}_j)$:

$$e(\mathbf{w}_j) = ||\mathbf{s}_j - \mathbf{w}_j^T\mathbf{X}||_2^2 \quad j \ \epsilon \ G_x \tag{5.1.5}$$

where j is the index that runs through the sources of one group and G_x denotes group x of the estimated sources. The latter way is what we use to select the sources. So, in order to estimate the sources, $\mathbf{s}_i$, we create a large number of reference signals $\mathbf{r}_j = exp(-(t-j)^2/\sigma^2)$, say T, each having a latency at a different time sample. Hence, we create T of those references (Figure 5.1) shows an example one).

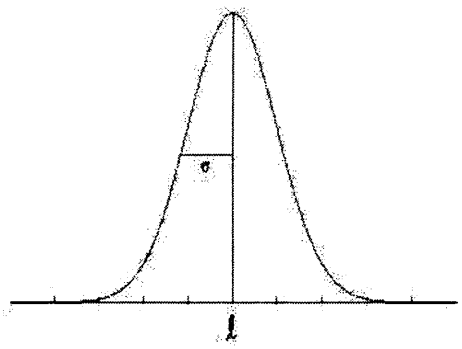

Figure 5.1. Gaussian spike with variance σ^2 and latency l, used to model the ERP components

Then we compute T (note that $T \gg m$) filters such as:

$$\mathbf{w}_j^T = \arg\min_{\mathbf{w}_j} ||\mathbf{r}_l - \mathbf{w}_j^T\mathbf{X}||_2^2, \quad \mathbf{y}_j = \mathbf{w}_j^T\mathbf{X} \quad j = 1...T \tag{5.1.6}$$

Again we note that we assume that the signals $\mathbf{y}_j$ which have a similar

latency to that of a true source $\mathbf{s}_i$ correspond to that source only. Also, since the ERP components have distinct latencies we expect the signals $\mathbf{y}_j$ to be grouped into m clusters, equivalent to the number of sources. To cluster the T signals $\mathbf{y}_j$ we use the following algorithm:

for j=1 to T

- measure $l(j)$, the latency of $\mathbf{y}_j$
- if $l(j) - l(j-1) < \beta$, then $\mathbf{y}_j$ and $\mathbf{y}_{j-1}$ belong to the same group, β is a threshold selected empirically
- if $l(j) - l(j-1) > \beta$, then $\mathbf{y}_j$ belongs to a different group than $\mathbf{y}_{j-1}$

Then, within each group we select the signal $\mathbf{y}_j$ with the minimum error as defined in Equation (5.1.5) and obtain c signals $\mathbf{y}_c$, and it is expected that $c = m$. Note that the obtained sources can be a scaled version of the original sources $\mathbf{s}_i$ because their scales depend on the amplitude of the reference signals.

This procedure extracts spike-like waves from the data. The main advantage is that by sweeping all the data at relatively small intervals every spike-like wave will be extracted. This coincides with ERP signals, which are known to have transient waveforms. Also, this method is robust to differences in modelling of the spikes. It extracts the sources which are most similar to the reference signal.

5.1.2 Forward matrix

To obtain the forward vectors (also known as scalp maps, the columns of the forward matrix $\mathbf{H}$) we use the following novel procedure. First

we compute $\mathbf{R}$ as the cross-correlation between the data matrix and the output sources matrix:

$$\mathbf{R}_{xy} = \mathbf{X}\mathbf{Y}^T = \mathbf{H}\mathbf{S}\mathbf{Y}^T \tag{5.1.7}$$

where $\mathbf{Y}$ is a matrix whose rows are the signals $\mathbf{y}_c$. The estimated sources $\mathbf{Y}$ can be written as $\mathbf{Y} = \mathbf{D}\mathbf{S}$, where $\mathbf{D}$ is a diagonal matrix describing the scaling factor of each of the sources:

$$\mathbf{D} = \begin{pmatrix} d_1 & 0 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 & 0 \\ 0 & 0 & . & 0 & 0 \\ 0 & 0 & 0 & . & 0 \\ 0 & 0 & 0 & 0 & d_c \end{pmatrix} \tag{5.1.8}$$

If we multiply $\mathbf{R}$ by the autocorrelation matrix of $\mathbf{Y}$ we can obtain a scaled version of the scalp maps:

$$\begin{aligned} \mathbf{R}\mathbf{R}_y^{-1} &= \mathbf{H}\mathbf{S}\mathbf{Y}^T(\mathbf{Y}\mathbf{Y}^T)^{-1} \\ &= \mathbf{H}\mathbf{D}^{-1}\mathbf{Y}\mathbf{Y}^T(\mathbf{Y}\mathbf{Y}^T)^{-1} = \mathbf{H}\mathbf{D}^{-1} \end{aligned} \tag{5.1.9}$$

The permutation does not have any effect on the solution since the ordering of the sources is arbitrary. Hence, the i_{th} scaled scalp map will correspond to the scaled i_{th} source.

5.1.3 Least squares estimation of the position of the source

We now show how to calculate the position of the source using the modified LS method. The first step is to convert the elements of the $\mathbf{H}$ matrix to estimates of the distances between the electrodes and the

sources. We use an isotropic propagation model of the source where it attenuates with the $3rd$ power of the distance [33]. To convert to the distance we perform the following operation:

$$r_j = \frac{1}{h_j^{1/3}} \tag{5.1.10}$$

where h_j is the j_{th} elemement of a specific column of the $\mathbf{H}$ matrix. The point q is the solution to the following least squares problem:

$$E(\mathbf{q}, M) = \sum_{j=1}^{n} [M||\mathbf{q} - \mathbf{A}_j||_2 - r_j]^2 \tag{5.1.11}$$

where we know $\mathbf{A}_j$ the positions of the electrodes, and r_j the scaled distances, $E(q)$ is the squared error. The factor M denotes the scaling that arises from the algorithm as discussed in the previous section and also from the fact that our model (5.1.11) did not consider the electrical properties of the head. We desire to minimise the error function and it should ideally be zero. The derivatives with respect to $\mathbf{q}$ and M are:

$$\nabla E_{\mathbf{q}} = 2 \sum_{j=1}^{n} (\mathbf{q} - \mathbf{A}_j)(M^2 - M \frac{r_j}{||\mathbf{q} - \mathbf{A}_j||_2}) \tag{5.1.12}$$

$$\nabla E_{\mathbf{M}} = 2 \sum_{j=1}^{n} M||\mathbf{q} - \mathbf{A}_j||_2^2 - ||\mathbf{q} - \mathbf{A}_j||_2 r_j \tag{5.1.13}$$

We employ an iterative procedure to estimate q:

$$\mathbf{q} = \mathbf{q} - l \nabla E_{\mathbf{q}} \tag{5.1.14}$$

and

$$M = M - m \nabla E_M \tag{5.1.15}$$

where l and m are the learning rates. The solution for $\mathbf{q}$ and M is unique subject to having an appropriate number of electrodes. This method performs better than the localisation method of last chapter.

5.2 Results for simulated EEG

In this part we evaluate the performance of the algorithm from using synthetic EEG data. We use a similar setup to that of the previous chapter; we mix 3 sources in 30 electrodes. Again, the head model is a realistic model obtained through the Brainstorm software platform [72]. The width of the sources is $\sigma = 0.1$. In Figure 5.2 we show the sources, and in Figure 5.3 we show an example mixture.

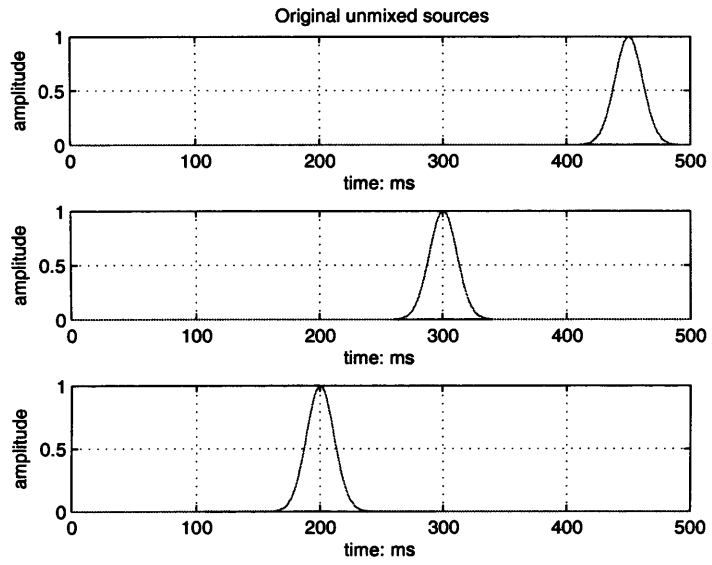

Figure 5.2. Three Gaussian-spike sources peaking at 200ms, 300ms and 450ms.

The first experiment is to apply the algorithm assuming we know exactly the width of the original sources. In other words the σ of Equation (5.1.4) is the true one. We also assume there is no noise for the first experiment. In Figure 5.4 we see the resulting scatter diagram.

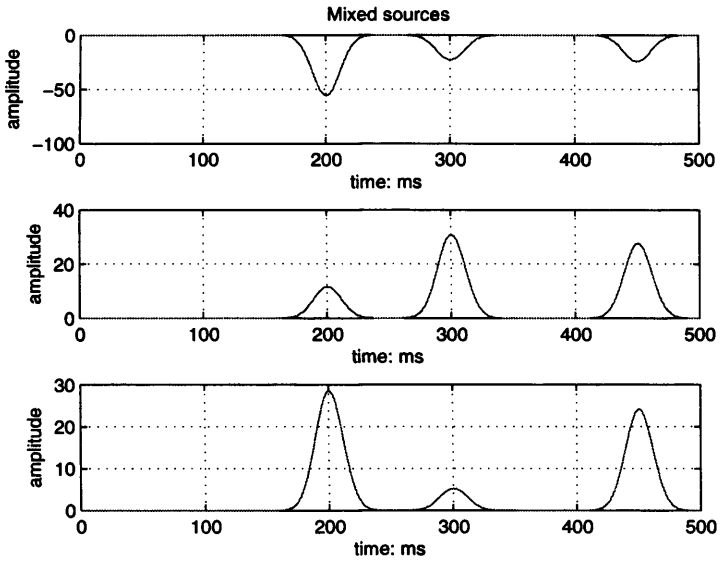

Figure 5.3. Three mixtures from the sources of Figure 5.2.

The horizontal axis shows the latency of the produced signal, as in Equation 5.1.6; the vertical axis is the peak amplitude of the produced signal. Note that the amplitude is not important since the algorithm tries to extract a signal with an amplitude as close to the amplitude of the reference signal. In Figure 5.5 we use a slightly different width, $\sigma = 0.5$ and in Figure 5.6 we use $\sigma = 5$. For illustration we show the three different spikes in Figure 5.7.

It is evident from these plots, for the noiseless case, even for a quite different reference signal, that the algorithm estimates the latencies of the sources perfectly.

In Figure 5.8 we show the obtained sources using the proposed algorithm for the three different cases.

Next experiment is to demonstrate the behaviour of the algorithm when the mixtures are corrupted with noise. In the ERP context this noise can be background EEG or measurement noise. In Figure 5.9 we show the scatter diagram for $12dB$ SNR and in Figure 5.10 for $0dB$

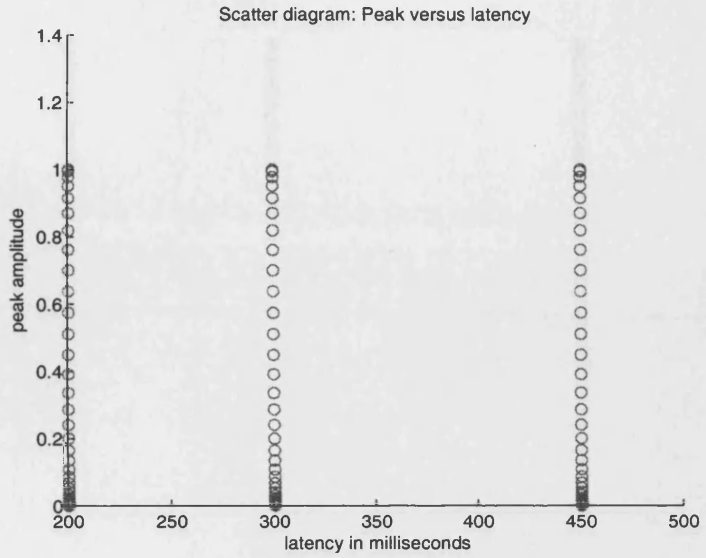

Figure 5.4. Scatter diagram of the latencies of the extracted sources and their peak amplitudes. Here we use the actual width of the source $\sigma = 0.1$. The clustering of the extracted sources at the correct latencies is evident.

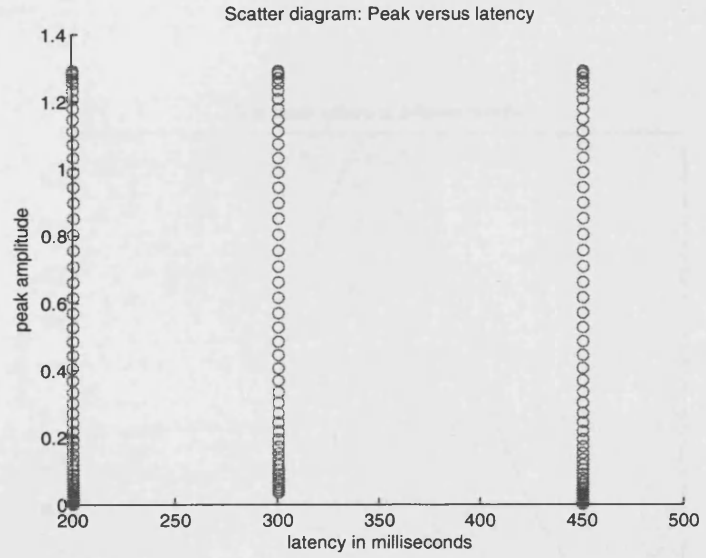

Figure 5.5. Scatter diagram of the latencies of the extracted sources and their peak amplitudes. Here we use a higher width than the actual width of the source $\sigma = 0.5$. The resulting clustering of the extracted sources at the correct latencies is evident, although in this case there were less sources at 300ms.

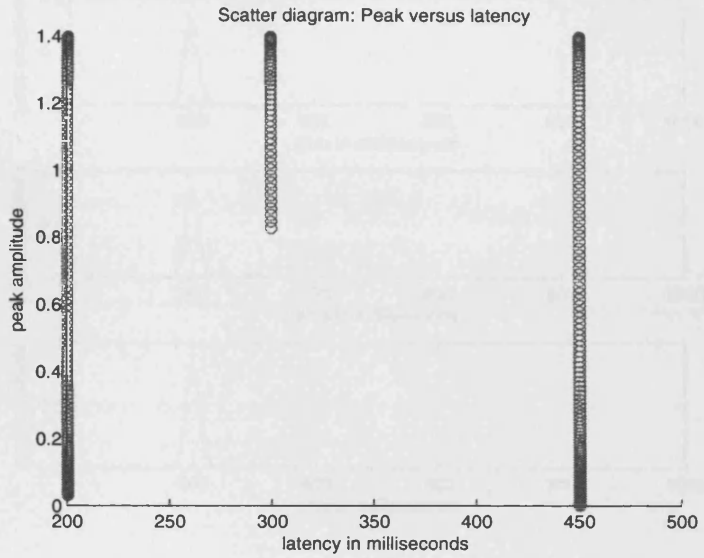

Figure 5.6. Scatter diagram of the latencies of the extracted sources and their peak amplitudes. Here we use a substantially different width from that of the actual source $\sigma = 5$. The clustering of the extracted sources at the correct latencies is still evident. Here, many of the extracted $\mathbf{y}_l$ signals that previously had latency of 300ms have moved to the other two clusters.

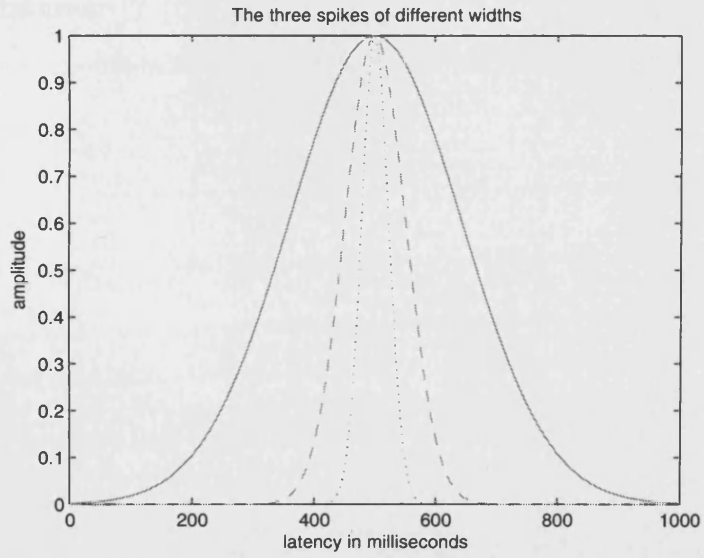

Figure 5.7. The three Gaussian spikes used in the simulations with widths $\sigma_1 = 0.1$, $\sigma_2 = 0.5$, *and* $\sigma_3 = 5$

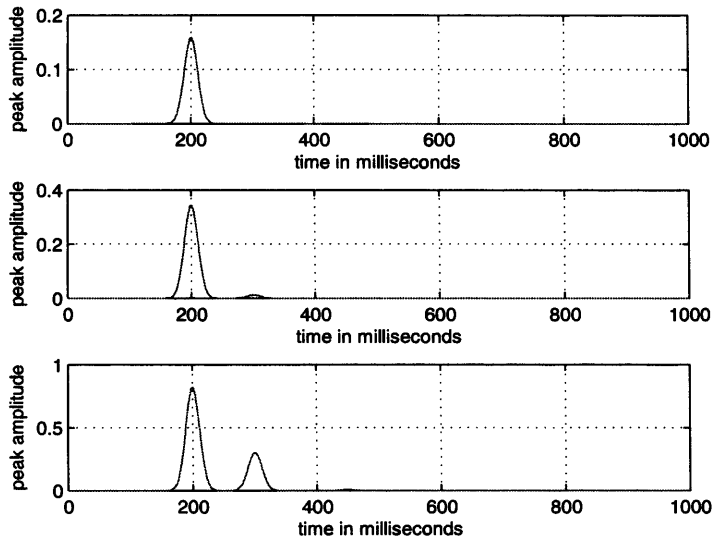

Figure 5.8. The estimated source signal using the three reference signals of Figure 5.7. Although the reference signals are different from the true source the latency however is the correct one.

SNR. In Figure 5.11 we show the result of clustering of the solutions for when two of the sources overlap (see Figure 5.13), simulating the overlap between a real P3a and P3b, and in Figure 5.12 we add noise to the mixtures.

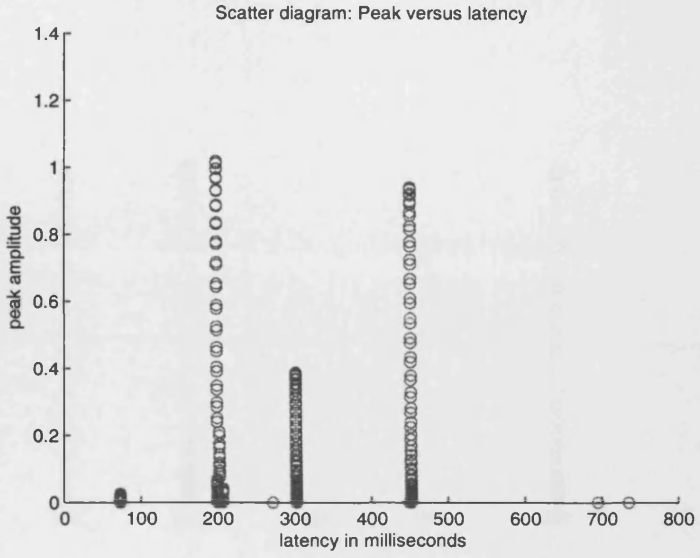

Figure 5.9. Scatter diagram of the latencies of the extracted sources and their peak amplitudes. The clustering of the extracted sources at the correct latencies has degraded in this case for $12dB$ SNR.

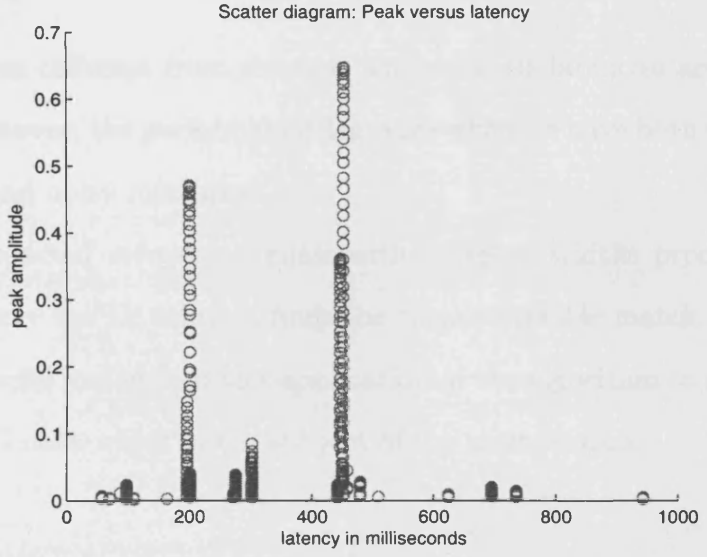

Figure 5.10. Scatter diagram of the latencies of the extracted sources and peak amplitudes. The clustering of the extracted sources at the correct latencies has degraded in this case for $0dB$ SNR.

5.2.1 Discussion

The algorithm performs very well in terms of latency estimation even for noisy mixtures. It exhibits robustness and even if the extracted

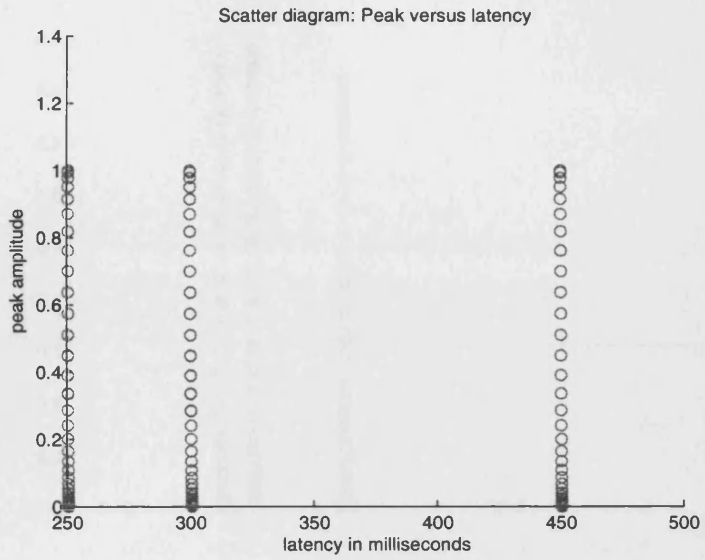

Figure 5.11. Scatter diagram of the latencies of the extracted sources and their peak amplitudes. The clustering of the extracted sources at the correct latencies is perfect for this case where the two sources are correlated.

sources are different from the true sources their latencies are the true ones. However, the performance degrades when we have both correlated sources and noisy mixtures.

As expected reference signals with different widths produce good results since the LS solution finds the closest possible match. This fact is very useful and encourages application of the algorithm to real single trial EEG data which is the subject of the next section.

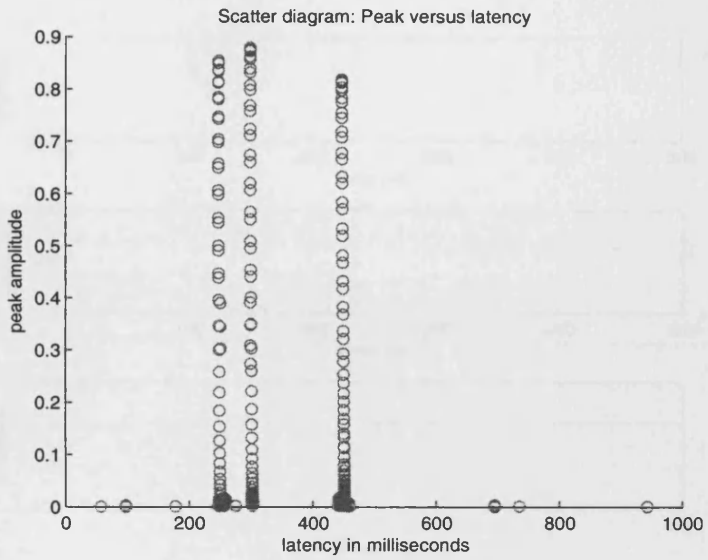

Figure 5.12. Scatter diagram of the latencies of the extracted sources and their peak amplitudes. The clustering of the extracted sources at the correct latencies has degraded in this case for $0dB$ SNR and when the two sources are correlated.

5.3 Results on real single trial EEG data

When applied to real data the algorithm follows a step by step procedure. The procedure will be automatic and will return important features of the extracted components. Here we describe the full algorithm although some of the functions can be skipped for a particular application. For example, it may be only necessary to obtain just the latency of the P3a subcomponent. In that case we create a T_{P3a} number of filters, hence we need a T_{P3a} number of reference signals, ranging from $j = 250ms \;\; to \;\; 300 \;\; ms$. Hence we skip the scalp map and location estimation. The steps of the full procedure for all components and their scalp maps and locations are:

- Choose the width σ, the number of reference signals to be created, maximum T, and the latency range of those references.

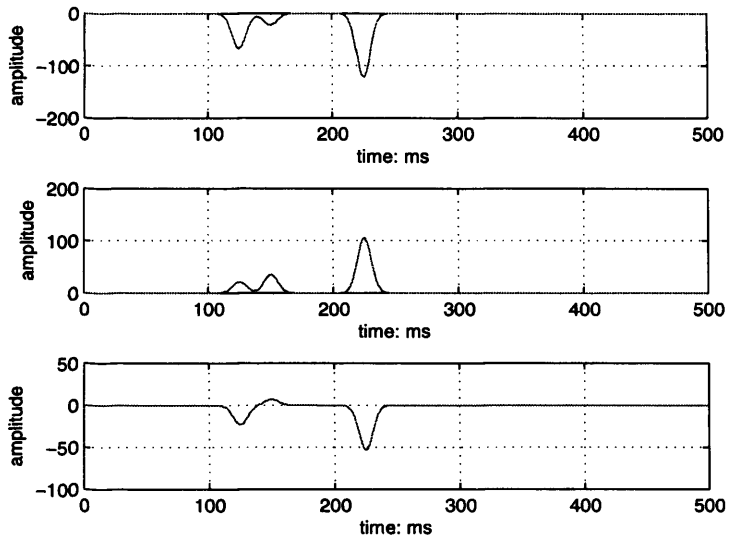

Figure 5.13. The extracted sources for the case of Figure 5.11. In this case the first two sources overlap in time simulating the overlap between the P3a and P3b subcomponents.

- Compute the T filters $\mathbf{w}_l$, each corresponding to a reference signal.
- Measure the latency of each output signal $\mathbf{y}_l$, group them in c clusters according to the algorithm in Section 5.1.1, and average the signals within each cluster.
- Measure the latency, compute the scalp maps for the averaged outputs of each cluster.
- Compute the $3D$ locations of the desired components using the method of Section 5.1.3.

5.4 Single-Trial P300 estimation

In this section we show the performance of the algorithm on real single trial EEG data. In Figure 5.14 we show the scatter diagram for a schizophrenic patient. Observing the figure we see various clusters of

components. In Figure 5.15 we have zoomed in to the latencies of interest, around P3a and P3b. In Figure 5.16 we show the zoomed in scatter diagram of a different trial.

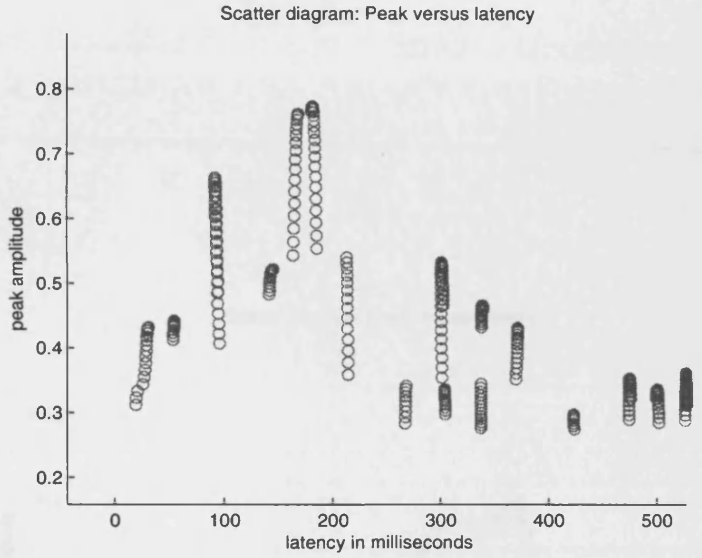

Figure 5.14. Scatter diagram of the latencies of the extracted sources and their peak amplitudes.

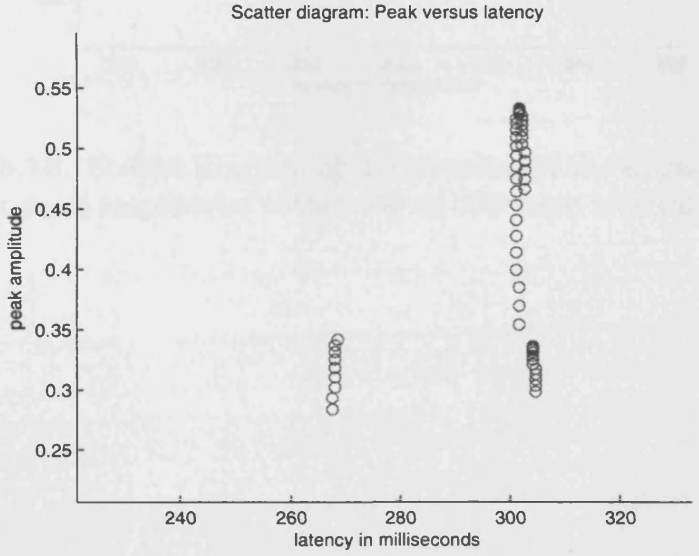

Figure 5.15. Scatter diagram of the latencies of the extracted sources and their peak amplitudes within 240 to 320 msec interval.

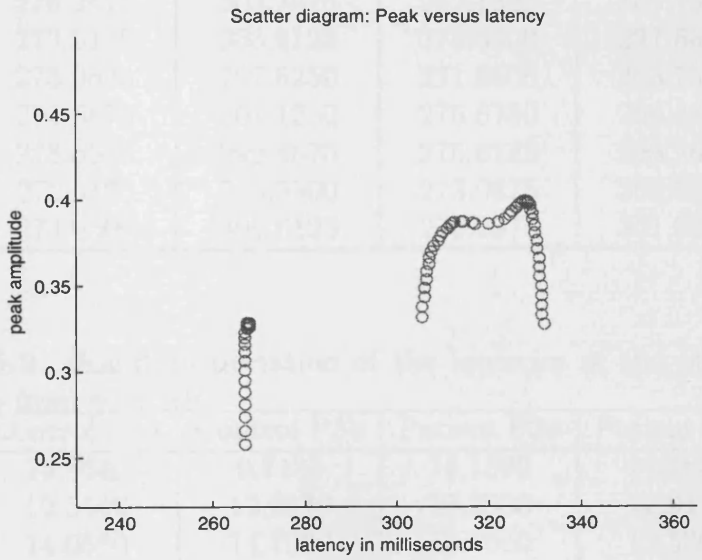

Figure 5.16. Scatter diagram of the latencies of the extracted sources and their peak amplitudes within 240 to 360 msec interval.

We applied the algorithm to the data from 10 schizophrenic patients and 10 normal subjects. In Table 5.1 we show the mean latencies of the P3a and P3b of 40 trials. In Table 5.2 we show the standard deviation of those latencies.

Table 5.1. Average latencies of the obtained P3a and P3b from 40 trials.

Control P3a	Control P3b	Patient P3a	Patient P3b
274.4625	300.9000	275.5125	296.5125
274.9875	299.9875	277.4500	296.6375
271.5250	300.3750	275.7500	300.9500
276.2875	303.7875	272.1625	303.7500
273.5125	305.6125	275.6500	297.6625
273.0625	297.6250	271.5500	303.7500
275.6875	301.1250	276.6750	296.4875
278.6500	299.4875	275.6125	298.7625
273.3125	303.3500	273.0875	300.6500
273.6500	301.6125	273.2875	301.0875

Table 5.2. Standard deviation of the latencies of the obtained P3a and P3b from 40 trials.

Control P3a	Control P3b	Patient P3a	Patient P3b
10.8640	9.7133	12.1390	14.1630
12.5440	12.2270	16.2530	12.9170
14.0550	14.7300	10.7660	12.5300
13.0660	12.0330	12.9450	15.6520
16.4980	13.7720	11.4390	13.6550
14.6290	12.6480	15.3490	16.0510
12.3930	12.7240	16.5020	15.6420
14.9210	10.2970	8.6012	15.7280
14.7560	13.8890	12.9810	15.0160
13.6530	12.8460	12.6620	15.1100

The means of latencies and standard deviations are shown in tables 5.3 and 5.4 respectively.

Table 5.3. Means of the latencies of the obtained P3a and P3b from 40 trials calculated using Table 5.1.

Control P3a	Control P3b	Patient P3a	Patient P3b
274.5138	301.3863	274.6737	299.6250

Table 5.4. Means of the standard deviations of the obtained P3a and P3b from 40 trials calculated using Table 5.2.

Control P3a	Control P3b	Patient P3a	Patient P3b
13.7379	12.9821	12.4695	14.6464

Table 5.5. Means of kurtosis of the obtained P3a and P3b from 40 trials.

Control P3a	Control P3b	Patient P3a	Patient P3b
-0.0846	0.1483	0.4388	1.1923

5.4.1 Discussion

The main findings of the current study are supported by findings via clinical studies and examinations well supported in the literature. One main result is that control subjects and schizophrenic patients have similar latencies in terms of P3a and P3b. The latency therefore cannot distinguish between the two groups.

In contrast to the mean latencies of the two groups there are other features available only through single-trial estimation that can separate the two populations. Most importantly, the standard deviation of the latencies of the P3b subcomponent can distinguish between the two groups with 80% accuracy. We also observe that the kurtoses of the patients' P3a and P3b signals are significantly higher than that of the healthy individuals.

5.5 Single-Trial P300 location tracking for two subjects

In the paper [86] we applied our localisation algorithm to estimate the locations of the P300 subcomponents for a schizophrenic patient and for a control subject. In Figure 5.17 we see the different locations of the P3a and P3b signals for a schizophrenic patient and in Figure 5.18 the locations of the P3a and P3b for a control subject. It is observed that the locations of the control subject exhibit less variability than that of the schizophrenic patient. In this section we just aim to exhibit the performance of the algorithm and we do not aim for a qualitative explanation of the differences between patients and control subjects.

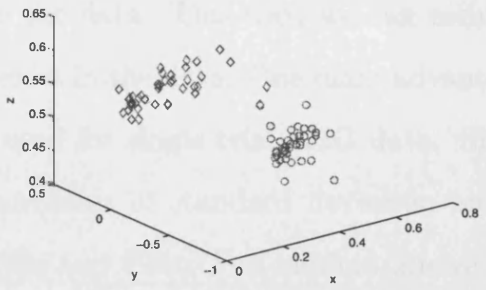

Figure 5.17. The locations of the of P3a and P3b for a schizophrenic patient. The diamonds represent the P3a and the circles the P3b. The x axis denotes right to left (positive to negative), the y axis front to back (positive to negative), and the z axis is up to down (positive to negative).

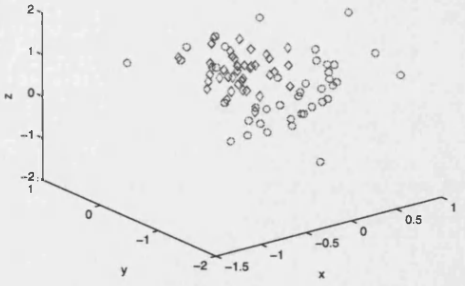

Figure 5.18. The locations of the of P3a and P3b for a normal subject are shown in this figure. The diamonds represent the P3a and the circles the P3b. The x axis denotes right to left (positive to negative), the y axis front to back (positive to negative), and the z axis is up to down (positive to negative).

5.6 Conclusions

We used a newly developed method for the latency estimation of ERP components. We performed simulations to evaluate the performance of the algorithm in generated data. The results were successful so we proceeded to real single trial EEG data. We used the algorithm to investigate P300 latency variability in schizophrenia. The algorithm works by obtaining a reference signal to be used as template for an

ERP component. Then we project a number of delayed versions of the reference onto the data. This way, we can estimate various ERP components that exist in the data. One main advantage of the method is that it can be used for single trial EEG data. Single trial analysis allowed the measurement of standard deviation and kurtosis for the subcomponents, P3a and P3b. The method allowed for various measurements and most importantly by using the standard deviation of the P3b there was 80% correct classification between controls and patients. The results are consistent with studies suggesting that latency variability does not underlie P300 amplitude reduction in schizophrenia, but rather reflects a separate abnormality with potential diagnostic value.

Chapter 6

LOCALISATION USING SPATIAL NOTCH FILTERS

Localisation of the sources in EEG has attracted great attention by itself, separately from the issue of extraction or separation. As described in Chapter 3 the traditional methods of complete source reconstruction suffer from the inverse problem. Here, we develop a method based on spatial notch filtering (SNF) which does not suffer from the inverse problem and targets specific sources. So far, a method does not exist in which a desired signal is used for localisation of brain sources. The work in this chapter is supported by the following publications [88] [89]

6.1 Spatial Notch Filtering

The method proposed in this paper uses prior information of the shapes of the ERP components and utilises the principles of spatial filtering. As it is widely observed, ERP signals are transient waves time-locked at approximate latencies after an ERP eliciting event. A good approximation for the ERP components is to model them as Gaussian shaped signals (with certain latencies and variances). The proposed spatial filter used is designed to cut-off any activity from a specific location.

It will be shown later that if a correct model is used for the ERP signals and at the same time the activity from the location of an ERP component is filtered out, then the algorithm will point to the correct location of that component.

The construction of the algorithm is similar to that of the LCMV spatial filter method. In the LCMV method the filter output is minimised, while the activity from one location passed. Here, we try to force the filter output to be as similar as possible to our reference ERP component while we cut off the activity from a location. It turns out that if the reference signal matches the true ERP and the true location is within our solution space (i.e. it is included in the grid search) then the algorithm will always attain a saddle point at the correct location. This implies that if the ERP model is not exact, for example, the reference has a slightly different latency or width, the algorithm will try to find the closest match between the reference signal and the extracted one.

The proposed method does not need any prior assumptions about the number of sources, and more importantly it does not require to have a noise subspace (i.e. overdetermined system, more sensors than sources) as the LCMV and MUSIC method do. Our method works for more sources than sensors too. In contrast to the MN methods, our algorithm is not an inverse method so there is no need for regularisation of the solution. The motivation of this method is to investigate the efficacy of our algorithm in high noise conditions and quantify how useful the application of the reference signal is. A number of different cases are examined for synthetically generated data and real EEG. Specifically, we examine the effect of noise and the effect of correla-

tion between different ERP components and explore various set-ups. Finally, we apply the algorithm to real EEG data.

6.2 Sources and EEG Model

The proposed method is based on a head and source model which describes the propagation of the brain sources to the sensors. The sources are modelled as current dipoles and their propagation to the sensors is mathematically described by an appropriate forward model [33] [90].

We model the EEG signal as an $n \times T$ matrix, where n is the number of electrodes and T is the number of time samples:

$$\mathbf{X} = \mathbf{HMS} + \mathbf{N} = \sum_{j=1}^{m} \mathbf{H}_j \mathbf{m}_j \mathbf{s}_j + \mathbf{N} \tag{6.2.1}$$

The term $\mathbf{HMS}+\mathbf{N}$ of Equation (6.2.1) is the matrix form of the model and $\mathbf{H}$ is an $n \times 3m$ matrix describing the forward mixing model of the m sources to the n electrodes and $\mathbf{N}$ is the noise matrix. $\mathbf{H}$ is further decomposed into m matrices $\mathbf{H}_j$ as:

$$\mathbf{H} = [\ \mathbf{H}_1 \ \ldots \ \mathbf{H}_j \ \ldots \ \mathbf{H}_m] \tag{6.2.2}$$

where $\mathbf{H}_j$ is an $n \times 3$ matrix whose each column describes the potential at the electrodes due to the j_{th} dipole for each of the three orthogonal orientations. For example, the first column of $\mathbf{H}_j$ describes the forward model of the x component of the j_{th} dipole when the y and z components are zero. Similarly, $\mathbf{M}$ is a $3m \times m$ matrix describing the orientation of the m dipoles and is decomposed as:

$$\mathbf{M} = \begin{pmatrix} \mathbf{m}_1 & 0 & 0 & 0 & 0 \\ 0 & ... & 0 & 0 & 0 \\ 0 & 0 & \mathbf{m}_j & 0 & 0 \\ 0 & 0 & 0 & ... & 0 \\ 0 & 0 & 0 & 0 & \mathbf{m}_m \end{pmatrix} \tag{6.2.3}$$

where $\mathbf{m}_j$ is a 3×1 vector describing the orientation of the j_{th} dipole. Finally, $\mathbf{s}_j$ which is a $1 \times T$ vector, is the timecourse of the j_{th} dipole and $\mathbf{N}$ is the combination of the measurement noise and modelling error.

In addition to the forward model, we create a model for the time-course of the ERP sources as well. It can be seen by many studies that they exhibit transient behaviour which can be modelled by a Gaussian shaped signal [49] [84]. Here, we model the ERP references as (for source i):

$$r_i(t) = exp(-(t - l_i)^2/\sigma_i^2) \tag{6.2.4}$$

where l_i is the latency of the i^{th} source and σ_i is the spike width (Figure 6.1).

Note that we define the EEG as a sum of sources, some of which are ERP components. The subscript i in Equation (6.2.4) refers to one of the m sources. Generally, ERPs are not correlated with each other and occur at distinct latencies. An exception is the P3a and P3b subcomponents, which overlap slightly but their peaks are distinct in time. To use the reference model properly, estimation of their latencies and widths is necessary. These parameters should correspond to the true ERP shape of each component. In this work we focus on the P300 subcomponents (P3a and P3b) for which an approximate estimate is

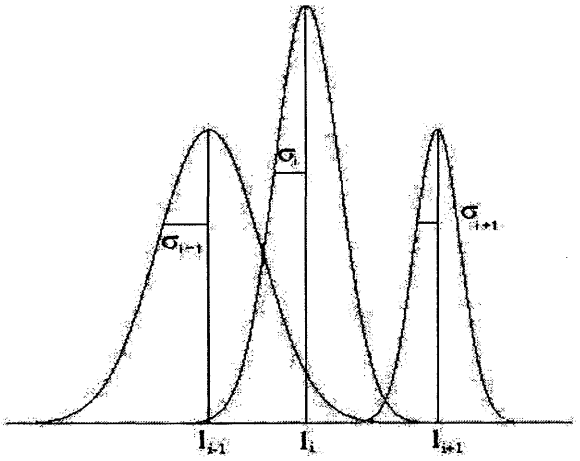

Figure 6.1. Three different Gaussian shaped spikes used to model the ERP components

made by inspection of the data [86]. The spike shapes do not need to be very accurate since the algorithm finds the closest match.

6.3 Proposed Method

The designed algorithm is based on spatial notch filtering and minimising the distance between the reference signal and a filtered version of the EEG. This is equivalent to a beamformer which tries to set a null in the location of the desired source. We perform a constrained optimisation technique in which the primary cost function is the Euclidean distance between the reference signal and the filtered EEG:

$$f_d(\mathbf{w}_i) = ||\mathbf{r}_i - \mathbf{w}_i^T\mathbf{X}||_2^2 \tag{6.3.1}$$

where $\mathbf{w}_i$ refers to the filter for extracting the i_{th} source. Also, note that $\mathbf{r}_i$ is a vector containing the T samples of $r_i(t)$. The minimum point can be obtained by the classic least squares minimisation and is given by:

$$\mathbf{w}_{i,opt} = (\mathbf{X}\mathbf{X}^T)^{-1}\mathbf{X}\mathbf{r}_i^T \tag{6.3.2}$$

This method designs a filter $\mathbf{w}_{i,opt}$, which is of dimensions $n \times 1$, and gives an estimate of the i_{th} reference signal that exists in the data. However, this procedure alone does not include any spatial information unless we obtain all the filters $\mathbf{w}_{i,opt}$ for all the sources. This way, we can construct a matrix $\mathbf{W}$, similar to the separating matrix in an ICA framework, which could be converted to the forward matrix $\mathbf{H}$ [86]. In this work we wish to estimate the location of a source which matches our reference signal without having to estimate the $\mathbf{w}_{i,opt}$ filters for

other sources.

If we minimise $f_d(\mathbf{w})$ subject to $f_c(\mathbf{w}) = \mathbf{0}$, i.e.:

$$f_c(\mathbf{w}) = \mathbf{w}^T\mathbf{H}(p) = \mathbf{0} \tag{6.3.3}$$

where $f_c(\mathbf{w})$ is a 1×3 function and $\mathbf{H}(p)$ is the forward matrix of a dipole at location p, and perform a grid search over a number of locations then the algorithm will point to the true location of the reference signal (as shown later). Note that $\mathbf{H}(p)$ denotes the forward matrix for spatial location p while $\mathbf{H}_i$ denotes the submatrix from Equation (6.2.2). For example, if source i is at location q then $\mathbf{H}(q) = \mathbf{H}_i$. By imposing such a constraint on the original cost function we force the filter output to have minimum energy from a particular location. Such a constraint function can be thought of as a spatial notch filter. A spatial notch filter removes any signal coming from a specific location. The stopband behaviour and accuracy of the filter depends on the number of electrodes and the spacing between the sources in the grid.

Hence, in fact, we are designing an adaptive null beamformer which scans a number of locations and tries to find the closest match to our reference signal. However, at the same time, we are testing for the absence of the reference signal. This expresses the main novelty of this work. When the proposed beamformer scans a particular location in the candidate source space it tries to match the reference signal while placing a null in that particular location. If the reference signal does not originate from that location the beamformer will try to find the closest match but will be influenced by the three degrees of freedom (one degree of freedom for each orthogonal dipole orientation) that have been used by placing the null in that location. So, the solution differs from the

optimum filter that extracts the reference with minimum error. In this case, our adaptive beamformer will not be equal to $\mathbf{w}_{i,\,opt}$ of Equation (6.3.2). The $\mathbf{w}_{i,\,opt}$ filter places nulls on the locations of the undesired sources (assume $\mathbf{r}_i$ is our desired source at location $\mathbf{H}_i$ and assuming no noise). That is because the LS algorithm, which tries to extract only our desired source $\mathbf{r}_i$, nulls the mixing vectors of all sources other than $\mathbf{r}_i$, which is the model of source i:

$$\mathbf{w}_{i,\,opt}^T \mathbf{H}_j \mathbf{m}_j = \mathbf{0} \quad j = 1...m \text{ and } j \neq i \tag{6.3.4}$$

where j runs over the range of sources (ERPs) as in Equation (6.3.1). $\mathbf{H}_j\mathbf{m}_j$ is the mixing vector of source j. At some point during the grid search, the forward matrix $\mathbf{H}(p)$ takes the value of the forward matrix of source i, our desired source. At this point $\mathbf{H}(p) = \mathbf{H}_i$, which will place another null in the location of our desired source. So, at this point our adaptive beamformer (with primary and constraint functions) $\mathbf{w}$ satisfies:

$$\mathbf{w}^T \mathbf{H}_j \mathbf{m}_j = \mathbf{0} \quad j = 1...m \tag{6.3.5}$$

Note that for source i $\mathbf{w}^T\mathbf{H}_i\mathbf{m}_i = 0$ because $\mathbf{w}^T\mathbf{H}_i = \mathbf{0}$. Hence, the only result of the filter will be a zero signal, which can only be obtained by a filter equal to the null vector. That is because we have cancelled out all other signals. This happens only when the beamformer has been constrained to be at the location of our desired source. For any other location, the beamformer will try to do its best according to the conditions imposed by the constraint. So, we steer our beamformer constraint over a number of candidate source locations and at some

point it fails to give any output (i.e. it is equal to the null vector). That location will be the location of the desired signal. In other words, the optimum point of the process is where the algorithm fails to find a solution.

We will now show that for the correct reference and location the filter $\mathbf{w}$ is forced to zero. The constrained problem can be posed as:

$$\min_{\mathbf{w}} \quad f_d(\mathbf{w}) \quad \text{subject } to \quad f_c(\mathbf{w}) = \mathbf{0} \tag{6.3.6}$$

This constrained problem can be converted to an unconstrained optimisation procedure by using Lagrange multipliers. Consequently, Equation (6.3.6) is converted to (in the following equations, index i has been dropped for convenience):

$$F(\mathbf{w}) = f_d(\mathbf{w}) + f_c(\mathbf{w})\mathbf{q} = ||\mathbf{r} - \mathbf{w}^T\mathbf{X}||_2^2 + \mathbf{w}^T\mathbf{H}(p)\mathbf{q} \tag{6.3.7}$$

where $\mathbf{q}$ is a 3×1 vector of Lagrange multipliers. The derivative of $F(\mathbf{w})$ w.r.t. $\mathbf{w}^T$ is:

$$\frac{\partial F(\mathbf{w})}{\partial \mathbf{w}^T} = \frac{\partial}{\partial \mathbf{w}^T}\{\mathbf{r}\mathbf{r}^T - 2\mathbf{r}\mathbf{X}^T\mathbf{w} + \mathbf{w}^T\mathbf{X}\mathbf{X}^T\mathbf{w} + \mathbf{w}^T\mathbf{H}(p)\mathbf{q}\} \tag{6.3.8}$$

where i has been discarded for simplicity. This becomes:

$$\frac{\partial F(\mathbf{w})}{\partial \mathbf{w}^T} = -2\mathbf{r}\mathbf{X}^T + 2\mathbf{w}^T\mathbf{X}\mathbf{X}^T + \mathbf{q}^T\mathbf{H}(p)^T \tag{6.3.9}$$

To obtain the minimum we set Equation (6.3.9) to zero and obtain:

$$\mathbf{w}^T = \frac{1}{2}\left(2\mathbf{r}\mathbf{X}^T - \mathbf{q}^T\mathbf{H}(p)^T\right)\mathbf{C}_x^{-1} \tag{6.3.10}$$

where $\mathbf{C}_x = \mathbf{X}\mathbf{X}^T$ is the covariance matrix of $\mathbf{X}$. If we substitute Equation (6.3.10) into Equation (6.3.5) we obtain:

$$\mathbf{w}^T\mathbf{H}(p) = \frac{1}{2}\left(2\mathbf{r}\mathbf{X}^T - \mathbf{q}^T\mathbf{H}(p)^T\right)\mathbf{C}_x^{-1}\mathbf{H}(p) = 0 \tag{6.3.11}$$

which will give us the Lagrange multipliers $\mathbf{q}$:

$$\mathbf{q}^T = 2\mathbf{r}\mathbf{X}^T\mathbf{C}_x^{-1}\mathbf{H}(p)(\mathbf{H}(p)^T\mathbf{C}_x^{-1}\mathbf{H}(p))^{-1} \tag{6.3.12}$$

Now, we substitute $\mathbf{q}^T$ into Equation (6.3.10) and we obtain the full expression for the filter:

$$\mathbf{w}^T = \left(\mathbf{r}\mathbf{X}^T - \mathbf{r}\mathbf{X}^T\mathbf{C}_x^{-1}\mathbf{H}(p)(\mathbf{H}(p)^T\mathbf{C}_x^{-1}\mathbf{H}(p))^{-1}\mathbf{H}(p)^T\right)\mathbf{C}_x^{-1} \tag{6.3.13}$$

which splits into two parts; the first part is the solution to the primary cost function $f_d(\mathbf{w})$ and the second part is due to $f_c(\mathbf{w})$:

$$\mathbf{w}^T = \mathbf{w}_{opt}^T - \mathbf{r}\mathbf{X}^T\mathbf{C}_x^{-1}\mathbf{H}(p)(\mathbf{H}(p)^T\mathbf{C}_x^{-1}\mathbf{H}(p))^{-1}\mathbf{H}(p)^T\mathbf{C}_x^{-1} \tag{6.3.14}$$

We now proceed to show that if $\mathbf{r}$ corresponds to a source $\mathbf{s}_i$, which is uncorrelated with the other sources, and the forward matrix $\mathbf{H}_i$ is included in the grid search, then $\mathbf{w}$ will be forced to zero. This happens only for the conditions mentioned above. Consider the product $\mathbf{r}\mathbf{X}^T$

(ignore the noise for the moment) which is:

$$\begin{aligned}\mathbf{r}\mathbf{X}^T =& \mathbf{r}\Big(\sum_{j=1}^{m}\mathbf{H}_j\mathbf{m}_j\mathbf{s}_j\Big)^T \\ =& \sum_{j=1}^{m}\mathbf{r}\mathbf{s}_j^T\mathbf{m}_j^T\mathbf{H}_j^T\end{aligned} \tag{6.3.15}$$

So if the sources are uncorrelated then we obtain:

$$\mathbf{r}\mathbf{X}^T = \mathbf{r}\mathbf{s}_i^T\mathbf{m}_i^T\mathbf{H}_i^T \tag{6.3.16}$$

By substituting that into Equation (6.3.14) we achieve:

$$\mathbf{w}^T = \mathbf{w}_{opt}^T - \mathbf{r}\mathbf{s}_i^T\mathbf{m}_i^T\mathbf{H}_i^T\mathbf{C}_x^{-1}\mathbf{H}(p)(\mathbf{H}(p)^T\mathbf{C}_x^{-1}\mathbf{H}(p))^{-1}\mathbf{H}(p)^T\mathbf{C}_x^{-1} \tag{6.3.17}$$

Now, at some point during the grid search the matrix $\mathbf{H}(p)$ will take the value of $\mathbf{H}_i$ which will give through Equation (6.3.17), the following:

$$\begin{aligned}\mathbf{w}^T =& \mathbf{w}_{opt}^T - \mathbf{r}\mathbf{s}_i^T\mathbf{m}_i^T\mathbf{H}_i^T\mathbf{C}_x^{-1} \\ =& \mathbf{w}_{opt}^T - \mathbf{r}\mathbf{X}^T\mathbf{C}_x^{-1} \\ =& \mathbf{w}_{opt}^T - \mathbf{w}_{opt}^T = \mathbf{0}\end{aligned} \tag{6.3.18}$$

So, $\mathbf{w} = \mathbf{0}$ only for the location corresponding to the desired signal. A measure of how close $\mathbf{w}$ is to the null vector will point to the correct location. The procedure is to calculate Equation (6.3.14) for all locations obtained from a forward model and choose the solution with $\mathbf{w}$ closest to the null vector. The closeness measure we use is the norm of $\mathbf{w}$ since it reflects the distance of $\mathbf{w}$ to the origin, which is the null vector.

6.3.1 Correlated sources

In this section we explore the cases where the desired source is correlated with some other sources. For example, the P3a and P3b usually overlap temporally. In such a case the $\mathbf{w}$ vector will not become zero at the correct location. It will be biased by the cross-correlated terms as shown in the summation of Equation (6.3.15). After some algebra, we get for the filter $\mathbf{w}$ (at the location i of our desired source):

$$\mathbf{w}^T = -\mathbf{r} \sum_{j=1\, j\neq i}^{m} \mathbf{s}_j^T \mathbf{m}_j^T \mathbf{H}_j^T \mathbf{C}_x^{-1} \mathbf{H}_i (\mathbf{H}_i^T \mathbf{C}_x^{-1} \mathbf{H}_i)^{-1} \mathbf{H}_i^T \mathbf{C}_x^{-1} \qquad (6.3.19)$$

So, at the correct location of source $\mathbf{s}_i$ the norm of $\mathbf{w}$ will depend on the cross-correlation between the desired source $\mathbf{r} = \mathbf{s}_i$ and the rest of the sources as well as the rest of the product in Equation (6.3.19). In such a case it is possible for the norm of $\mathbf{w}$ to be minimised in a wrong location (however not equal to the null vector). Note that the effect of the correlation between sources will influence the beamformer for every location. If we denote the beamformer as $\mathbf{w}_C$ when there are correlated sources and $\mathbf{w}$ for the uncorrelated case then we can write:

$$\mathbf{w}_C^T = \mathbf{w}^T - \mathbf{r}\bar{\mathbf{X}}^T \mathbf{C}_x^{-1} \mathbf{H}(p) (\mathbf{H}(p)^T \mathbf{C}_x^{-1} \mathbf{H}(p))^{-1} \mathbf{H}(p)^T \mathbf{C}_x^{-1} \qquad (6.3.20)$$

where $\bar{\mathbf{X}}$ is the matrix containing all the sources that are correlated with our desired source:

$$\bar{\mathbf{X}} = \sum_{j=1}^{c} \mathbf{H}_j \mathbf{m}_j \bar{\mathbf{s}}_j \qquad (6.3.21)$$

where in this case $\bar{\mathbf{s}}_j$ for $j = 1...c$ is the subset of c sources that are correlated with the desired source. The effect of source correlation is investigated in section 6.4.

6.3.2 Noise effect

If we include the noise effect in Equation (6.3.15) it changes to:

$$\begin{aligned} \mathbf{r}\mathbf{X}^T =& \mathbf{r}\Big(\sum_{j=1}^{m} \mathbf{H}_j \mathbf{m}_j \mathbf{s}_j + \mathbf{N}\Big)^T \\ =& \sum_{j=1}^{m} \mathbf{r}\mathbf{s}_j^T \mathbf{m}_j^T \mathbf{H}_j^T + \mathbf{r}\mathbf{N}^T \end{aligned} \tag{6.3.22}$$

If we denote the beamformer as $\mathbf{w}_N$ when the noise effect is considered and $\mathbf{w}$ the beamformer for the noiseless case then we can write:

$$\mathbf{w}_N^T = \mathbf{w}^T - \mathbf{r}\mathbf{N}^T \mathbf{C}_x^{-1} \mathbf{H}(p)(\mathbf{H}(p)^T \mathbf{C}_x^{-1} \mathbf{H}(p))^{-1} \mathbf{H}(p)^T \mathbf{C}_x^{-1} \tag{6.3.23}$$

The last term in the right hand side of Equation (6.3.23) is influenced by the correlation between the desired source and the noise sequence. The effect depends on the statistics of the noise, the number of samples, and the rest of the product in Equation (6.6.2). The effect of the noise is investigated in section 6.4 as well.

6.4 Experimental Results

In this section we apply the algorithm to a simulated EEG signal containing a number of ERP components and we localise the P3a and P3b from real EEG data. We investigate a joint ICA-SNF procedure in which the desired signal is obtained directly from ICA.

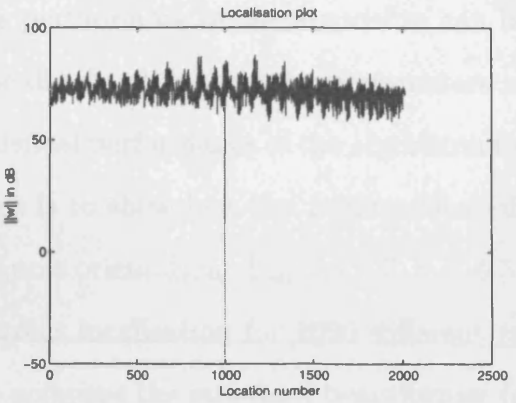

Figure 6.2. Localisation plot for one source uncorrelated with other sources in a noise free environment. The location number refers to a geometrical location in a three-dimensional grid within the brain.

6.4.1 Simulated EEG

Here, we investigate the ability of the algorithm for correct localisation of the sources in various scenarios. A forward model is obtained using the **BrainStorm** software [72]. We use a 3 layer spherical head model with conductivities of $0.33\mu S/cm$, $0.0042\mu S/cm$, $0.33\mu S/cm$, for scalp, skull and brain respectively. We create thirty-two Gaussian pulses in thirty-two different locations with random orientations, peaking at different latencies and using thirty electrodes. This is to have an underdetermined system. We consider several different cases in order to evaluate the effect of noise and correlation between the sources. There are 2700 voxels. The source we are looking for is originally placed at location numbered 1000. For the simple case of no noise and uncorrelated[1] sources we obtain accurate localisation of the source as depicted in Figure 6.2.

So far, we have used a single orientation set-up for the mixing of

[1]we define correlation between vectors $\mathbf{x}$ and $\mathbf{y}$ as $r_{\mathbf{x},\mathbf{y}} = \frac{\mathbf{x}^T\mathbf{y}}{||\mathbf{x}||_2||\mathbf{y}||_2}$ where $||.||_2$ denotes Euclidean norm.

the sources. The performance of the algorithm can be affected by the orientation of the dipole while the other parameters are fixed. Now we consider the statistical performance of the algorithm for various random orientations. This is to show how the performance of the algorithm is affected by the dipole orientation. Figures 6.3, 6.4, 6.5 and 6.6 show the percentage of correct localisation for 1000 different randomly oriented dipoles. We also compare the standard beamformer (dashed line) with our algorithm (bold). Figure 6.3 shows the performance of the algorithm as the SNR increases for a constant correlation (10%) between two sources while Figure 6.4 shows the same for 30% correlation. In Figures 6.5 and 6.6 we show the performance of the algorithm with respect to the correlation between the sources for constant SNR values (15dB and 5dB). Again, the standard LCMV algorithm is compared with our algorithm. Note that for all the figures we have the same number of sources (32) using the same setup.

It can be seen from the plots that the algorithm is very robust in terms of noise but the performance degrades as the level of correlation increases. Compared to the LCMV method, our algorithm demonstrates much better performance for high noise levels. However, LCMV has a better performance for high correlation values.

6.4.2 Real EEG

In the section we used the same recording as in Section (4.1.2). In Figures 6.7 and 6.8 we see the locations of the P3a and P3b sources for a number of normal individuals and schizophrenic patients. It can be seen that the locations of the P3a and P3b for the schizophrenic patients are less distinct than the locations for the healthy individuals.

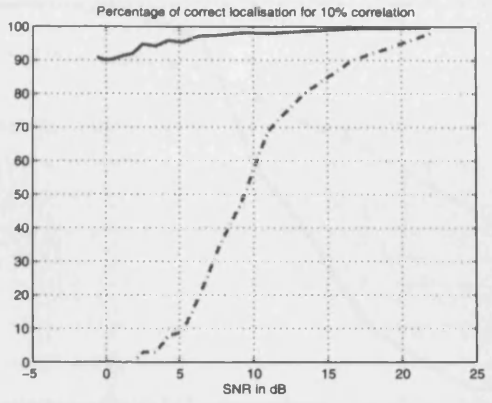

Figure 6.3. Percentage of successful localisation for various SNRs for our algorithm (bold) and the LCMV (dashed). The purpose is to evaluate the performance of the algorithm for different orientations of the sources. We used the same noise sequence for 1000 different orientations and various SNR values. Here, the correlation is 10%.

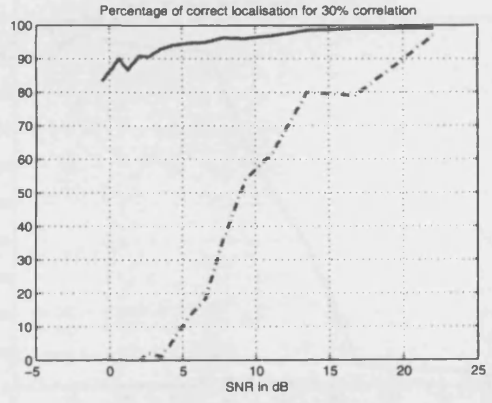

Figure 6.4. Percentage of successful localisation for various SNRs for our algorithm (bold) and the LCMV (dashed). The purpose is to evaluate the performance of the algorithm for different orientations of the sources. We used the same noise sequence for 1000 different orientations and various SNR values. Here, the correlation is 30%.

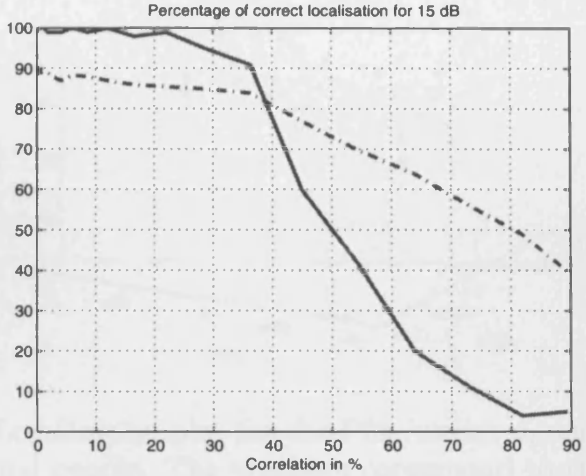

Figure 6.5. Percentage of successful localisation for various correlation values for our algorithm (bold) and the LCMV (dashed). The purpose is to evaluate the performance of the algorithm for different orientations of the sources. We used the same noise sequence for 1000 different orientations and various correlation values. Here, the SNR is 15dB.

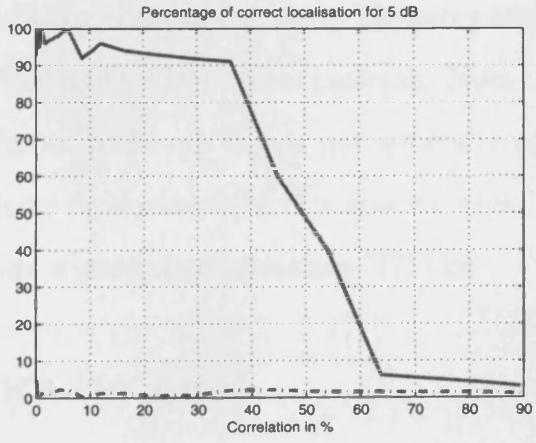

Figure 6.6. Percentage of successful localisation for various correlation values for our algorithm (bold) and the LCMV (dashed). The purpose is to evaluate the performance of the algorithm for different orientations of the sources. We used the same noise sequence for 1000 different orientations and various correlation values. Here, the SNR is 5dB.

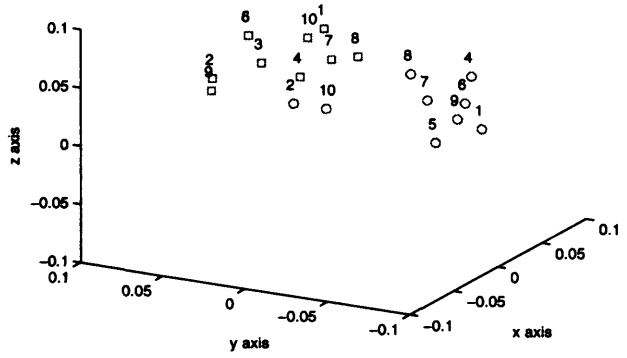

Figure 6.7. Localisation plot for the P3a, circles o, and P3b, squares □, for 10 normal people. The numbers correspond to the control subject's number (e.g. o^1 shows the location of the P3a for control subject number 1). The three axis refer to the geometrical coordinates in meters. The y axis determines front-back of the head, x axis is left-right and z is the vertical position. Units are in meters.

The average distance for the normal subjects was 12cm while for the patients it was 8cm. The standard deviation of the distances was 0.03 for the normal subjects and 0.036 for the schizophrenic patients. In Figure 6.9 we see a bar chart of the mean distances and standard errors between the normal and schizophrenic patients. Note that by source we mean a single dipole; although this is not generally correct. However, if the brain activity is concentrated in a specific area, the single dipole model can provide a good approximation [17] [38].

6.5 A hybrid ICA-SNF system

Until now we had knowledge of the desired source beforehand. In this section we investigate the case where we obtain the desired source using the Infomax algorithm. We perform experiments on simulated data and use the same setup as in last section.

We use the original unconstrained Infomax algorithm as described

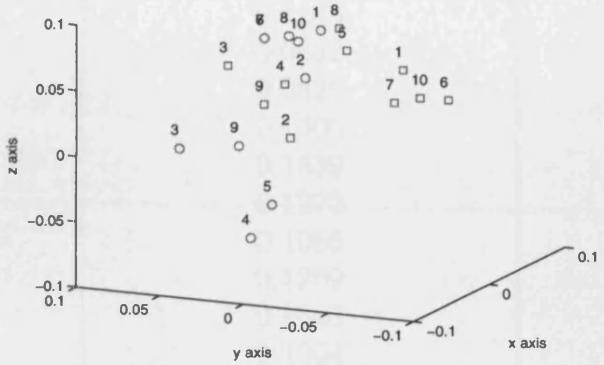

Figure 6.8. Localisation plot for the P3a, circles ○, and P3b, squares □, for 10 schizophrenic patients. The numbers correspond to the schizophrenic patient's number (e.g. $\circ^1$ shows the location of the P3a for schizophrenic patient number 1). The three axis refer to the geometrical coordinates in meters. The y axis determines front-back of the head, x axis is left-right and z is the vertical position. Units are in meters.

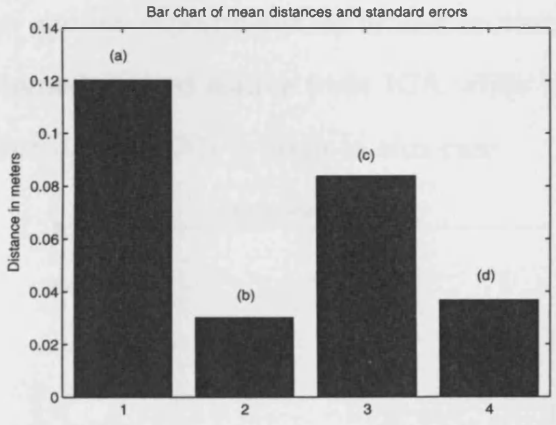

Figure 6.9. Bar chart of the distances of normal subjects and schizophrenic patients. (a) shows the mean distances between P3a and P3b for the normal subjects, (b) shows the standard deviation of the distances from the mean, (c) shows the mean distances between P3a and P3b for the schizophrenic patients and (d) shows the standard deviation of the distances from the mean

Table 6.1. Distance in meters between the P3a and P3b locations for control subjects and schizophrenic patients.

Subject Number	Distance(meters) for Control	Distance(meters) for Patient
1	0.1145	0.1063
2	0.0632	0.0447
3	0.0825	0.0663
4	0.1300	0.1217
5	0.1539	0.1269
6	0.1273	0.1095
7	0.1086	0.0985
8	0.1269	0.0316
9	0.1640	0.0332
10	0.1304	0.1005

in section 3.5. After the sources are estimated we use a very simple method to choose the one closest to our desired latency. We compute the correlation between all of our desired sources and a rough template that peaks at the desired latency. Then we choose from the estimated sources the one with the highest correlation.

Finally we use that source as our desired source for the SNF algorithm. We run similar experiments as in last section. In Figure 6.10 we see the obtained desired source from ICA while in Figure 6.11 we see the true source. The SNR is 20db in this case.

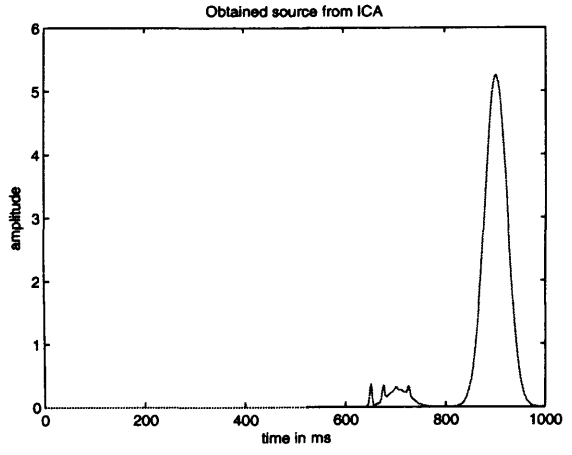

Figure 6.10. The obtained desired source from ICA(Infomax).

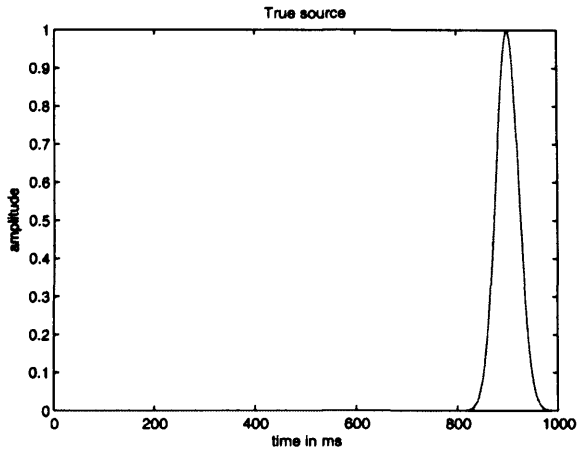

Figure 6.11. The estimation of the true source using ICA.

We run the simulation 1000 times and the mean Euclidean distance between the true source and the estimated one from ICA is 0.7 while a typical Euclidean distance between two sources that are near in latency is 1.3. Compared to the case where we exactly know the desired source the algorithm exhibited a 30% drop in localisation success for 20dB SNR while for low noise this figure increased to 45%.

6.6 Correlated Sources

If there are two or more sources correlated with our desired signal then the performance of the algorithm degrades as observed in the previous sections. In Equations (6.6.1) and (6.6.2) we saw the relation between the obtained $\mathbf{w}$ for the correlated and uncorrelated cases.

$$\mathbf{w}_C^T = \mathbf{w}^T - \mathbf{r}\bar{\mathbf{X}}^T\mathbf{C}_x^{-1}\mathbf{H}(p)(\mathbf{H}(p)^T\mathbf{C}_x^{-1}\mathbf{H}(p))^{-1}\mathbf{H}(p)^T\mathbf{C}_x^{-1} \tag{6.6.1}$$

$$\bar{\mathbf{X}} = \sum_{j=1}^{c} \mathbf{H}_j\mathbf{m}_j\bar{\mathbf{s}}_j \tag{6.6.2}$$

Lets assume that there are only two sources that are correlated and we know their waveforms. Then Equation (6.6.1) breaks down to:

$$\mathbf{w}_C^T = \mathbf{w}^T - \mathbf{r}\bar{\mathbf{s}}_j^T\mathbf{m}_j^T\mathbf{H}_j^T\mathbf{C}_x^{-1}\mathbf{H}(p)(\mathbf{H}(p)^T\mathbf{C}_x^{-1}\mathbf{H}(p))^{-1}\mathbf{H}(p)^T\mathbf{C}_x^{-1} \tag{6.6.3}$$

Hence, in order to overcome the problem with correlation we need an estimation of $\mathbf{m}_j$ and $\mathbf{H}_j$. The procedure for the estimation is as such:

- Find $\mathbf{w}_C^T = \mathbf{w}_{opt}^T - \mathbf{r}\mathbf{X}^T\mathbf{C}_x^{-1}\mathbf{H}(p)(\mathbf{H}(p)^T\mathbf{C}_x^{-1}\mathbf{H}(p))^{-1}\mathbf{H}(p)^T\mathbf{C}_x^{-1}$ as in Equation (6.3.14) for every location p,
- For every such $\mathbf{w}_C$ we include another loop that runs through all locations, use index q here, and find $\mathbf{w}_q^T$ of Equation (6.6.3),
- The only unknown is $\mathbf{m}_q$ which is found by estimating the proper $\mathbf{m}_q$ that minimises $\mathbf{w}_q$,
- Among all $\mathbf{w}_q$ which are obtained through a minimisation process,

we find the one with the smallest norm.

In other words we minimise Equation (6.6.3) with respect to the location of the desired source, the location of the correlated source and the orientation of the correlated source.

6.7 Conclusions

In this chapter we developed an algorithm for localisation of the P300 subcomponents. We modified the LS approach by using a desired signal and a spatial notch filter. The desired signal is designed based on prior knowledge of the shape of the P300 subcomponents. The algorithm is in fact a special kind of beamformer and is a constrained version of the original LS solution. It points to the correct location when we have a suitable model of the actual sources and the sources are uncorrelated. As seen in section 6.4, the correlation between the sources and increasing the noise level can degrade the performance of the algorithm but even in these cases the algorithm achieves high localisation accuracy.

We applied the algorithm to real EEG data and compared the locations of the P300 subcomponents between normal subjects and schizophrenic patients. We achieved that the locations of the P300 subcomponents are mixed for the schizophrenic patients, whereas the locations are more distinct for a healthy subject. Moreover, the comparison between our proposed beamformer and the LCMV method showed an improved quality of the results when using our method.

We also combined ICA and SNF in which the desired signal is obtained directly from an ICA algorithm. The results were encouraging with small errors even for high noise level.

Chapter 7

SUMMARY, CONCLUSIONS AND FUTURE WORK

The field of EEG signal processing and more specifically the extraction and localisation of the P300 subcomponents have attracted great attention in the last two decades. Advances in signal processing and the application of well known methods in the field of ERP has enhanced the analysis of brain signals. Moreover, the creation of software packages such as Brainstorm [72] and EEGLAB [91] has made EEG analysis easier even for non-experts. As a result, the possibilities for research are endless. Although the signals of interest in our thesis, namely the P3a and P3b subcomponents, have very valuable diagnostic information, they have not been effectively explored by experts in ERP analysis. Also, methods specifically designed for their automatic recognition have not been developed.

ERP signal processing suffers from poor spatial resolution, low SNR and regarding our research, the temporal overlap of the P3a and P3b subcomponents. In this thesis we developed three methods that dealt with those problems. All of the developed methods used prior knowledge of the shapes of the P300 subcomponents. However, the implementation of the prior knowledge has been substantially different between

the methods. There has been a gap in existing methods regarding the use of prior information when dealing with ERP signal processing. In this thesis we aimed to close that gap.

In section 4.1 we enhance the conventional Infomax algorithm by adding a constraint function to the original Infomax cost function. The aim of the constraint function is to aid Infomax in the extraction of the desired P300 subcomponents (see section 4.1). The resulting algorithm's outputs resemble the true sources more closely than those of the unconstrained Infomax. Also, there are occasions that unconstrained Infomax fails to produce any of the desired sources and incorporating the constraint alleviates that problem. For an investigation of the efficacy of the algorithm on simulated data see section 4.1.2 and on real EEG data see section 4.1.3. Another inherent problem with placing the constraint function is the reliable use of the prior information that we have. In section 4.1.1 we describe how we obtained that information. The BSS algorithms provide an estimate of the mixing matrix of the propagation medium from the sources to the electrodes. This in turn can provide an estimate of the location of the sources if there is proper knowledge of the propagation process. In section 4.1.4 we developed a simple method that can localise the extracted sources from the EEG. The method, albeit not entirely exact, has the ability to distinguish between sources from different locations.

In section 4.2 we apply a constraint function this time to the BSE method. BSE methods extract sources one at a time and this is desirable for a number of reasons. An algorithm can be designed for specific situations; the algorithm is much simpler and faster to converge and it is used to concentrate on the signals of interest. In section 4.2 we

apply the same constraint function to the basic BSE cost function. The algorithm was applied on real EEG data and the main advantage over unconstrained BSE is that it produces an output resembling a P300 subcomponent every time.

In chapter 5 we developed a novel method that is especially designed for the single-trial estimation of the P300 subcomponents. Unlike the reference signals used in Chapter 4 we used generic templates that resemble the P3a and P3b signals. The reason for that is that we wanted the process to be completely blind in the sense that we do not use any information from the EEG. In other words we desired to create an algorithm that can be used at any occasion without any information from the EEG. In section 5.1 we introduced an algorithm for the extraction of the sources together with a localisation algorithm that is an enhanced version of the localisation algorithm of section 4.1.4. In section 5.2 an investigation was performed on the validity of the method regarding simulated data and the performance was satisfactory and the resulting latency estimation was succesfull even in the presence of noise and in the case where the sources are correlated. Section 5.3 demonstrates the application of the algorithm on real EEG data. We measured various features of the extracted sources and we found variability between the P300 subcomponents of control subjects and schizophrenic patients. Results on the localisation algorithm are shown in section 5.4 where we investigated the efficacy of the algorithm on single trial data.

In Chapter 6 a novel method (SNF) that localises brain sources when we have a good knowledge of their shape was presented. It is based on the principles of LS and spatial filtering. The results were very satisfactory and a localisation method that uses desired signals does

not exist in the literature. In section 6.3 we described the SNF method from basic principles and gave a proof of its validity. In sections 6.3.1 and 6.3.2 we investigated the performance for the cases when there is noise and when some sources are correlated. Section 6.4 proceeded with experiments that demonstrated the localisation success of the algorithm in various scenarios; high-low SNR, high-low correlation. SNF exhibits very good performance even in the presence of noise and correlated sources. It was empirically verified that the localisation success is higher than that of the LCMV method (see section 3.4). A combined ICA-SNF method was shown in section 6.5 where the extracted sources from ICA as desired signals are used in SNF. The results were good but there is room for further investigation. Finally, in section 6.6 we described a solution to the problem of correlated sources.

7.1 Future work and development

As mentioned before the possibilities for ERP research are endless within the field of signal processing. Similarly, the methods shown in this thesis have great potential for further research and development for a variety of applications.

The CBSS algorithm of Chapter 4 can be modified in many ways to better enhance its performance. Firstly, the pdfs of the sources in an ICA framework can be optimised to better fit the pdfs of the true sources (see section 3.5). Also, other constraint function can be applied such as temporal smoothing or spatial decorrelation. Regarding CBSE (section 4.2), other BSE cost functions can be used such as FastIca [42].

The algorithm of Chapter 5 can be combined with ICA to extract the sources instead of the LS algorithm. In other words, Gaussian

spikes should be still used as templates but an ICA algorithm should be responsible for the extraction. Moreover, an in depth study into a classification algorithm for the various features produced from the algorithm would be highly desirable. This can lead to an automatic on-line classification algorithm for the distinction between schizophrenic patients and healthy subjects. Furthermore, the classification algorithm can be extended to distinguishing between types of schizophrenia and the sex of the patient.

The novel method in Chapter 6 deserves further study and exploration. It has to be mentioned that the algorithm is not limited in the EEG field. It can be used in any scenario where we have some knowledge of the propagation process between the sources and sensors (in our case electrodes). For example, the location of a sound source could be estimated provided we know the locations of the microphones and the approximate waveform of the sound source. More importantly, the algorithm can be modified to work in the frequency domain; in effect there is no difference since we can treat the spectrum of the signal as any other signal. In section 6.3.1 we observe that knowledge of the noise would lead to improvement in the localisation success rate. A common method used in other signal processing methods [38] is to estimate the noise statistics prior to the onset of the EEG signal and use this estimate as the noise within the current EEG dataset range. Finally, the ICA-SNF method should be applied to real EEG data for the automatic classification of the P300 subcomponents.

BIBLIOGRAPHY

[1] D. Salisbury, M.E Shenton, and R.W. McCarley, "P300 topography differs in schizophrenia and manic psychosis," *Biological Psychiatry*, vol. 45, pp. 98–106, January 1999.

[2] B. E. Swartz and E. S. Goldensohn, "Timeline of the history of EEG and associated fields," *Electroencephalography and Clinical Neurophysiology*, vol. 106, no. 2, pp. 173–176, 1998.

[3] I. B. Levitan and L. K. Kaczmarek, *The Neuron.* Oxford University Press, 1997.

[4] N. A. Carlson, *Foundations of Physiological Psychology.* Needham Heights, Massachusetts: Simon and Schuster, 1992.

[5] A. Kok, J. Ramautar, M. De Ruiter, G. Band and K. Ridderinkhof, "ERP components associated with successful and unsuccessful stopping in a stop-signal task," *Psychophysiology*, vol. 41, no. 1, pp. 9–20, 2004.

[6] J. Polich, "Clinical application of the P300 event-related brain potential," *American Journal of Physical Medicine & Rehabilitation*, vol. 15, no. 1, pp. 133–161, 2004.

[7] J. Dien, K. Spencer, and E. Donchin, "Localization of the event-related potential novelty response as defined by principal components

analysis," *Cognitive Brain Research, Brain Research*, vol. 17, no. 3, pp. 637–650, 2003.

[8] D. Friedman, Y. Cycowicz, and H. Gaeta, "The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty," *Neuroscience & Biobehavioral Reviews*, vol. 25, no. 4, pp. 355–373, 2001.

[9] T. Frodl-Bauch, R. Bottlender, and U. Hegerl, "Neurochemical substrates and Neuro anatomical Generators of the Event-Related P300," *Neuropsychobiology*, vol. 40, no. 2, pp. 86–94, 1999.

[10] M. Comerchero and J. Polich, "P3a and P3b from typical auditory and visual stimuli," *Clinical Neuropsychobiology*, vol. 110, no. 1, pp. 24–30, 1999.

[11] E. Niedermeyer and F. L. da Silva, *Electroencephalography, Basic Problems, Clinical Applications, and Related Fields.* LW&W, 1999.

[12] B. Turetsky, E. Colbath, and R. Gur, "P300 subcomponent abnormalities in schizophrenia: I. Physiological evidence for gender and subtype specific differences in regional pathology," *Biological Psychiatry*, vol. 43, no. 2, pp. 84–96, 1998.

[13] B. Turetsky, T. Cannon, and R. Gur, "P300 subcomponent abnormalities in schizophrenia: III. Deficits In unaffected siblings of schizophrenic probands," *Biological Psychiatry*, vol. 47, no. 5, pp. 380–390, 2000.

[14] Y. Jeon and J. Polich, "Meta-analysis of P300 and schizophrenia: patients, paradigms, and practical implications," *Psychophysiology*, vol. 40, no. 5, pp. 684–701, 2004.

[15] J. C. Mosher, P. S. Lewis, and R. M. Leahy, "Multiple Dipole Modelling and Localisation from Spatio-Temporal MEG Data," *IEEE Transactions on Biomedical Engineering*, vol. 39, pp. 541–557, June 1992.

[16] J. C. Mosher and R. M. Leahy, "EEG and MEG Source Localisation using Recursively Applied (RAP) Music," *Proc. Thirtieth Annual Asilomar Conference on Signals, Systems and Computers*, November 3-6 1996.

[17] B. He, T. Musha, Y. Okamoto, S. Homma, Y. Nakajima, and T. Sato, "Electric Dipole Tracing in the Brain by Means of the Boundary Element Method and Its Accuracy," *IEEE Transactions on Biomedical Engineering*, vol. 34, pp. 406–414, June 1987.

[18] M. Scherg and D. Von Cramon, "Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model," *Electroencephalography Clinical Neurophysiology*, vol. 62, pp. 32–44, January 1985.

[19] I. S. Yetik, A. Nehorai, J.D. Lewine, and C.H. Muravchik, "Distinguishing between moving and stationary sources using EEG/MEG measurements with an application to epilepsy," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 839–851, May 2005.

[20] I. S. Yetik, A. Nehorai, C. H. Muravchik, and J. Haueisen, "Line-source modeling and estimation with magnetoencephalography," *IEEE International Symposium on Biomedical Imaging: Nano to Macro*, vol. 2, pp. 1339 – 1342, April 2004.

[21] R. D. Pascual-Marqui, "Review of Methods for Solving the EEG

Inverse Problem," *International Journal of Bioelectromagnetism*, vol. 1, pp. 75–86, January 1999.

[22] R. D. Pascual Marqui, C. M. Michel and D. Lehmann, "Low Resolution Elecromagnetic Tomography: A New method for Localising Electrical Activity in the Brain," *International Journal of Psychophysiology*, vol. 18, pp. 49–65, 1994.

[23] H. Liu, P. H. Schimpf, G. Dong, X. Gao, F. Yang and S. Gao, "Standardized Shrinking LORETA-FOCUSS(SSLOFO): A New Algorithm for Spatio-Temporal EEG Source Reconstruction," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 1681–1691, October 2005.

[24] B. D. Rao, "Analysis and Extensions of the Focuss algorithn," *Signals, Systems and Computers,*, vol. 2, pp. 1218–1223, November 1996.

[25] I. F. Gorodnitsky, J. S. George, and B. D. Rao, "Neuromagnetic source imaging with FOCUSS: A recursive weighted minimum norm algorithm," *Electrencephalography Clinical Neurophysiology*, vol. 95, pp. 231–251, October 1995.

[26] H. Liu, X. Gao, P. Schimpf, F. Yang and S. Gao, "A recursive algorithm for the three-dimensional imaging of brain electric activity: shrinking LORETA-FOCUSS," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 1794–1802, October 2004.

[27] R. D. Pascual-Marqui, "Standardized low resolution brain electromagnetic tomography (sLORETA): technical details," *Methods and Findings in Experimental and Clinical Pharmacology*, vol. 24, pp. 5–12, 2002.

[28] U. Schmitt, A. K. Louis, F. Darvas, H. Buchner and M. Fuchs, "Numerical Aspects of Spatio-Temporal Current Density Reconstruction from EEG-/MEG-Data," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 314–324, April 2001.

[29] Y. Serinagaoglu, D. H. Brooks and R. S. Macleod, "Baeysian Solutions and Performance Analysis in Bioelectric Inverse Problems," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 1009–1020, June 2005.

[30] S. Baillet and Line Garnero, "A Bayesian Approach to Introducing Anatomo-Functional Priors in the EEG/MEG Inverse Problem," *IEEE Transactions on Biomedical Engineering*, vol. 44, pp. 374–385, May 1997.

[31] J. Daunizeau, J. Mattout, D. Clonda, B. Goulard, H. Benali, and J. -M. Lina, "Bayesian Spatio-Temporal Approach for EEG Source Reconstruction: Conciliating ECD and Distributed Models," *IEEE Transactions on Biomedical Engineering*, vol. 53, pp. 503–516, March 2006.

[32] P. O. Ranta-Aho, "Single-Trial Estimation of Multichannel Evoked-Potential Measurements," *IEEE Transactions on Biomedical Engineering*, vol. 50, pp. 189–196, February 2003.

[33] J. Sarvas, "Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem," *Physiology Medical Biology*, vol. 32, pp. 11–22, January 1987.

[34] N. v. Ellenrieder, C. H. Muravchik and A. Nehorai, "A Meshless Method for Solving the EEG Forward Problem," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 249–257, February 2005.

[35] M. Fuchs, M. Wagner and J. Kastner, "Boundary element method volume conductor models for EEG source reconstruction," *Clinical Neurphysiology*, vol. 112, no. 8, pp. 1400–1407, 2001.

[36] D. Ent, J. C. de Munck, A. L. Kaas, "A Fast Method to Derive Realistic BEM Models for E/MEG Source Reconstruction," *IEEE Transactions on Biomedical Engineering*, vol. 48, pp. 1434–1443, December 2001.

[37] J. C. Mosher, R. M. Leahy and P. S. Lewis, "EEG and MEG: Forward Solutions for Inverse Methods," *IEEE Transactions on Biomedical Engineering*, vol. 46, pp. 245–259, March 1999.

[38] B. D. van Veen, W. van Drogelen, M. Yuchtman and A. Suzuki, "Localisation of Brain Electrical Activity via Linearly Constrained Minimum Variance Spatial Filtering," *IEEE Transactions on Biomedical Engineering*, vol. 44, pp. 867–880, September 1997.

[39] G. Van Hoey, R. Van de Walle, B. Vanrumste, M. D'Havse, I. Lemahieu and P. Boon, "Beamforming techniques applied in EEG source analysis," *Proceedings ProRISC99*, vol. 10, pp. 545–549, 1999.

[40] S. Lemm, B. Blankertz, G. Curio and K. -R. Muller, "Spatio-Spectral filters for Improving the Classification of Single Trial EEG," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 1541–1548, September 2005.

[41] H. Ramoser, J. Muller-Gerking and G. Pfutscheller, "Optimal Spatial Filtering of Single Trial EEG During Imagined Hand Movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, pp. 441–446, December 2000.

[42] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing, Learning Algorithms and Applications.* John Wiley, 2002.

[43] A. J. Bell and T. J. Sejnowski, "An information-maximisation approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1004–1034, 1995.

[44] J. F. Cardoso, "Blind Signal Separation: Statistical Principles," *Proceeding of the IEEE*, vol. 9, pp. 2009–2025, October 1998.

[45] P. Comon, "Independent component analysis, A new concept?," *Signal Processing*, vol. 36, pp. 287–314, April 1994.

[46] K. J. Pope, R. E. Bogner, "Blind Signal Separation I. Linear, Instantaneous Combinations," *Digital Signal Processing*, vol. 6, pp. 5–16, January 1996.

[47] A. Hyvarinen, E. Oja, "Independent component analysis: algorithms and applications," *Neural Networks*, vol. 13, no. 4-5, pp. 411–430, 2000.

[48] A. Cichocki, R. Unbehauen and E. Rummert, "Robust learning algorithm for blind separation of signals," *Electronics Letters*, vol. 30, pp. 1386–1387, August 1994.

[49] S. Makeig, A. J. Bell, T. Jung and T. J. Sejnowski, "Independent Component Analysis of Electroencephalographic data," *Advances in Neural Information Processing Systems*, vol. 8, no. 17, pp. 145–151, 1996.

[50] S. Makeig, T. Jung, A. J. Bell, D. Ghahremani and T. J. Sejnowski, "Blind separation of auditory event-related brain responses

into independent components," *Proceedings National Acadamy of Science*, vol. 94, pp. 10979–10984, September 1997.

[51] J. Anemuller, T. J. Sejnowski and S. Makeig, "Complex independent component analysis of frequency-domain electroencephalographic data," *Neural Networks, Special Issue*, vol. 16, pp. 1311–1323, November 2003.

[52] A. Ossadtchi, S. Baillet, J.C. Mosher, D. Thyerlei, W. Sutherling, and R.M.Leahy, "Automated Interictal Spike Detection and Source Localisation in MEG using ICA and Spatio-Temporal Clustering," *Clinical Neurophysiology*, vol. 115, pp. 508–522, October 2003.

[53] S. Baillet, L. Garnero, G. Marin, and J. -P. Hugonin, "Combined MEG and EEG Source Imaging by Minimization of Mutual Information," *IEEE Transactions on Biomedical Engineering*, vol. 46, pp. 522–534, May 1999.

[54] W. Lu, J. C. Rajapakse, "Approach and Applications of Constrained ICA," *IEEE Transactions on Biomedical Engineering*, vol. 16, pp. 203–212, January 2005.

[55] C. J. James and O. J. Gibson, "Temporally Constrained ICA: An Application to Artifact Rejection in Electromagnetic Brain Signal Analysis," *IEEE Transactions on Biomedical Engineering*, vol. 50, pp. 1108–1116, September 2003.

[56] L. Spyrou, S. Sanei, and A. Sumich, "Separation and Localization of Auditory And Visual P300 Sources In Schizophrenia Patients Via Constrained BSS," *Proceedings of IEEE, ICASSP 2005*, vol. 5, pp. 613–616, March 2005.

[57] S. Sanei, J. A. Chambers, M. A. Latif, and L. Shoker, "Localization of abnormal EEG sources using blind source separation partially constrained by the locations of known sources," *IEEE Signal Processing Letters*, vol. 13, pp. 117 – 120, March 2006.

[58] Z. Chen, A. Cichocki, and T.M. Rutkowski, "Constrained non-Negative Matrix Factorization Method for EEG Analysis in Early Detection of Alzheimer Disease," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 5, May 2006.

[59] A. Cichocki, R. Zdunek, and S. Amari, "New Algorithms for Non-Negative Matrix Factorization in Applications to Blind Source Separation," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 5, May 2006.

[60] Daniel D. Lee and H. Sebastian Seung, "Algorithms for Non-negative Matrix Factorization," *NIPS*, vol. 5, pp. 556–562, 2000.

[61] V. Samar, K. P. Swartz, and M. R. Raghuveer, "Multiresolution Analysis of Event-Related Potentials by Wavelet Decomposition," *Brain and Condition*, vol. 27, pp. 398–438, 1995.

[62] Z. Nenadic and J. W. Burdick, "Spike Detection Using the Continuous Wavelet Transform," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 74–87, January 2005.

[63] B. Causevic, R. E. Morley, M. V. Wickerhauser and A. E. Jacquin, "Fast Wavelet Estimation of Weak Biosignals," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 1021 – 1032, June 2005.

[64] E. L. Glassman, "A Wavelet-Like Filter Based on Neuron Action Potentials for Analysis of Human Scalp Electroencephalographs," *IEEE*

Transactions on Biomedical Engineering, vol. 52, pp. 1851 – 1862, November 2005.

[65] M. G. Jafari and J. A. Chambers, "Fetal Electrocardiogram Extraction by Sequential Source Separation in the Wavelet Domain," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 390– 400, March 2005.

[66] R. Kus, M. Kaminski, K. J. Blinowska, "Determination of EEG Activity Propagation: Pair-Wise versus multichannel Estimate," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 1501–1509, September 2004.

[67] R. Kus, M. Kaminski, K. J. Blinowska, "Model Selection in Spatio-Temporal Electromagnetic Source Analysis," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 414 – 420, March 2004.

[68] L. J. Waldorp, H. M. Huizenga and R. P. P. P. Grasman, "The Wald Test and Cramer-Rao Bound for Misspecified Models in Electromagnectic Source Analysis," *IEEE Transactions on Signal Processing*, vol. 53, pp. 3427–3435, September 2005.

[69] M. Hansson, T. Gansler, and G. Salomonsson, "Estimation of Single Event-Related Potentials Utilizing the Prony Method," *IEEE Transactions on Biomedical Engineering*, vol. 43, pp. 973 – 981, October 1996.

[70] V. Garoosi, B.H. Jansen, "Development and evaluation of the piecewise Prony method for evoked potential analysis," *IEEE Transactions on Biomedical Engineering*, vol. 47, pp. 1549 – 1554, December 2000.

[71] S. Aviyente, L.A.W. Brakel, R.K. Kushwaha, M. Snodgrass, H. Shevrin, W.J. Williams, "Characterization of event related potentials

using information theoretic distance measures," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 737– 743, May 2004.

[72] "BrainStorm, Matlab toolbox available [Online] @ http://neuroimage.usc.edu/brainstorm/,"

[73] L. Spyrou, S. Sanei, M. Jing, and A. Sumich, "Separation and localisation of P300 sources and the subcomponents using constrained blind source separation," *EURASIP Journal on Advances in Signal Processing*, vol. 1, pp. 89–98, January 2006.

[74] L. Spyrou and S. Sanei, "A Robust Constrained Method for the Extraction of P300 Subcomponents," *Proceedings of IEEE, ICASSP 2006*, vol. 2, pp. II–II, May 2006.

[75] S. Sanei and L. Spyrou, "Localization of P300 Sources in Schizophrenia Patients Using Constrained BSS," *Proceedings of the 5th Conference in Independent Component Analysis (ICA2004)*, vol. 3195, pp. 177–184, September 2004.

[76] D. G. Luenberger, *Optimization by Vector Space Methods.* John Wiley, 1997.

[77] W. Lu and J.C. Rajapakse, "Constrained ICA," *Advances in Neural Information Processing Systems*, vol. 13, pp. 570–576, 2000.

[78] W. Lu and J.C. Rajapakse, "Approach and Applications of Constrained ICA," *IEEE Transactions on Neural Networks*, vol. 16, pp. 203– 212, January 2005.

[79] B. Karoumi, A. Laurent, F. Rosenfeld, T. Rochet, A. M. Brunon and J. Dalery, "Alteration of event related potentials in siblings discor-

dant for schizophrenia," *Schizophr. Res.*, vol. 41, no. 2, pp. 325–334, 2000.

[80] J. Yordanova, "The Relationship Between P300 and Event-related Theta EEG Activity," *Psycoloquy, Memory Brain*, vol. 7, no. 25, p. article 7, 1996.

[81] J. C. Mosher, R. M. Leahy, and P. S. Lewis, "Forward solutions for inverse methods," *IEEE Transanctions on Biomedical Engineering*, vol. 46, pp. 245–259, March 1999.

[82] C. -G. Bnar, R. N. Gunn, C. Grova, B. Champagne, and J. Gotman, "Statistical Maps for EEG Dipolar Source Localization," *IEEE Transanctions on Biomedical Engineering*, vol. 52, pp. 401–413, March 2005.

[83] I. D. Coope, "Reliable computation of the points of intersection of n spheres in Rn," *ANZIAM, Journal of the Australian Mathematical Society*, vol. 42, no. 5, pp. C461–C477, 2000.

[84] B. H. Jansen, A. Allam, P. Kota, K. Lachance, A. Osho, and K. Sundaresan, "An Exploratory Study of Factors Affecting Single Trial P300 Detection," *IEEE Trans. on Biomed. Eng.*, vol. 51, pp. 975–978, June 2004.

[85] M. A. Latif, S. Sanei, J. Chambers and L. Spyrou, "Partially Constrained Blind Source Separation for Localization of Unknown Sources Exploiting Non-Homogeneity of the Head Tissues," *Journal of VLSI Signal Processing Systems*, vol. 49, pp. 217 – 232, November 2007.

[86] L. Spyrou, S. Sanei, and C. Cheong-Took , "Estimation and Lo-

cation Tracking of the P300 Subcomponents from Single-Trial EEG," *Proceedings of IEEE, ICASSP 2007*, vol. 2, pp. 1149–1152, April 2007.

[87] L. Spyrou and S. Sanei, "A Novel Approach For The Single-Trial EEG Estimation Of The P300 Subcomponents," *Proceedings of the 5th European Symposium. on BioMedical Engineering*, vol. 1, July 2006.

[88] L. Spyrou and S. Sanei, "A new beamforming approach for the localisation of event related potentials," *Proceedings of EUSIPCO*, pp. 2489–2493, September 2007.

[89] L. Spyrou and S. Sanei, "Localisation of Event Related Potentials Incorporating Spatial Notch Filters," *IEEE Transactions on Biomedical Engineering*, accepted and ready for publication.

[90] P. H. Schimpf, J. Haueisen, C. Ramon, and H. Nowak, "Realistic computer modeling of electric and magnetic fields of human head and torso," *Parallel Computing*, vol. 24, pp. 1433–1460, 1998.

[91] A. Delorme and S Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics," *Journal of Neuroscience Methods*, vol. 134, pp. 9–21, 2004.