# Localization of Brain Signal Sources using Blind Source Separation

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

Mohamed Amin Latif

Supervisor: Dr. S. Sanei

Centre of Digital Signal Processing

Cardiff University

December 10, 2006





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

**DECLARATION**

This work has been not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed.....(candidate) Date 11/12/06

**STATEMENT 1**

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed.....(candidate) Date 11/12/06

**STATEMENT 2**

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit reference.

Signed.....(candidate) Date 11/12/06

**STATEMENT 3**

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed.....(candidate) Date 11/12/06

**STATEMENT 4**

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan after expiry of a bar on access approved by the Graduate Development Committiee.

Signed.....(candidate) Date 11/12/06

## Abstract

Reliable localization of brain signal sources by using convenient, easy, and hazardless data acquisition techniques can potentially play a key role in the understanding, analysis, and tracking of brain activities for determination of physiological, pathological, and functional abnormalities. The sources can be due to normal brain activities, mental disorders, stimulation of the brain, or movement related tasks.

The focus of this thesis is therefore the development of novel source localization techniques based upon EEG measurements. Independent component analysis is used in blind separation (BSS) of the EEG sources to yield three different approaches for source localization. In the first method the sources are localized over the scalp pattern using BSS in various subbands, and by investigating the number of components which are likely to be the true sources. In the second method, the sources are separated and their corresponding topographical information is used within a least-squares algorithm to localize the sources within the brain region. The locations of the known sources, such as some normal brain rhythms, are also utilized to help in determining the unknown sources. The final approach is an effective BSS algorithm partially constrained by information related to the known sources. In addition, some investigation have been undertaken to incorporate non-homogeneity of the head layers in terms of the changes in electrical and magnetic characteristics and also with respect to the noise level within the processing methods. Experimental studies with real and synthetic data sets are undertaken using MATLAB and the efficacy of each method discussed.

# ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| ALF | Adaptive standardized LORETA/FOCUSS |
| AP | Action potential |
| BOLD | Blood oxygenation level dependent |
| BEM | Boundary element method |
| BSS | Blind source separation |
| CBF | Cerebral blood flow |
| CJD | Jakob Creutsfeldt disease |
| CNS | Central nervous system |
| CRLB | Cramer Rao lower bound |
| CT | Computerized X-ray tomography |
| DC | Direct current |
| $d\ell$ | Infinitesmal length of conductor |
| ECD | Equivalent current dipole |
| EEG | Electroencephalography |
| EP | Evoked potential |

| | |
|---|---|
| EPSP | Excitatory postsynaptic potential |
| EVD | Eigen-value decomposition |
| FA | Factor analysis |
| FEM | Finite element method |
| FFT | Fast fourier transform |
| fMRI | Functional magnetic resonance imaging |
| FN | Frobenius norm |
| FOCUSS | FOCal underdetermined system solver |
| Hct | Hematocrit (the percent of the red blood cells in whole blood in the body) |
| HOS | High order statistics |
| Hz | Hertz |
| ICA | Independent component analysis |
| IC | Independent component |
| ICr | Information criterion |
| IPSP | Inhibitory postsynaptic potential |
| JAD | Joint approximation diagonalization |
| JADE | Joint approximate diagonalization of eignematrices |
| JD | joint diagonalization |
| KL | Kullback-Leibler |
| $L_1$ | Norm of a current vector J, defined as $\|\mathbf{J}\|_1 = \sum_{r=1}^{n} \|\mathbf{J}_r\|$ |
| $L_2$ | Norm of a current vector, defined as $\|\mathbf{J}\|_2 = \sqrt{\sum_{r=1}^{n} \|\mathbf{J}_r\|^2}$ |
| LD | linear distributed |
| LR | Likelihood ratio |
| LORETA | Low resolution brain electromagnetic tomography |
| LS | Least squares |
| MCS | Most contributed signal |
| MEG | Magnetoencephalography |

| | |
|---|---|
| MF | magnetic fields |
| MR | Magnetic resonance |
| MRBSS | Multi-resolution blind source separation |
| MRI | Magnetic resonance imaging |
| MSP | Model selection procedures |
| MUSIC | Multiple signal classification |
| MLE | Maximum likelihood estimation |
| MVUE | Minimum variance unbiased estimator |
| $\mu volts$ | Micro volts |
| $\mu_0$ | Magnetic permeability of free space |
| NGA | Natural gradient Algorithm |
| PCA | Principle component analysis |
| pdf | Probability density function |
| PET | Positron emission tomography |
| sLORETA | Standardized LORETA |
| SNR | Signal to noise ratio |
| SOBI | Second order blind identification |
| SQUID | Super-conducting quantum interference device |
| ssLOFO | Standardized shrinking LORETA-FOCUSS |
| V | Volts |
| Voxles | Discrete volume elements |

# OPERATORS

| | |
|---|---|
| $det(\cdot)$ | Determinant of a matrix |
| $diag(\cdot)$ | Diagonal of a matrix |
| $offdiag[\cdot]$ | Off-diagonal elements of a matrix |
| $\ell_p$ | $L_p$ norm |
| $\ell_1$ | $L_1$ norm |
| $\ell_2$ | $L_2$ norm |
| $E\{\cdot\}$ | Expectation |
| $(\cdot)^T$ | Transpose |
| $(\cdot)^H$ | Hermitian/Conjugate transpose |
| $(\cdot)^*$ | Conjugate |
| $\hat{x}$ | Estimated sample |
| $\lvert\cdot\rvert$ | Absolute value |
| $\lVert\cdot\rVert$ | Euclidean norm |
| $\lVert\cdot\rVert_F$ | Frobenius norm |
| $(\cdot)^\dagger$ | Moore-Penrose pseudoinverse |
| arg | Argument |
| min | Minimum |
| $sign$ | Signum |
| $\nabla$ | Gradient operator |

# PUBLICATIONS

### Journal Papers

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", to be appear in The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology, 2006.

M A Latif, S Sanei, J Chambers, 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, no. 3, pp. 117-120, 2006.

### Conference Papers:

M A Latif, S Sanei, and J Chambers, "Localization of abnormal EEG sources incorporating constrained BSS", Proc. of the international conference on artificial neural networks (ICANN2005), Warsaw, Poland, pp. 703-709, 2005.

M A Latif, and S Sanei, "Localization of brain abnormal signal sources using blind source separation", Proc. of the Engineering in Medicine and Biology Conference (EMBC05), Shanghi, China, 2005.

M A Latif, S Sanei, and L Shoker,"An iterative ICA-based algorithm for EEG source localization", Proc. of the 2nd International Conference on advances in Medical Signal and Information Processing (MEDSIP 2004), Malta, 2004.

S Sanei, M A Latif, and L Shoker,"A new constrained BSS algorithm for separation of EEG signals with eye-blinking artifact", Proc. of the third IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM2004), Barcelona, Spain, 2004.

L Shoker, S Sanei, and M A Latif,"Removal of eye blinking artifacts from EEG incorporating a new constrained BSS algorithm",Proc. of the third IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM2004), Barcelona, Spain, 2004.

M A Latif, and S Sanei, "An iterative BSS algorithm for an efficient separation of EEG and localization of the sources", Proc. of the Institute of Physics and Engineering in Medicine (IPEM) Meeting in Signal Processing Applications in Clinical Neurophysiology, York, UK, 2004.

# STATEMENT OF ORIGINALITY

As far as the author aware the majority of the work presented in Chapters 3 to 5 represents original contribution to the area of Blind Source Separation (BSS), with partial constraints, and integration of certain effects of nonhomogeneous layers in localization of brain source signals. The originality is partially supported by two journal and six conference papers. Different approaches for source localization have been proposed, as listed below:-

- In Chapter 3, topographic localization of sources over the scalp pattern using BSS in various subbands and by investigating the number of components which are likely to be the true sources is investigated. In this chapter, detection of the number of sources by implying spatial and frequency constraints is also investigated, and the result is compared with the Akaike Information Criterion.
- In Chapter 4, the sources are separated based on correlation and their corresponding topographical information has been used by a least-squares algorithm to localize the sources within the brain region.
- In Chapter 5, the locations of the known sources, such as some normal brain rhythms, have been utilized in two different approaches for determination of the unknown sources. The final approach refers

to development of an effective BSS algorithm partially constrained by certain information related to the known sources. In addition some investigation has been undertaken to incorporate non-homogeneity of the head layers in terms of the changes in electrical and magnetic characteristics and also with respect to the noise level. The experiments have been carried out using MATLAB and the related results presented at the end of each section.

# ACKNOWLEDGEMENT

First of all, I would like to thank my supervisor Dr. Saeid Sanei for his invaluable suggestions, guidance, patience and continuous enthusiastic help and encouragement throughout my research work, which without his support it would not have been possible to carry out my research work.

I am also owe a great deal of gratitude to my co-supervisor Prof. Jonathon Chambers for all his support and suggestions and help in my through out the research work.

Special thanks to my research companions Mr. Leor Shoker, Mr. Loukianos Spyrou, and Mr. Kianoush Nazarpour for their valuable discussion in research and supports.

I am most grateful to my wife Dr. Loabat Amiri for her constant support and encouragement.

# Contents


**1 LOCALIZATION OF BRAIN SIGNAL SOURCES INTRODUCTION** **23**

1.1 Neurophysilogical Basis of the EEG . . . . . 25

1.2 Why Localization is Important? . . . . . 29

1.3 Decomposition of the EEGs in to their Original Sources . . . . . 33

1.4 Objectives of the Research . . . . . 37

1.5 Layout of the Thesis . . . . . 39

**2 BRAIN SOURCE LOCALIZATION OVERVIEW** **40**

2.1 Background . . . . . 40

2.1.1 The concept of dipole modeling of the brain sources . . . . . 40

2.2 EEG Source Signal Decomposition Based on Blind Source Separation 43

2.2.1 Pre processing . . . . . 46

2.2.2 Diagonalization of eignematrices for fourth order cumulant . . 47

2.3 Brain Source Localization Methods . . . . . 51

2.3.1 Equivalent current dipole model . . . . . 53

2.3.2 Linear distributed method . . . . . 54

2.4 Finite Element Method . . . . . 63

2.5 Other Modalities for Detection of Brain Sources . . . . . 64

3 TOPOGRAPHIC LOCALIZATION OF BRAIN SIGNAL SOURCES 66
3.1 Separation . . . . . 66
3.1.1 Separation of the EEG source signals using iterative SOBI algorithm . . . . . 69
3.2 Filtering . . . . . 72
3.3 Localization Over the Scalp using Back Projection . . . . . 75
3.4 Detection of the Number of Sources by implying Spatial and Frequency Constraints . . . . . 76
3.5 Conclusions . . . . . 89

4 LOCALIZATION IN A 3D SPACE BASED ON THE CORRELATION MEASUREMENTS 90
4.1 The Information Held by Mixing Matrix A . . . . . 93
4.2 PCA Rather Than ICA . . . . . 97
4.3 Variance Test . . . . . 98
4.4 Localization of the Brain Sources . . . . . 100
4.4.1 Finding permutation **R** and mixing matrix **A** . . . . . 100
4.5 Separation and Localization of the EEG Signals Based on Correlation Measurement . . . . . 101
4.5.1 Estimation of the point of intersection using the signals' contributions . . . . . 105
4.6 Incorporating Non-homogeneity of the Head Tissues . . . . . 110
4.6.1 Non-homogenity problem . . . . . 111
4.6.2 Non-homogeneity considered as Nonlinearity of the head . . . 112
4.7 Conclusions . . . . . 120

5 PARTIALLY CONSTRAINED METHODS 122
5.1 Extension of the Method Based on the Correlation Measurements . . 122

5.2 Partially Constrained BSS for Brain Source Localization . . . . . 128
5.3 Selection of the Known Sources . . . . . 144
5.3.1 Conclusions . . . . . 147

**6 SUMMARY, CONCLUSIONS, AND FUTURE WORK 149**
6.1 Summary . . . . . 149
6.2 Conclusions . . . . . 151
6.3 Future work . . . . . 152

**7 APPENDICES 154**

# List of Figures


1.1 Schematic structure of a neural cell (neuron). The neuron consists of a cell body, dendrites, axon and synapse. Adopted from [7]. . . . . 26

1.2 Schematic drawing of an action potential, adopted from [8], indicating the resting potential and the rising phase passing the threshold together with overshoot. In approximately 30mV the potential suddenly drops down. All this occurs within one millisecond. . . . . 27

1.3 Anatomy of the brain, adopted from [3], [4], and [5]. . . . . 30

1.4 The division of sensory (left) and motor (right) function in the cerebral cortex, adopted from [13]. . . . . 30

1.5 Some examples of the EEG waves in different frequencies. . . . . 31

1.6 Cocktail party problem . . . . . 33

1.7 ICA model for a cocktail party problem . . . . . 34

1.8 ICA unmixing operation showing the localized sources within the brain, the scalp sensors and the extracted IC components. . . . . 37

2.1 The magnetic field **B** at each electrode is calculated with respect to the moment of the dipole and the distance between the center of the dipole and the electrode. . . . . 41

2.2 The magnetic field **B** at each electrode is calculated with respect to the accumulated moments of the $m$ dipoles and the distance between the center of the dipoles' volume and the electrode. . . . . 42

3.1 Topographic localization based on back projection . . . . . 67
3.2 EEG signals from a patient with CJD symptoms . . . . . 73
3.3 Filtered signals at Delta band, from a patient with CJD symptoms (DFT is used for transformation). . . . . 73
3.4 EEG recordings of a finger movement of a normal awake person. . . . 74
3.5 Space (electrodes)-frequency representation of filtered signals related to finger movement. AR with order of 8 is used for transformation. Hence, the highest peak amplitude at the electrode location seven in electrode axis with frequency of 8 Hz is considered as the known source. 74
3.6 A 2D topological view of the brains ICs for a patient with CJD. The dark red are the highly active areas: (a) IC1 (b) IC2 (c) IC6 (d) IC9. 76
3.7 A 2D topological view of the brains ICs for a patient in state of drowsiness. The dark red region represent high activity areas; frontal and temporal: (a) IC1 (b) IC7 (c) IC8 (d) IC9 . . . . . 77
3.8 A 2D topological view of the brain's ICs for a patient with focal seizure. The locations of most dominant sources are (a) IC5, (b) IC8, (c) IC9, and (d) IC10. . . . . 78
3.9 Four synthetic waveforms in 8, 9 ,10 , and 11 Hz representing the source signals in Alpha band . . . . . 78
3.10 Mixtures of source and noise signals . . . . . 79
3.11 Plot of eigenvalues of the mixed signals . . . . . 79
3.12 Distribution of the source signals . . . . . 80
3.13 Distribution of the noise signals . . . . . 80

3.14 Akaike information criterion when the noise information is available. Plot of $n$ sources from $m$ mixtures against ICr, where the Criterion reaches to a minimum value. . . . . 83
3.15 Akaike information criterion when the noise information is not available. Plot of $n$ sources from $m$ mixtures against ICr, where the Criterion reaches to a minimum value. . . . . 83
3.16 Distribution of the source signals in a patient just before focal seizure. 84
3.17 Distribution of the noise signals in a patient just before focal seizure. 84
3.18 Distribution of the seizure related source signals in a patient during focal seizure. . . . . 85
3.19 Distribution of the noise signals in a patient during focal seizure. . . . 85
3.20 The topography of the projected ICs for a patient with focal epilepsy before seizure (a) IC8 (b) IC9 (c) IC10, in these figure the amplitude of the peak in frequency domain at this IC is at the highest peak. . . 86
3.21 The topography of the projected ICs for a patient with focal epilepsy during seizure (a) IC5 (b) IC10, in this figure the amplitude of frequency at this IC is at the highest peak. . . . . 87
3.22 Plot of ICr against the number of sources before focal seizure. "o" indicates ICr of this particular EEG data. The first minimum value in the curve, which is the minimum of ICr versus $m$ identifies the number of independent sources. In this case the number of sources are three. . . . . 87
3.23 Plot of ICr against the number of sources during focal seizure. "o" indicates ICr of this particular EEG data. The first minimum value in the curve, which is the minimum of ICr versus $m$ identifies the number of independent sources. In this case the number of sources are two. . . . . 88

4.1 Signal: $s_1(t) = \sin(40\pi t)$ . . . . . 94
4.2 Signal: $s_2(t) = \sin(24\pi t)$ . . . . . 94
4.3 Recovered signals by using the described separation algorithm . . . . 95
4.4 The first n number of Eigenvalues of the 4th order cumulant tensor . 99
4.5 The cross-correlations between the estimated sources and the mixtures and the corresponding correlation charts . . . . . 100
4.6 Convention 10/20 . . . . . 102
4.7 Correlation between 1st contributor and the mixed signals . . . . . 103
4.8 Blocking each contributor after correlation. . . . . 104
4.9 Definite contributors to the mixed signals. . . . . 104
4.10 Electrode position within a Cartesian coordinate . . . . . 106
4.11 Cross-section of 3 spheres with an intersection point (left) and without a common intersection point (right) adopted from [110] . . . . . 108
4.12 Intersection of three sphere in a common point represents location of a source . . . . . 108
4.13 Part of scalp including the electrodes and the location of a source to be identified . . . . . 109
4.14 Homogeneous spherical model of the brain . . . . . 114
4.15 The graph of relationship between cross correlation of the source signal verses square of distance between the electrode at the scalp and sources inside the head in decimeters. The slope for Homogeneous spherical model of the brain is shown in doted line and an asymptotically decaying curve between the same cross correlation considering the non-homogenous case shown in solid line. . . . . 116
4.16 Concentric spherical head model by Rush and Driscoll (1969). The model contains a region for the brain, scalp and skull, each of which is considered to be homogeneous. . . . . 116

4.17 Curve of Non-Homogeneous spherical model of the brain considering the effect of noise shown as solid line, and for the homogenous case shown as doted line . . . . . 118

5.1 The original electrode signals during an ictal period. . . . . 126

5.2 Representing the normalized correlation between the estimated sources and the mixtures (electrode signals), e.g the second bar in the third frame is the normalized correlation between the second mixture and the third output. . . . . 127

5.3 The separated dominant sources from the signals in Fig. 5.1; the top figure represents the normal Alpha rhythm. The middle one is clearly an epileptic seizure signal with a frequency of between 6-7 Hz. The bottom signal is a Delta rhythm. . . . . 127

5.4 A section of the scalp including three electrodes and locations of the two sources . . . . . 130

5.5 The estimated source locations when (a) there is only one out of three sources known, and (b) two sources are known. "•" represents the sensors locations, "∗" shows the actual locations of sources, and "□" represents the estimated locations of the sources. x, y, z are towards the front, lateral-right, and plannar views. . . . . 134

5.6 Error versus the number of known sources . . . . . 135

5.7 Convergence of the unmixing matrix $\mathbf{W}$; Plot of Error against number of iteration, Error=$\|\mathbf{W}(t+1)-\mathbf{W}(t)\|_F$, where $\mathbf{W}$(t) is old unmixing matrix, and $\mathbf{W}$(t+1) is new unmixing matrix after each iteration. . . 136

5.8 Least sqaures convergence plot, Error = $\|\mathbf{f}-\mathbf{d}\|_2^2$, where $\mathbf{f}$ is the source and $\mathbf{d}$ is inverse of correlations between the estimated source and electrode position. . . . . 136

5.9 *error − distance* plot against values of $\alpha$ in (4.10) described in section "Non-homogeneity considered as Nonlinearity of the head" for homogeneous spherical model of the brain . . . . . 138

5.10 The estimated sources using BSS are shown in by " □", which are overlapping the " ∗" representing the actual sources, and "•" represent the locations of the sensors for a homogeneous spherical model of the brain . . . . . 139

5.11 A synthetic source signals: (a) Original Signals, (b) The result of BSS; the sources are scaled and permuted. The additional noise level is 0.1 percent. . . . . 140

5.12 *error − distance* between the original and estimated source locations with the effect of 2 percent noise added to the mixture signals. An average value of $\alpha = 0.7$ gives a minimum *error − distance*. . . . . 141

5.13 The effect of nonlinearity parameter $\alpha$ and *error − distance* of estimated sources: the location of the original source is shown in star, the triangles indicating the noisy estimation of the sources with nonlinear parameter value of $\alpha = 0.7$. The black points are the electrode locations. The numbers shown on each location indicates the original order of each signal. . . . . 141

5.14 Distance error between original and estimated first, second, and third source locations respectively, with the effect of 5 percent noise to the mixture signals. An average $\alpha = 0.7$ gives a minimum *error − distance* between the original source and the estimated one. . . . . 142

5.15 Selection of highest amplitude level in Alpha rhythm as for the known source . . . . . 143

5.16 Localization of the real EEG sources (a) The top view, and (b) The lateral perspective view of the locations of electrodes shown by " • ", the locations of the known sources are shown by "□" and the estimated locations of the unknown sources are shown by " ∗ " . . . . 144

5.17 Localization of the real EEG sources with the effect of non-homogeneity; (a) the top view, and (b) the lateral perspective view of the locations of electrodes shown by " • ", the location of the known source shown by "□" and the estimated locations of the unknown sources shown by " ∗ ". The calculated nonhomogeneous (nonlinear) effect on establishing an estimated location for the unknown sources are shown in " △ ". Each location is marked with a number and some of the sources found are overlapping . . . . . 145

5.18 Implementation of the dipole fitting method from the EEG lab, which confirms the location of some of the sources computed using the proposed partially constrained method. . . . . 146

5.19 Projection of kurtosis to the scalp topography. The red patch in Beta range is the harmonic of the signals in Theta range. . . . . 146

# List of Tables

4.1 Geometrical locations of the electrodes over the scalp within a cartesian coordinate . . . . . 107

5.1 Error verses known sources . . . . . 135

5.2 Error values for a set of $\alpha$ for homogeneous spherical model of the brain 139

# Chapter 1

# LOCALIZATION OF BRAIN SIGNAL SOURCES INTRODUCTION

Medical signal processing has an important role in technological science and enhancement of new systems such as computerized x-ray tomography scanner (CT), electroencephalographs (EEGs), magnetoencephalographs (MEGs), functional magnetic resonance imaging (fMRI), optical recording, and wireless communication. As advanced and reliable methods for recording and visualization are more in demand, methods such as blind source separation (BSS) become a potential enabler for new technological applications especially in biomedical engineering, econometrics, exploration seismology, communication technology, medical imaging, remote sensing, and voice enhancement.

Biologists in particular, neuroscientists, require reliable techniques, which are able to extract and separate useful information contained within brain source signals, by using non-invasive recordings of human brain activities such as EEG and MEG for

instance, in order to understand the brain's ability to sense, recognize, store and recall patterns and crucial elements of learning [129], or locating sources of abnormalities such as due to epilepsy [130], Parkinson's disease [132], and schizophrenia [131]. This can help in the design of a course of treatment. Therefore, the neurologist requires a technique to find the brain source signals and their locations, and to be able to identify instants of time during which dynamic changes are taking place. The result of hemodynamic responses in the brain can be recorded by the existing brain imaging modalities such as positron emission tomography (PET) and fMRI with a good spatial resolution, on the order of millimeters. But due to the slowness of hemodynamic responses of neural activity, the location of sources can only be monitored with a very low temporal resolution. Direct measurements of the electromagnetic fields produced by the neuronal activity on the other hand have a temporal resolution of less than 1 ms. The spatial resolution however, depends on the number of EEG electrodes.

A number of methods for localization of electromagnetic sources have been investigated by researchers, among them methods based on an equivalent current dipole (ECD) model assumption of the sources have been very well established. In this approach it is assumed that the EEG/MEG signals are generated by a small number of focal sources which are modelled as single temporally independent [62], fixed, or reoriented dipoles. In this approach the number of focal sources is considered less in number than the available measurements (i.e. the mixing system is overdetermined). For the models with a large number of sources, some approximation techniques should be used. The solutions depend on the initial estimate of the locations and orientations of dipoles. In all ECD-based models, the solution relies on a given number of dipoles, but in reality, the actual number of sources cannot be determined (*a priori*). Another method is to consider all possible fixed source locations as an *a priori*. This assumption would generate a model known as the linear

distributed (LD) model. This continuous current source model is more biologically approved than the ECD model [140]. However, the accuracy of such an algorithm is dependent on the number of both sources and sensors, otherwise the solution will be ill-posed. This means that there are many possible solutions, and hence an issue of uniqueness.

EEG source localization requires a solution to an ill-posed inverse problem. Therefore many solutions can lead to the same response. An effective and simple technique for both separation and localization of the EEG sources therefore requires a priori information to constraint the solutions. This may be achieved by incorporating some physiological properties of the mixing system or by fusing the available information about the known sources and their locations into the algorithm. This approach is exploited in the major contributions within the thesis.

## 1.1 Neurophysilogical Basis of the EEG

### Basic physiology

The brain has a mass of 1-2 kg in an adult person. It consumes 25 percent of the body's total glucose, 20 percent of the body's total oxygen, its cortex has a surface area of approximately $3000cm^2$ and there are about $3 \times 10^{10}$ neurons in the human's brain cortex [5]. Within the central nervous system (CNS) there are 100 billion nerve cells [6]. The most common nerve cells, or neurones, as shown in Figure 1.1, may vary in detailed structure from one kind to another, but typically comprise a cell-body, containing a nucleus, a long nerve fiber called an axon which carries nerve-impulses away from the cell-body, and many short branching processes, known as dendrites, which receive impulses from other cells and transfer them to the cell body [5].

Inside the neuron there is a high concentration of potassium ions ($K^+$), meanwhile

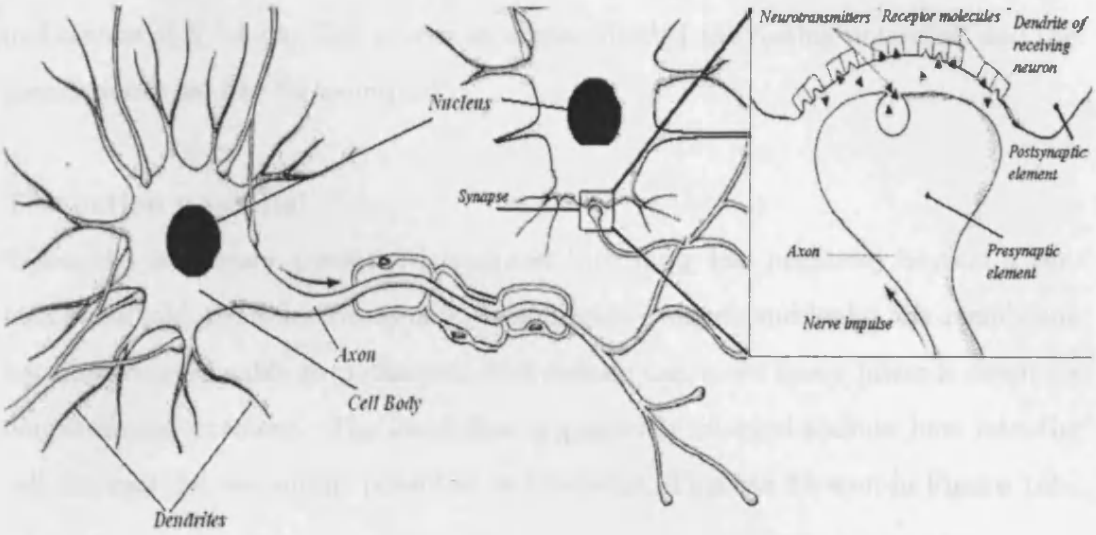



Figure 1.1: Schematic structure of a neural cell (neuron). The neuron consists of a cell body, dendrites, axon and synapse. Adopted from [7].

outside the neuron the potassium concentration is low. On the contrary, sodium ($Na^+$) has a much lower concentration within the nerve cell than within the extracellular fluid. These ions have a natural tendency to diffuse down their concentration gradients, as such potassium $K^+$ leaks out of the cell and sodium $Na^+$ inwards. The cell membrane has the property of being selectively much more permeable to potassium $\mathbf{K}^+$ ions and negatively charge chloride ($Cl^-$) ions, but is impermeable to sodium $Na^+$. Since an unequal amount of $Cl^-$ is distributed across the cell wall a voltage gradient forms at the junction between the intracellular plasma and the outside of the cell. Therefore potassium diffuses out of the cell, but sodium cannot so readily diffuse inwards. This imbalance of ionic movements causes a net loss of positive charge from the cell, creating a difference of electrical potential across the cell membrane. An equilibrium is established when the negative potential inside the membrane with respect to the outside becomes large enough to prevent the contin-

ued escape of $K^+$ ions. This occurs at about -70mV ( the resting potential) and the membrane is said to be 'polarized'.

### The action potential

When the membrane potential is reduced (becoming less negative) beyond a certain threshold the selective ionic permeabilities change: suddenly, the membrane becomes impermeable to potassium, but sodium can move freely inwards down its concentration gradient. The local flow of positively charged sodium ions into the cell reverses the membrane potential at this point. This can be seen in Figure 1.2.

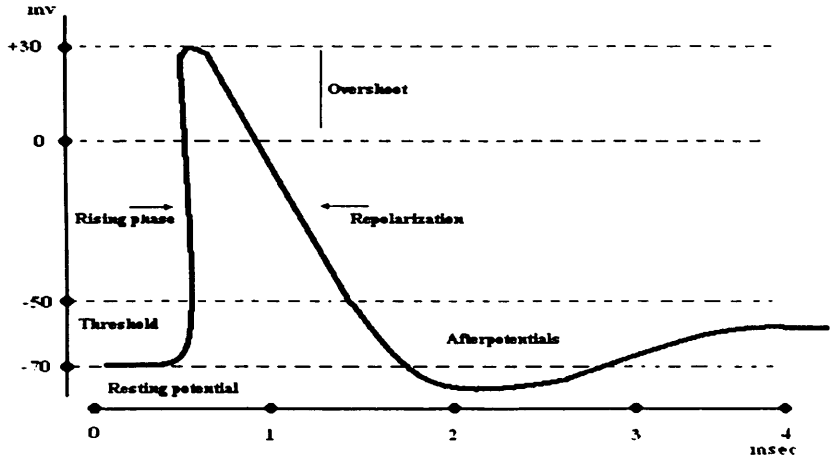



Figure 1.2: Schematic drawing of an action potential, adopted from [8], indicating the resting potential and the rising phase passing the threshold together with overshoot. In approximately 30mV the potential suddenly drops down. All this occurs within one millisecond.

When the inside of the membrane has become some 30mV positive with respect to the outside, the selective permeabilities rapidly return to their previous states. The inflow of sodium ceases and potassium again diffuses outwards, thus restoring the

original membrane potential, and indeed producing a short-lived overshoot. This sequence of events lasts approximately 1 ms and is termed the action potential (AP). An AP occurring at a point in the cell membrane sets up local electric currents which in turn depolarize adjacent parts of membrane. Consequently, the AP spreads across the surface of the cell and in particular is swiftly propagated down the axon. The repeat operation of AP would tend to cause a gradual accumulation of sodium and loss of potassium within the cell. This is corrected by active transport of these ions across the cell membrane by a biological pump. The whole process consumes a large amount of energy which demands a high level of metabolic activity, hence it requires constant supply of oxygen and glucose for its survival. The AP transmits signals along the nerve fibres, different mechanisms are present in transmission between nerve cells. The axon of a neuron typically divided into several terminal branches which make contact with the dendrites and cell-bodies of other neurons. These points of close contact are called 'synapses'. When an AP passing down the axon reaches the synapse, a chemical substance called a 'neuron-transmitter' is released, which diffuses across the very small distance to the membrane of the adjacent 'post-synaptic' neuron and produces changes in the physical properties of its membrane. The membrane of the nerve cell potential is subject to various fluctuations caused by synaptic activities (i.e. micro-electrode values of 60-70mV [5], with negative polarity).

These spontaneous activities are recorded at the electrode location and referenced against another recoded electrode location or ear lobe (i.e. silent electrode), with a frequency limit of up to approximately 100Hz within few milliseconds. This signal is referred to as the electroencephalograph (EEG) measurement.

Measuring the brain activity can be recorded in different ways

- Spontaneous: The normal EEG, which contains some activities that are un-

correlated with any stimuli. This can be observed both during and in between stimulations, and is called spontaneous activity.

- Evoked potentials (EPs) are those components of the EEG that arise in response to a stimulus ( which may be electric, auditory, or visual). The amplitude of evoked potentials is usually smaller than that of the spontaneous activities, and they are rarely visible in a single recording. These signals are also mostly below the noise level and thus not readily distinguished. One must use a train of stimuli together with signal averaging to improve the signal to noise ratio (SNR).

- Induced: Similar to EP, induced potentials are directly caused by the experimental stimuli. They appear with varying latency or phase.

## 1.2 Why Localization is Important?

The human brain weights approximately 1.5 kg [12]. It is divided into two hemispheres. The hemispheres are divided into four lobes: frontal, parietal, occipital, and temporal lobes, as shown in Figure 1.3. The largest part of the brain is the cerebrum which has a surface area of approximately 1600 $cm^2$, and its thickness is 3mm. Each section of the brain consists of different neural types which are responsible for activities or functionalities of different part of human bodies. For example, the visual sensors are located in the occipital lobe, and the sensory area and motor area are located on both sides of the central fissure. As shown in Figure 1.4.
The highly expanded surface of the brain neurons are arranged in functional columns within the distinct cellular layers that make up the cerebral cortex in each of the two hemispheres. This fact has allowed for the functional classification of the cortex into areas mediating awareness in each of the sensory modalities, control of voluntary

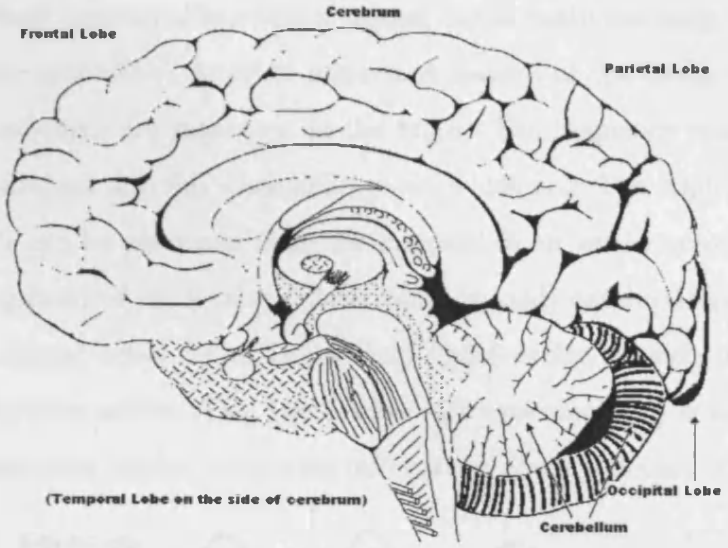



Figure 1.3: Anatomy of the brain, adopted from [3], [4], and [5].

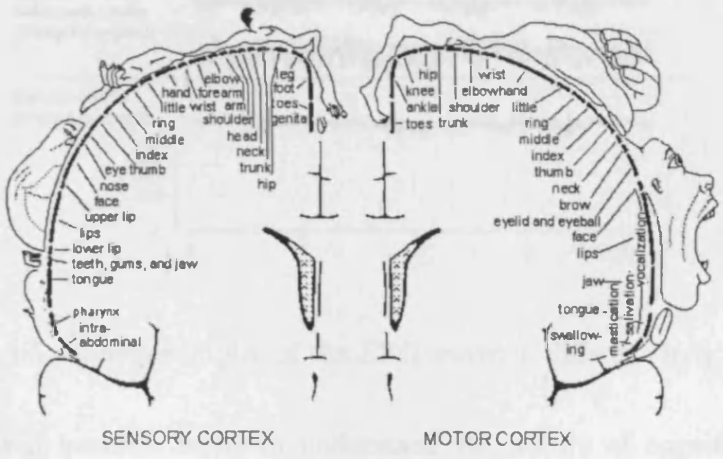



Figure 1.4: The division of sensory (left) and motor (right) function in the cerebral cortex, adopted from [13].

movements, storage of information, and the capacity for higher intellectual activities. Knowledge about functionality of each section of the brain can help to detect the source of these activities. Another important feature of the brain activities are the frequencies which are generated in the brain. The frequency spectrum of the brain can be divided into following main parts, which are: The Alpha waves (8-13 cycles/second) can be measured from the occipital in an awake eye-closed person. The frequency band of Beta wave (13-30 cycles/second) can be detected over the parietal and frontal lobes. The Delta band (0.5-4 cycles/second) is detected in infants and sleeping adults. The Theta wave (4-8 cycles/second) is mostly seen in children and sleeping adults. Acquiring information about different brain activities

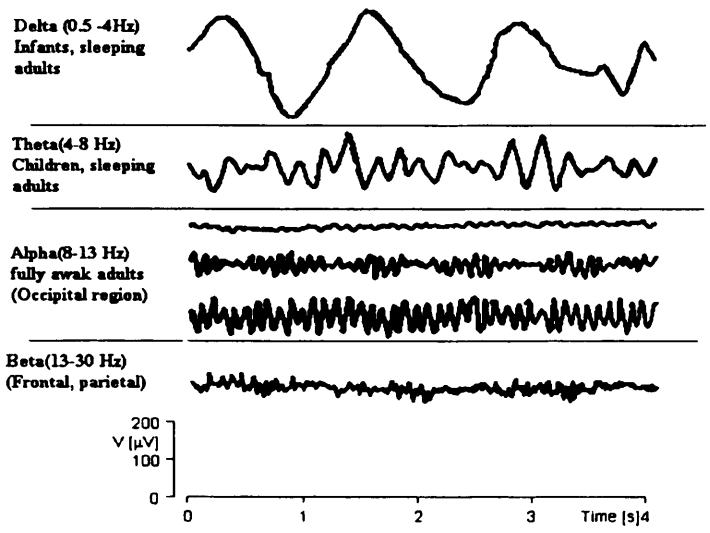



Figure 1.5: Some examples of the EEG waves in different frequencies.

with the related location helps to understand the nature of cognition, memory, and other related brain functionality. Biologists and neuroscientists require reliable techniques which are able to obtain information about a human's brain activity and their location, non-invasively but instantly. The result of slowness in hemodynamics

responses causes the information regarding actual location of the sources to be lost, without being monitored. From the other side, the signal captured with the aid of the EEG / MEG measurements reflects the brain's neuronal electrical activity within an instant in time. The neuronal electrical current in the brain consists of tangential and radial components. MEG can only capture the tangential sources, but EEG is able to convey the tangential and radial characteristics of the sources. In the middle of the sphere the sources are always radial, and such sources anywhere in the sphere do not produce any external magnetic field. EEG, by contrast, would pick up all these currents (tangential, radial, and deep). In addition MEG can be easily affected and distributed by head motion and by the external field. An algorithm is therefore required to extract and separate and localize useful information from superimposed signals within the brain with spatial and temporal synchrony, using EEG recordings. To localize a source there are some physiological aspects which should be considered.

The source signals measured over the scalp penetrate and pass through the white and grey matters, cortex region, and blood fluid, skull, and scalp each of which have different conductivity or resistivity, before reaching the sensors these mixed signals are subject to noise and attenuation. The noise may be generated from other underlying unwanted signals, or physical muscle movements in the body such as eye blink and heart beats. The noise could also be generated from the outside body environment, such as electrodes connected to the head, or measurement device noise. In order to achieve an accurate localization, all the above aspects should be taken into consideration.

## 1.3 Decomposition of the EEGs in to their Original Sources

Brain source separation in some ways resembles the cocktail party problem. Normally the human auditory system is able to block other speeches and even the background noise in the room and concentrate on only one person at a time in the party. This can be represented as a model which is able to separate sources from each other as shown in Figure 1.6. It is required to design a machine to separate individual speeches. This is the well known problem often referred to as the "Cocktail Party Problem". It is essential to mention that in the cocktail party problem there is a time lag between each received signal to each sensor, where as in this work the source signals from the brain are considered instantaneous. The travelling time for the signals is therefore much less than the sampling interval. The solution to these kinds of problems is called "Blind Source Separation" (BSS). The term "Blind" means that, one knows very little, if anything, about the observed data, and makes little assumptions on the source signals. This can be described as a


Figure 1.6: Cocktail party problem

block diagram model as shown in Figure 1.7. The observed data which are mixtures

of some unknown sources are separated at the right end of Figure 1.7. They are called independent components (IC) of the observed data. ICs exploit statistical independence within a set of measurements. Independent component analysis (ICA) is a computational technique for revealing hidden factors such as statistical independence that underlies sets of measurements or source signals. ICA is the most widely used algorithm for BSS. Generally, BSS estimates the original source signals even if they are not completely mutually statistically independent, while the objective of ICA is to determine such transformation which assures that the output signals are as independent as possible.

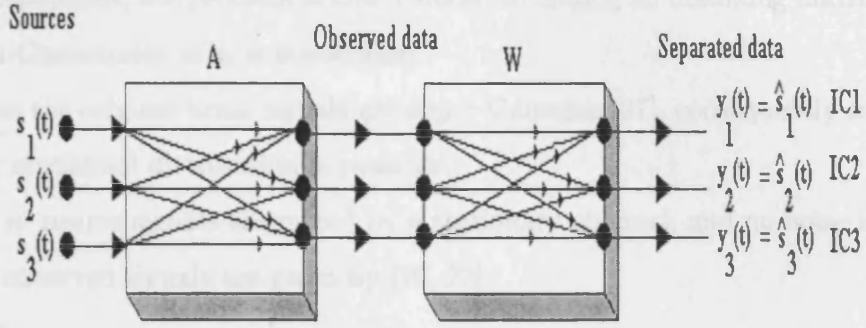



Figure 1.7: ICA model for a cocktail party problem

EEG measurements at the scalp electrodes are a sum of the large number of brain cell (neuron) potentials. During the EEG examination, an issue of interest is from where the brain source signals are originated. Measuring sources directly from within the brain requires inserting electrodes inside the head invasively by surgery. This is not desirable since it would be a risk to the subject and is not recommended in early stages of brain disease development. A better solution would be to obtain the desired signals from the EEG obtained on the scalp non-invasively.

A pictorial representation of separating the ICs of the sources from the sensors on the scalp is shown in Figure 1.8. The sensors are different linear combinations of the sources, and sources are statistically independent (independent components). Another restriction for ICs extracted from the brain is that they generally must have non-Gaussian distribution, since two joint Gaussian distributions of unit variance have completely symmetric distribution. Therefore, it does not contain any information on directions of the columns of the mixing matrix; consequently the mixing matrix can not be estimated. According to the central limit theorem, the sum of independent random variables has a distribution that is closer to Gaussian than the distribution of the source signal. As the components are known to be non-Gaussian and independent, the problem is now reduced to finding an unmixing matrix so that the non-Gaussianity of $s_i$ is maximized.

Hence as the original brain signals are super-Gaussian [27], consequently separation of their statistical distribution is possible.

When $n$ source signals are mixed by a stationary channel, and no noise is present the $m$ observed signals are given by [16, 22]

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{1.1}$$

where $\mathbf{x}(t) = [x_1(t) \ldots, x_m(t)]^T$ is the m-dimensional vector of observed signals (or sensors), $\mathbf{s}(t) = [s_1(t), \ldots, s_n(t)]^T$ is the vector of source signals (or sources), $[.]^T$ denotes vector transpose, $\mathbf{n}(t) = [n_1(t) \ldots, n_m(t)]^T$ and $t$ denotes the discrete time index. $\mathbf{A}$ is an $m \times n$ dimensional mixing matrix, containing information about sensors and direction of arrival of the source signals. The solution is found with the following limitations:

1. The energy of the sources cannot be determined because a scalar multiplier in $s_i$ results in the same $\mathbf{x}$ as a scaling of the $i^{th}$ column in $\mathbf{A}$.

2. Sign of the sources cannot be determined for the same reason.

3. The order of the sources cannot be determined, because swapping two sources results in the same $\mathbf{x}$ if the corresponding columns in $\mathbf{A}$ are swaped accordingly.

Equation (1.1) is for instantaneous mixing system. In a convolutional model where the mixing system varies with time this equation changes to:$\mathbf{x}(t) = \mathbf{H}(t) * \mathbf{s}(t)$ where $*$ denotes convolution operation. However the mixing model for the EEG signals is considered instantaneous as the time lag may be only within a few milliseconds. Therefore for an instantaneous mixing model $\mathbf{s}$ can be computed as

$$\mathbf{y}(t) = \hat{\mathbf{s}}(t) = \mathbf{W}\mathbf{x}(t) \tag{1.2}$$

where $\mathbf{y}(t) = \hat{\mathbf{s}}(t)$ is the estimated vector source signals, $\mathbf{W} = \mathbf{RDA}^{-1}$ is the $n \times m$ dimensional unmixing matrix, or inverse of mixing matrix, subject to permutation (denoted by $\mathbf{R}$) and scaling (denoted by $\mathbf{D}$) of the columns.
Although the skull attenuates the EEG signals, this does not affect the linear relation between the potentials in the brain and the potentials measured over the scalp [25]. Electrical or magnetic fields propagate to the sensors (electrodes) through the brain and skull without significant delays. Therefore, an instantaneous BSS can be adopted. Moreover, the environment and the source signals for the above case are non-stationary (e.g. semi-stationary or approximately cyclostationary). Additive Gaussian noise also affects the signals.

The EEG can be decomposed mathematically into a set of spatial and temporal independent components [28]. If the decomposition is confined to describing the distribution of components across the scalp, no model of head or its contents is required [30]. The decomposition can be done by producing components that are mutually orthogonal, such as principal component analysis (PCA), common spatial

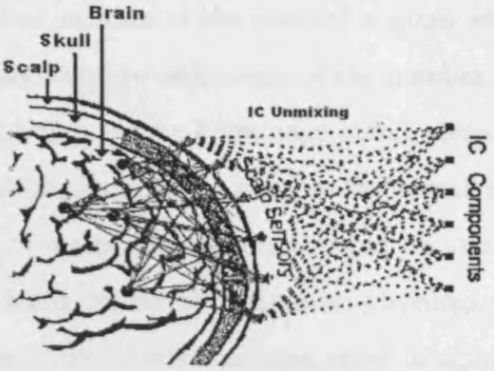



Figure 1.8: ICA unmixing operation showing the localized sources within the brain, the scalp sensors and the extracted IC components.

pattern decomposition [28], and ICA. ICA algorithms presently applied to brain data depend on several assumptions about the underlying processes which requires some *a priori* knowledge that may not be fully realizable. Therefore, it is desirable to alter the ICA methodology by weakening its existing assumptions to permit more proper modelling of EEG dynamics. Any *a priori* knowledge about the sources' statistics or their locations may be mathematically incorporated into the conventional ICA algorithms as constraints. The noise and interferences involved in the EEGs can be exploited in the localization process if their characteristics and properties can be determined.

## 1.4 Objectives of the Research

The objective of this reseach is to find a solution to estimation of the locations of the neural activities from the EEG data, based on Blind Source Separation (BSS) technique. Three approaches are followed for this purpose. In the first approach the estimated components are projected and localized over the brain. In the second

approach the geometrical location of the selected sources within the brain are estimated. This involves an accurate estimation of the number of the sources. Finally, some constraints regarding *a priory* knowledge about some of the sources are embedded into the separation algorithm to improve the overall separation/ localization results. Identification of a single local source could be straightforward, i.e. those in sensory projection areas, or known anatomico-physilogical constraints could be used. It is also possible to use constraints from other imaging modalities [15]. However, when the source area is extended or when several areas are active at the same time, the patterns may be very complex. One can then identify the first source using different modalities or specific physiological conditions as a constraint, and then try to locate the remaining sources.

To describe the objectives, it is useful to mention that the equivalent current dipole (ECD) models are inefficient in accounting for the synchronous activation of broader cortical areas, and are unable to localize more deep sources [81], [17]. Alternatively a method described as the linear distributed approach considers all the possible source locations simultaneously and it has been considered in this work as the main inverse solution to the localization problem.

Concluding, the enquiry in this thesis should therefore consider the nature of the sources as a linear distributed model. The information which describes the location of certain known sources can then be used as the constraint within the proposed BSS algorithm. Hence this should lead to a solution to the ill-posed inverse problem of source localization. Non-homogeneity of the head tissues, on the other hand, should also be exploited.

## 1.5 Layout of the Thesis

In Chapter 2, a brain source localization overview is given, and the background work covers the bioelectromagnetism, and concept of dipole modelling of the brain sources. Furthermore, BSS methods using PCA and ICA are discussed, and the JADE algorithm [29] is described. In the same chapter some brain localization methods and different modalities are explained. In Chapter 3, an iterative separation technique based on second order blind identification is described, topographic localization of sources plus the issues about the number of sources with corresponding results are explained. In this chapter detection of the number of sources by implying spatial and frequency constraints is also investigated, and the result of independent sources are compared with the Akaike's information criterion. Chapter 4, describes: localization based on correlation; the information inherent within the mixing matrix $\mathbf{A}$; the issues of number of sources; and also the issue of attenuation and noise caused non-homogeneity of the head tissues, with corresponding results. In Chapter 5, two methods based on partial constraints with some known brain sources is introduced, and some related simulations such as separating EEG recording with constraint on some known brain sources, and utilization of update equations in BSS for separation and consequently localization of sources are achieved. It is also demonstrated how EEG readings can help to approximately localize sources based on back projection of kurtosis on to the electrodes over the scalp. At the end of the chapter some results are drawn. In Chapter 6 conclusions and future work are presented. Last but not least in Chapter 7, the Appendices non-Gaussianity, least-squares estimation concept from information theory are outlined.

# Chapter 2

# BRAIN SOURCE LOCALIZATION OVERVIEW

## 2.1 Background

### 2.1.1 The concept of dipole modeling of the brain sources

In practice a way to investigate the functionality of a living organisms is to build a model that shows the operational behavior as close as possible to the real one. In most of the cases the hypothesis features make interactions between several variables, which have a same characteristics difficult to solve experimentally. The basic laws of science should be used to control a model's behavior (e.g., Ohm's law, Ampere, etc.).

The reason to construct a model is to handel better the interpretation of a practical phenomenon. The experiment that can be carried out on a model may not be possible to be done on a real living tissues. One can understand the model performance better than a real observable event.

The basic model of a directional magnetic source is a dipole. For a dipole at location

l, the magnetic field observed at electrode $i$ at location $\mathbf{r}(i)$ is achieved as shown in Figure 2.1

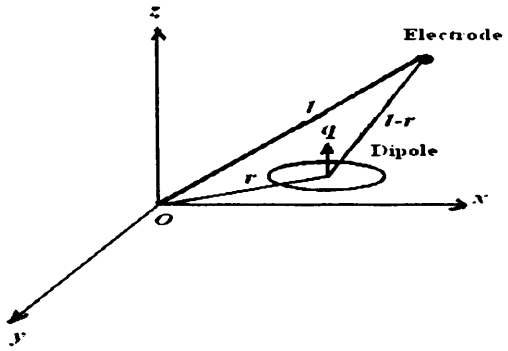



Figure 2.1: The magnetic field $\mathbf{B}$ at each electrode is calculated with respect to the moment of the dipole and the distance between the center of the dipole and the electrode.

$$\mathbf{B}(i) = \frac{\mu_0 \mathbf{q} \times (\mathbf{r}(i) - \mathbf{l})}{4\pi |\mathbf{r}(i) - \mathbf{l}|} \qquad for\ i = 1, \ldots, n \tag{2.1}$$

where $\mathbf{B}$ is the magnetic field, $\mathbf{q}$ is the dipole moment, which is a vector whose magnitude is the product of the current in and the area of loop and whose direction is the direction of the thumb as fingers of the right hand follow the direction of the current. $\mu_0$ denotes the magnetic permeability of free space=$4\pi \times 10^{-7} N/A^2$. This is frequently used as the model for magnetoencephalographic data observed by magnetometers. Equation (2.1) can be extended to the case of a dipole volume. The effect of a volume containing $m$ dipoles on an electrode is shown in Figure 2.2. In the $m$-dipole case the magnetic field at point $j$ is obtained as:

$$\mathbf{B}(i) = \frac{\mu_0}{4\pi} \sum_{j=1}^{m} \frac{\mathbf{q}_j \times (\mathbf{r}(i) - \mathbf{l}_j)}{|\mathbf{r}(i) - \mathbf{l}_j|} \qquad for\ i = 1, \ldots, n \tag{2.2}$$

where $n$ is the number of electrodes and $\mathbf{l}_j$ represents the location of the $j^{th}$ dipole. The matrix $\mathbf{B}$ can be considered as $\mathbf{B} = [\mathbf{b}(1), \mathbf{b}(2) \ldots \mathbf{b}(n)]$.

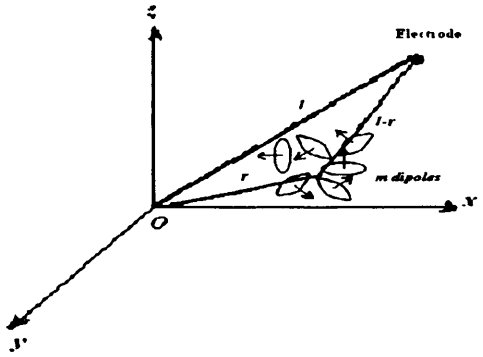



Figure 2.2: The magnetic field **B** at each electrode is calculated with respect to the accumulated moments of the $m$ dipoles and the distance between the center of the dipoles' volume and the electrode.

Single or multiple current dipole fits using Least-Square or multiple signal classification (MUSIC) based approaches [113] are certainly the most popular source modelling techniques in EEG and MEG. But one of the key limitations of current dipole source models is that they may prove to be inefficient in accounting for the synchronous activation of broader cortical areas. The estimation of locations, orientations, and strengths of ECD are limited to only few parameters in all. In all ECD-based methods, the solution depends completely on the number of dipoles assumed but in reality, the actual number of ECD can not be determined a *priori*. Furthermore, the dipole or multi-pole model is unable to localize more deep sources [81]. The ECD may fit properly to the forward field produced by such an activation, but it is very likely that its location would be far from the centroid of the true activated patch, thus making the appraisal of the localization results rather problematic [17].

## 2.2 EEG Source Signal Decomposition Based on Blind Source Separation

An important and great challenge in biomedical engineering is to asses the physiological changes occurring inside the human brain non-invasively. Often the variations can be modelled as source signals that can be measured by techniques such as EEG, MEG and other non-invasive methods.

Brain signals are usually very weak and attenuate when passing through various medium such as skull. They are also non-stationary, and fluctuates irregularly to all directions, therefore no *a priori* knowledge of the desired signal is available, and the signal may also be distorted by noise and interferences. Moreover, they are usually mutually superimposed and their extraction is very difficult. These facts make extraction of the relevant information from diagnosis very complicated and challenging.

The BSS is a technique concerning the signals separability principles that can be applied to them without any knowledge about their origins. Basically, BSS extracts some source signals from a set of sensors in EEG, which receives a different combination of underlying signals or sources. The signals received at the sensors are considered as the mixture signals. The sources generated within the head, are not measured directly and as described earlier there is no *a priori* information about the mixing process leading to the observed mixtures. However this lack of knowledge about the generating process is compensated with strong statistical assumption about the independence between the sources.

Considering head as a closed box system where only a mixture of output is known through the sensors and the sources cannot be separated and examined one by one. In such system an input could be applied via natural senses: Visual, auditory (hearing), somatosensory (feeling), olfactory (smelling) or gustatory (testing), and the

output recorded with EEG sensors from scalp. At each EEG electrode a superposition of highly correlated neuron potentials (sources) are weighted sums, where the weights depend on the signal path from the neurons to the electrodes. A method to transform a multi-variable data set by subspace orthogonality, is referred to as principle component analysis (PCA).

To describe this mathematically lets consider $\lambda$ as the eigenvalue of a matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ and $\mathbf{x}$ is the corresponding eigenvector:

$$\mathbf{A}\mathbf{x} = \mathbf{x}\lambda \tag{2.3}$$

This is true only if

$$det|\mathbf{A} - \lambda\mathbf{I}| = 0 \tag{2.4}$$

Equation (2.4) can be expanded to a $N^{th}$ degree polynomial in $\lambda$, whose roots are eigenvalues. This means that there are always N eigenvalues, of which some can be equal. Equation (2.3) is valid for any multiple of $\mathbf{x}$, in the following format:

$$\mathbf{AV} = \mathbf{VD} \tag{2.5}$$

$\mathbf{V}$ is orthogonal matrix to $\mathbf{A}$ and $\mathbf{D}$ is the diagonal matrix containing the eigenvalues. The eigenvalues of $\mathbf{A}$ correspond to the power contribution of the principle components to the data set, this means that the first principal component is the one corresponding to the largest eigenvalue. For that reason the eigenvalues and eigenvectors are often sorted by power so that $\lambda_1$ is the first principle component, and so on. The orthogonality of $\mathbf{V}$ means that all the eigenvectors are perpendicular to each other and therefore form an orthonormal basis. It also means that the transpose of $\mathbf{V}$ is the same as its inverse.

$$\mathbf{V}^T = \mathbf{V}^{-1} \tag{2.6}$$

Matrix **A** is diagonalizable if:

$$\mathbf{D} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V} \tag{2.7}$$

**D** is a diagonal matrix. If **A** is already diagonal matrix, the eigenvalues are simply the values along the diagonal and **V** is a indentity matrix.

It is clear that most of the variance in the data set is along the first principle component.

The general application of PCA is to decorrelate the data and reduce the number of signals to analyze. In places where the signal and noise subspaces are separable PCA gives very good separation results. Its drawbacks is that it only manipulates the orthogonality of the data, whereas some other methods which will be described later such as independent component analysis (ICA) uses statistics of all orders to make the outputs statistically independent. PCA is used for certain post and pre processing steps such as exploring the characteristics of the recorded signals, dimensionality reduction, or noise elimination. PCA has the advantages that the analysis can be based on a simple second-order statistics only.

BSS is a technique concerning the signals separability principles that could apply without knowledge about their origins. Basically, BSS estimates some source signals from a set of sensor mixtures. The signals received at the sensors are considered the mixture signals. The sources generated within the head, are not measured directly and as described earlier there is no *a priori* information about the mixing process leading to the observed mixture or sensor available. However this lack of knowledge about the generating process is compensated with strong statistical assumption about the independence between the sources, which mostly in the separation result gives close information to the source or original signal.

In practice exact representation of statistical ICA can not be found. This leads to BSS, which tries to force the processed signals to be as independent as possible. The

ICA requires:

1. The source signals s(k) have zero mean and are mutually statistically independent.

$$P_Y(Y) = \prod_{i=1}^{N} P_{y_i}(y_i) \tag{2.8}$$

where $p_{y_i}(y_i)$ are marginal probability distribution functions and $P_Y(Y)$ is the joint probability distribution function (pdf).

2. The mixing matrix $\mathbf{A}$ in equation (1.1) has full column rank.

3. At most one source has Gaussian distribution.

### 2.2.1 Pre processing

In ICA pre-processing the data before applying different algorithms for estimation of the BSS parameters is an important practice. The pre-processing significantly simplifies the algorithm and can improve the resulting convergence.
Generally the mixtures and the sources are considered to be at zero mean, however if this is not the case, the data can be centered by subtracting the sample mean, forcing mean of the mixture to become zero.
Another pre-processing method in ICA is whitening of the data. Whitening is a coordinate transformation of data with an arbitrary multivariate density function into data with an spherical one [78]. The whitening process un-correlates the data.

#### Whitening

Whitening is a linear transformation process which eliminates second order correlation among the data. Hence vector $\tilde{\mathbf{x}}$ is white when

$$E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = \mathbf{I} \tag{2.9}$$

where $E\{.\}$ is the statistical mean or the expected value, $(.)^T$ is the transpose operation, and $\mathbf{I}$ is identity matrix. A popular transformation is done by using eigen-value decomposition (EVD) of the covariance matrix of the observed data. $E\{\mathbf{x}\mathbf{x}^T\}$=$\mathbf{G}\mathbf{D}\mathbf{G}^T$ where $\mathbf{G}$ is the orthogonal matrix of eigenvectors of $E\{\mathbf{x}\mathbf{x}^T\}$ and $\mathbf{D}$ is the diagonal matrix of its eigenvalues. The whitening can be done by:

$$\tilde{\mathbf{x}} = \mathbf{G}\mathbf{D}^{-1/2}\mathbf{G}^T\mathbf{x} \tag{2.10}$$

where the matrix $\mathbf{D}^{-1/2}$ is $diag(d_1^{-1/2}, \ldots, d_n^{-1/2})$.

Since

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{2.11}$$

then

$$\tilde{\mathbf{x}} = \mathbf{G}\mathbf{D}^{-1/2}\mathbf{G}^T\mathbf{A}\mathbf{s} = \tilde{\mathbf{A}}\mathbf{s} \tag{2.12}$$

Hence the new mixing matrix $\tilde{\mathbf{A}}$ is orthogonal, i.e.

$$E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = \tilde{\mathbf{A}}E\{\mathbf{s}\mathbf{s}^T\}\tilde{\mathbf{A}}^T = \tilde{\mathbf{A}}\tilde{\mathbf{A}}^T = \mathbf{I} \tag{2.13}$$

### 2.2.2 Diagonalization of eignematrices for fourth order cumulant

One approach for estimation of ICs consists of using higher order cumulant tensor. The tensor can be considered as generalization of the covariance matrix. The fourth order tensor is defined by the fourth order cumulant $cum(x_i, x_j, x_k, x_l)$. Thus the (fourth order) cumulant contains all the fourth order information, just as the covariance matrix gives all the second order information. Both moments and cumulants contain the same statistical information, because cumulant can be expressed in terms of sums of products of moments. It is usually preferable to work with cumulant because they present in a clearer way the additional information provided by higher-order statistics. Followings are some useful properties of cumulant:

- If $x_i$, $x_j$, are statistically independent random vectors having same dimension, then the cumulant of their sum $x_t = x_i + x_j$ is equal to the sum of their cumulants. This property also holds for the sum of more than two independent random vectors;

- For the distribution of the random vector or process x of a multivariate Gaussian, all its cumulants of order three and higher are identically zero.

Thus higher-order cumulants measure the departure of a random vector from a Gaussian random vector with an identical mean vector and covariance matrix. This property is highly useful, making it possible to use cumulants for extracting the non-Gaussian part of a signal. Moments and cumulants have symmetry properties that can be exploited to reduce the computational load in estimating them. A drawback in utilizing higher-order statistics is that reliable estimation of higher-order moments and cumulants requires much more sample than for second-order statistics. Another drawback is that higher-order statistics (i.e. such as kurtosis) can be very sensitive to outliers in the data. For example, a few data samples having the highest absolute values may largely determine the value of kurtosis. Which means that the single value makes kurtosis large, thus we see that the value of kurtosis may depend on only a few observation in the tails of the distribution, that may be erroneous irrelevant observation.

Note that if the $x_i$ are independent, all the cumulants with at least two different indices are zero. The $i, j^{th}$ element of the matrix given by the transformation, say $\mathrm{F}_{ij}$, is defined as

$$\mathbf{F}_{ij}(\mathbf{M}) = \sum_{kl} m_{kl}\ cum(x_i, x_j, x_k, x_l) \tag{2.14}$$

where $m_{kl}$ are the elements in the matrix $\mathbf{M}$ that is transformed. As any symmetric linear operator, the cumulant tensor has an eigenvalue decomposition (EVD). An

eigne-matrix of the tensor is, by definition, a matrix $\mathbf{M}$ such that

$$\mathbf{F}(\mathbf{M}) = \lambda \mathbf{M} \tag{2.15}$$

where $\lambda$ is a scalar eigenvalue. The cumulant tensor is a symmetric linear operator, since in the expression $cum(x_i, x_j, x_k, x_l)$, the order of the variables makes no difference. Therefore, the tensor has an eigenvalue decomposition. Let us consider the case where the data follows the ICA model, with whitened data:

$$\mathbf{z} = \mathbf{VAs} = \mathbf{W}^T\mathbf{s} \tag{2.16}$$

where $\mathbf{W}^T$ denotes the transposed whitened mixing matrix. This is because it is orthogonal, and thus it is the transpose of the separating matrix $\mathbf{W}$ for whitened data.

Joint approximate diagonalization of eigen-matrices (JADE) refers to the principle of solving the problem of equal eigenvalues of the cumulant tensor. Eigenvalue decomposition can be viewed as digonalization. The matrix $\mathbf{W}$ diagonalizes $\mathbf{F}(\mathbf{M})$ for any $\mathbf{M}$. In other words, $\mathbf{WF(M)W}_T$ is diagonal. This is because matrix $\mathbf{F}$ is a linear combination of terms of the form $\mathbf{w}_i\mathbf{w}_i^T$, assuming that the ICA model holds. Thus, we could take a set of different matrices $\mathbf{M}_i$, $i = 1, \ldots, k$, and try to make the matrices $\mathbf{WF}(\mathbf{M}_i)\mathbf{W}$ as diagonal as possible. The diagonality of a matrix $\mathbf{Q} = \mathbf{WF(M)W}_T$ can be measured, as the sum of the squares of off-diagonal elements: $\sum_{k \neq l} q_{kl}^2$. Equivalently, since an orthogonal matrix $\mathbf{W}$ does not change the total sum of squares of the elements in a matrix, minimization of the sum of squares of off-diagonal elements is equivalent to the maximization of the sum of square of diagonal elements. Thus, we can formulate:

$$J_{JADE} = (\mathbf{W}) = \sum_i ||diag(\mathbf{WF}(\mathbf{M}_i)\mathbf{W}^T)||^2 \tag{2.17}$$

where $||diag(.)||^2$ denotes the sum of squares of the diagonal. Maximization of $J_{JADE}$ is then one method of joint approximate diagonalization of the $\mathbf{F}(\mathbf{M}_i)$.

The EEG signals recorded from the scalp are contaminated by system noise, artifact from human internal signals, and adjacent electrode signals. In order to localize the abnormalities from an EEG reading, the signals should be captured and preprocessed and finally the sources detected and localized. In preprocessing a band pass filter between 0.5 to 40 Hz (i.e. 0.5-4 Hz Delta band, 4-8 Hz Theta band 8-13 Hz Alpha band, and 13-40 Beta band) is used to eliminate the system noise and other human internal signals. For EEG, multiplying the input data matrix by the 'Separating' matrix at the end of ICA training gives a new matrix whose rows, called the component activations, are the time courses of relative strengths or activity levels of the respective components onto each of the sensors. Four main assumptions underlie ICA decomposition of EEG time series:

1. Within the EEG frequency range the source signals are only attenuated by various head tissues; there is no time-delay involved.

2. Spatial projections of components are time invariant.

3. Source activations are statistically independent.

4. Statistical distributions of the sources within each EEG subband are not Gaussian.

As described in the section 2.2.1 titled "Negentropy", a measure of non-Gaussianity and an important concept in finding out how successful the ICA algorithm, is known as negentropy [86, 87]. It is known that the distribution with the highest entropy for the same variance, is Gaussian. The randomness of a signal can then be measured by its entropy.

## 2.3 Brain Source Localization Methods

Many anatomical, functional, pathological, and physiological abnormalities in the brain can be diagnosed using fMRI, PET and MEG/EEG recordings. The fMRI, and PET can only detect the haemodynamic changes in the brain and MEG/EEG detect the electric or electromagnetic signals of the brain. fMRI and PET are not able to reveal the transient activity of the brain sources due to their poor temporal resolution, and therefore the localization based on fMRI or PET may not give accurate solution to localization of the sources.

In this work methods described are based on electromagnetic signals in order to be consistent, and can be applied to both EEG and MEG data, but only EEG data will be used to illustrate the results.

Localization of EEG sources within the brain relies on information obtained at the sensors, and can be classified into two main categories:

- Exploiting physiological aspects of the EEGs
- Norm of subspace head volume

**Exploiting physiological aspects of the EEGs**

The brain consists of left and right hemispheres, each of which are divided into four lobes: frontal, parietal, temporal and occipital.

The EEG signals including normal and abnormal rhythms within the frequency range of 0.3 to more than 30 Hz within five main subbands, as described in Chapter 1, indicate the activities of various part of the body. These signals are dissipated from particular locations of the brain lobe, representing activities and behavior of the person. Electroencephalographers, or brain specialists can detect some of the

abnormalities occurring in the signal recorded from the EEG readings. Some other problem may be diagnosed by checking each frequency band. This is carried out by carefully filtering the EEG recordings according to the clinical expectation stem from the disease symptoms. In many cases the abnormalities cause some focal activities which are localized over the scalp topograph. However a three dimensional location of the source signal requires further processing of the data.

### Norm of subspace head volume

Norm of subspace head volume is based on geometrical distance measure. The distance measures between two $n$-dimensional vector $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_n(t)]^T$ and $\mathbf{y}(t) = [y_1(t), y_2(t), \ldots, y_n(t)]^T$ (i.e vectors of sensors and sources). To measure a model and to optimize the parameters of that model an estimation of the goodness of fit between measured and predicted data can be generated by $L_P$ norm [1], which can be expressed as:

$$d(x, y) = \sum_{i=1}^{n} (|x_i - y_i|^P)^{\frac{1}{P}} \tag{2.18}$$

The mostly used $L_P$ norms on as explained in section 2.2.2, are

- P=1: The L1,"absolute value" norm. The L1-norm minimizes the sum of the activity and it accounts better due to the fact that the measured data grows with the amount of individual activity and not with the square of it.

$$d(x, y) = \sum_{i=1}^{n} |x_i - y_i| \tag{2.19}$$

- P=2: The L2,"Euclidean", "least squares" norm, which is simply Root mean square "RMS level" of the distance. This method can be used for a smooth, coherent reconstruction of gradient of the distribution. The smoothed $L_2$ norm

represents a linear inverse solution with respect to the reduction of computation time. The solution to the localization of sources in low resolution electromagnetic tomography (LORETA) [52] [53] [128] is based on quadratic norm ($L_2$ norm). This method is called the minimum norm least-squares (MNLS) estimator.

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(|x_i - y_i|^2)} \tag{2.20}$$

- P=$\infty$: The L-infinity, "Chebyshev", "minimax" norm is a limiting case which becomes

$$d(x,y) = \max_{i=1,2,\ldots,n} |x_i - y_i| \tag{2.21}$$

In previous chapter the methods to model the sources in the brain has been briefly described, here those methods are used to localize the sources. These methods can be classified into two main categories:

- Equivalent current dipole
- Linear distributed method

### 2.3.1 Equivalent current dipole model

The EEG/MEG inverse problem is ill-posed. This means there is no unique solution to the problem. Therefore it requires a priori assumption regarding the source model to achieve a unique solution. The separation of EEG signal into multiple components according to temporal characteristics is performed using dipole fitting algorithm such as R-MUSIC [63], RAP-MUSIC [118], which use the concept of equivalent current dipole method (ECD). In ECD approach the EEG signals are considered as a relatively small number of focal sources, each of which can be modelled as a single fixed or moving dipole [81]. This helps to solve the inverse problem by considering

fewer unknown parameters than the independent measurements available. Further limitations of ECD model is that they may prove to be inefficient in accounting for the synchronous activation of broader cortical areas, therefore assumption is to consider the sources as temporally independent [62], and correlations of synchronized sources including prefect synchronization [64], hence the problem can be solved by using projection concept.

**Using projection concept**

Multiple signal classification [113], works due to the fact that signal vectors are orthogonal to noise subspace

$$\mathbf{s}_i^H \mathbf{v}_j = 0, i = 1, 2, \ldots, p, j = p+1, p+2, ..., M \tag{2.22}$$

where $\mathbf{s}$ is the signals subspace vector, $\mathbf{v}$ is the vector of noise subspace, and $(.)^H$ is the hermitian transpose.

The MUSIC [113] or recursive MUSIC (R-MUSIC) [63] algorithms are methods to locate dipolar sources from EEG recordings. The algorithms scan and extract the locations of the sources through a recursive search of the head volume for multiple local peaks in the projection metric, use subspace projections within a three-dimensional head volume, and computes projections onto an estimated signal subspace.

In ECD methods, the solution depends on the assumed number of dipoles, but in general, the actual number of ECD can not be known. Thus, they produce large localization errors.

### 2.3.2 Linear distributed method

In all equivalent current dipole based methods, the solution depends only on the number of dipoles assumed but in reality, the actual number of dipoles can not be known (as *a priori*). The implementation of few equivalent dipole sources may be

appropriate in modelling abnormalities such as focal epileptic foci or any other focal activities. However, in general, most dipole methods require correct estimation of the number of sources [81]. Another approach for source localization is to consider a priori all possible fixed source locations. This approach is called distributed source model or linear distributed (LD) method. By incorporating several aspects of functional neuroanatomy such as the location of active areas, the time course of their activities, and the nature of their interactions, the probability of estimating the locations more accurately is maximized. Because of the ill-posed nature of the inverse problem some constraints are required to ensure the likelihood of the data or the log posterior of the conditional estimators has a unique maximum. A good solution is one that jointly maximizes the likelihood of the data while minimizing a cost function of the constraints. The cost function can be made up of some prior information such as that the "best" estimate maximizes the log posterior (i.e., the most likely estimate given the data). When the constraints are chosen, a unique solution is obtained through minimizing the deviation from these constraints. However to establish uniqueness and stable condition often the constraint must be applied. Some of the constraint methods are listed below:

- 1. Weighted minimum norm [34].
- 2. sLORETA-FOCUSS [54].
- 3. Minimum variance spatial filtering [116].

Each of which are described in following sections.

### 2.3.2.1. Weighted minimum norm (WMN)

The linear instantaneous EEG signal can be modelled as shown in equation (??). The weighted minimum norm is the weights that normalizes the mixing matrix $\mathbf{A}$ in the $\ell_p$-norm sense. A minimum norm estimate of sources $\mathbf{s}$ can be calculated as a solution of the optimization problem of.

$$min\|\mathbf{s}\|. \tag{2.23}$$

The source estimation basis on measured sensor mixed data is an ill-posed problem. A regularization can be implemented in order to reduce the variance of the estimated result by allowing a slight bias. This can be achieved by Tikhonov or Wiener regularization [82] and the singular value decomposition methods.
Assume the singular value decomposition of:

$$\mathbf{A} = \mathbf{U}\Lambda\ \mathbf{V}^T \tag{2.24}$$

where $\mathbf{A}$ is the mixing matrix from equation (1.1), $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\Lambda$ is a diagonal matrix with their element values in descending order. By applying this to the mixing model we have:

$$\Lambda_n\mathbf{V}^T\mathbf{s} = \mathbf{U}_n^T\mathbf{x}, \tag{2.25}$$

where $\Lambda_n$ is the $n$ rows of $\Lambda$ and $U_n$ includes the $n$ first column of $\mathbf{U}$. In this equation the index $n$, is the regularization factor for optimization version to the problem in equation(2.23). For small $n$, greater mismatch between $\mathbf{x}$ and $\mathbf{As}$ can satisfy the constraint in equation (2.25).
The minimum source estimate minimizes the sum of the absolute values ($\ell_1$-norm) [83], leading to more focal source estimation rather than Euclidean ($\ell_2$) norm. An iterative method to calculate the norms with order between 1 and 2 could be used for estimates with properties between the $\ell_1$ and $\ell_2$.

The minimum $\ell_2$ -norm estimate can be considered as the maximum *a posteriori* probability estimate with Gaussian *a priori* source distribution, while the minimum $\ell_1$-norm estimate relates to an exponential *a priori* distribution.

The deep focal sources may be similar to superficial sources. Since sensors are more sensitive to the superficial sources, there would be errors for their estimation. This errors can be compensated by using a weighted norm [34], where the weight of each source component in vector s of equation (2.23) is proportional to the strength of the signal that a constant source in the same location produces, i.e. the Euclidean norm of a row of the matrix **H**. The weighted $\ell_1$-norm of s is:

$$\|\mathbf{s}\| = \sum_{j=1}^{N} \mathbf{w}_j \mathbf{x}_j \tag{2.26}$$

where $\mathbf{w}_j$ is the weight of signal in N number of locations and $\mathbf{s}_j$ is signal vector at source location j.

The disadvantage of the minimum $\ell_1$-norm estimate is that the optimization problem cannot be solved directly, as compared to minimum $\ell_2$-norm estimate. The orientations of the source signals assumed to be perpendicular to the cortical surface, and large number of reconstruction points helps to decrease the computational cost of the minimum $\ell_1$-norm estimate. The sparse signals are set and the minimum $\ell_2$-norm can be used to find the solution.

#### 2.3.2.2. sLORETA-FOCUSS

The low resolution brain electromagnetic tomography (LORETA) algorithm [52][53] corresponds to weighted minimum norm (WMN) method, which uses Laplacian operator into the weighting matrix in order to acquire a neurophysiologically (spatial) smooth results. An alternative to LORETA is a recently developed method called standardized LORETA (sLORETA) [54]. In this method the resolution matrix is

used to normalize a coarse WMN estimation, hence the source signal would be reconstructed in a noise free environment.

This is a technique for solving the EEG inverse problem, by combining a recursive process which takes the smooth estimate of sLORETA as initialization and then employs re-weighted minimum norm introduced by the focal underdetermined system solver (FOCUSS).

### Standardized FOCal Underdetermined System Solver (FOCUSS)

EEG and MEG in certain environments and conditions, not only produce temporal resolution in few milliseconds but also the spatial source localization can be estimated in accuracy of millimeters. The solution to estimate the distribution and source localization of sources within the brain from the electrode is an inverse ill-posed problem. External observed data at the electrode sensors are as the result of superposition of internal sources. Without any knowledge about the sources no unique solution for the internal sources can be determined. Furthermore limited availability of mathematically independent measurements, increases the underdetermined nature of the problem. This means that the number of internal sources are much greater than the number of sensors. Hence assumptions are made for solving an underdetermined problem. Models for localized energy sources are obtained from physiological evidence, which describes that the underlying activity is often limited in spatial extent, or arbitrarily distributed shaped areas.

A modelling technique to solve the underdetermined problems and localizing the energy of the sources can be referred to as tomographic reconstruction method using FOCUSS. FOCUSS can be implemented by using a forward model that a source is to be projected to a predetermined reconstruction region, in order to keep the linear relation between the sources.

The FOCUSS algorithm is a recursive linear estimation procedure, which finds lo-

calized solution by starting with a distributed estimate based on a weighted pseudo-inverse solution. The weights at each step are derived from the solution of the previous iterative step. The algorithm converges to a source distribution in which the number of sensors should not exceed the number of sources. The initialization determines to which of these localized solutions the algorithm converges. Finally, only a small number of winning elements remain non-zero, yielding the desired localized sources.

There has been some research, which claims that using methods such as adaptive LORETA/FOCUSS(ALF) [55] and standardized shrinking LORETA-FOCUSS (SSLOFO) [58] and some other related papers [57] [56] a number of various resolutions using different mesh intensities can be combined to achieve the localization of the sources with less computational complexity. The dipolar sources are presented using the approximate Laplace method [56].

FOCUSS [55] is a high resolution iterative WMN method that uses the information from the previous iterations. The FOCUSS algorithm can be summarized to the following steps [49]:

$$min\|\mathbf{C}\ \mathbf{S}\|_2^2,\ subjec\ to\ \mathbf{X}\ =\ \mathbf{L}\ \mathbf{S} \tag{2.27}$$

where

$$\mathbf{C} = (\mathbf{Q}^{-1})^T\mathbf{Q}^{-1} \tag{2.28}$$

and

$$\mathbf{Q}_i = \mathbf{B}\mathbf{Q}_{i-1}[diag(\mathbf{S}_{i-1}(1)\ldots\mathbf{S}_{i-1}(3n)] \tag{2.29}$$

where $\mathbf{B}$ is the diagonal matrix that compensates for deep sources ($\mathbf{B} = diag(\frac{1}{\|\mathbf{L}_1\|}\ldots\frac{1}{\|\mathbf{L}_{3n}\|})$) and $\mathbf{L}_i$ denotes the $i^{th}$ column of $\mathbf{L}$, and the solution at $i_{th}$ iteration becomes:

$$\mathbf{S}_i = \mathbf{Q}_i\mathbf{Q}_i^T\mathbf{L}^T(\mathbf{L}\mathbf{Q}_i\mathbf{Q}_i^T\mathbf{L}^T)^\dagger\mathbf{X} \tag{2.30}$$

$\mathbf{S}$ is the actual source amplitudes and $\mathbf{L}$ is an $m \times 3n$ matrix representing the forward transmission coefficients from each source to the array of sensors, and $\dagger$

denotes pseudo inverse. **L** has also been referred to as the system response kernel or the lead-field matrix. The iterations will stop when there is no significant change in the estimation. The result of FOCUSS is highly dependent on the initialization of the algorithm. In practice the algorithm converges close to the initialization point. A clever initialization of FOCUSS has been suggested to be the solution to LORETA [55].

As another option, standardized LORETA (sLORETA) achieves a unique solution to the inverse problem. It uses a different cost function, which is:

$$min[\|\mathbf{X} - \mathbf{LS}\|_2^2 + \lambda\|S\|_2^2] \tag{2.31}$$

Hence, sLORETA uses a zero-order Tikhonov-Phillips regularization [57], which provides a solution to the ill-posed inverse problems. Furthermore, there may be some more potential sources which deserve careful consideration form the re-weighted minimum norm sparse solution during iteration. The standardized shrinking LORETA-FOCUSS (SSLOFO) [50] can help eliminating nodes that do not contain any source activity, or if a node is incorrectly eliminated during iteration, to bring it back into solution space. This can be achieved by estimating sources using sLORETA, and initializing weighting matrix according to (2.28) and redefining, normalizing, and smoothing the solutions in various iterations until no changes occurs in two consecutive steps or the solution of any iteration is less sparse than the solution estimated by the previous iteration, and or the source strength of any node exceeds a threshold preset by the user. However, the accuracy of such algorithms is dependent on the number of both sources and sensors assuming fixed positions of the sources [114]. Moreover, in most of the above algorithms a head volume conductor model and a source model have to be defined. Therefore the associated computational complexity is generally very high.

### 2.3.2.3. Minimum variance spatial filtering

In this algorithm a localization method based on spatial filtering principle called linearly constrained minimum variance (LCMV) filtering, is exploited. These spatial filtering are designed to pass brain electrical activity from a specific location while attenuating activity originating at other locations.

The output power at a designed multiple spatial filter is an estimate of neural power activities from the spatial pass-band of the filter. The potential measured as the observed data vector $\mathbf{x}$ can be defined as:

$$\mathbf{x} = \sum_{i=1}^{L} \mathbf{H}(q_i)\mathbf{m}(q_i) + \mathbf{n} \tag{2.32}$$

where vector $\mathbf{m}(q_i)$ are the x, y, z components of the dipole moment $q_i$, and n is the measurement noise. $\mathbf{H}(q)$ represents the material and geometrical properties of the medium in which the sources are submerged.

By constructing three spatial filters for each location, one for each component for the dipole moment, the three component filter output $\mathbf{y} = \mathbf{m}(q_0)$ is the inner product of $\mathbf{W}(q_0)$ and $\mathbf{x}$:

$$\mathbf{y} = \mathbf{W}^T(q_0)\mathbf{x} \tag{2.33}$$

where $\mathbf{W}(q_0)$ is the spatial filter for narrowband volume element $Q_0$ centered at location $q_0$.

The LCMV can be expressed as:

$$\min_{\mathbf{W}(q_0)} tr\mathbf{C}(\mathbf{y}) \tag{2.34}$$

subject to orthogonality factor $\mathbf{W}^T(q_0)\mathbf{H}(q_0) = \mathbf{I}$ where $\mathbf{C}(\mathbf{y})$ is the covariance matrix of the dipole moment. Substituting for $\mathbf{C}(\mathbf{y})$:

$$\begin{aligned} &\min_{\mathbf{W}(q_0)} tr[\mathbf{W}^T(q_0)\mathbf{C}(\mathbf{x})\mathbf{W}(q_0)] \\ &subject\ to\ \ \mathbf{W}^T(q_0)\mathbf{H}(q_0) = \mathbf{I} \end{aligned} \tag{2.35}$$

The constraint in equation (2.35) can be incorporated into the main cost function to achieve an unconstrained Lagrangian $L(\mathbf{W}, \mathbf{L})$ as:

$$L(\mathbf{W}, \mathbf{L}) = tr\{\mathbf{W}^T\mathbf{C}\mathbf{W} + (\mathbf{W}^T\mathbf{H} - \mathbf{I})2\mathbf{L}\} \tag{2.36}$$

where $2\mathbf{L}$ is the lagrange multiplier, $(q_0)$, and $(\mathbf{x})$ are omitted for clarity. Denoting that $tr\mathbf{B} = tr\mathbf{B}^T$ for any square matrix B:

$$L(\mathbf{W}, \mathbf{L}) = tr\{\mathbf{W}^T\mathbf{C}\mathbf{W} + (\mathbf{W}^T\mathbf{H} - \mathbf{I})\mathbf{L} + \mathbf{L}^T(\mathbf{H}^T\mathbf{W} - \mathbf{I}\} \tag{2.37}$$

The matrix $\mathbf{C}$ is positive definite so the minimum of $L(\mathbf{W}, \mathbf{L})$ is attained by setting the first term to zero,

$$\mathbf{W} = -\mathbf{C}^{-1}\mathbf{H}\mathbf{L} \tag{2.38}$$

By substituting $\mathbf{W}$ in the constraint $\mathbf{W}^T\mathbf{H} = \mathbf{I}$, Lagrange multiplier $\mathbf{L}$ is now obtained

$$-\mathbf{L}^T\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H} = \mathbf{I} \tag{2.39}$$

or

$$\mathbf{L}^T = -(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1} \tag{2.40}$$

Substituting (2.40) into (2.38)

$$\mathbf{W}(\mathbf{q}_0) = [\mathbf{H}^T(\mathbf{q}_0)\mathbf{c}^{-1}(\mathbf{x})\mathbf{H}(\mathbf{q}_0)]^{-1}\mathbf{H}^T(\mathbf{q}_0)\mathbf{C}^{-1}(\mathbf{x}) \tag{2.41}$$

Using (2.41 in (2.33) gives an estimate of the moment at location $\mathbf{q}_0$. The estimated variance of the activity can be defined as:

$$\widehat{Var}(\mathbf{q}_0) = tr\{\mathbf{H}^T(\mathbf{q}_0)\mathbf{C}^{-1}(\mathbf{x})\mathbf{H}(\mathbf{q}_0)]^{-1}\} \tag{2.42}$$

To preform localization, the estimated variance as a function of location within the volume of brain can be accomplished by evaluating (2.42).

As the conclusion, the constrained methods described above are spatial inverse methods. These are not influenced by the temporal course of the sources, however, by processing only a single time sample, such methods usually have higher localization errors on independent sources in the presence of noise. On the other hand considering only spatial inverse methods may also give error in the location of the sources when continuous EEG recordings are processed. Implementing high resolution methods such as $\ell_1$-norm [65] and FOCUSS, may generate discontinuous signals because of their nonlinear characteristics. Low-resolution methods implementation such as WMN, LORETA, sLOERTA usually mix the source signals within their wide spread functions.

## 2.4 Finite Element Method

In nonhomogeneous head model the numerical solutions are often required for solving the localization problem as the signals propagate through different layers of head. The surface boundaries for each layer must be extracted from MRI to perform boundary element method (BEM) computation. BEM model consider the conductivity is isotropic and homogenous within each volume separated by the surfaces. In contrast, the finite element method (FEM) can model the anisotropy and inhomogeneity of the conductivity, which are known to be significant for layers of head tissues [35].

In the fusion techniques a source localization method involves a priori assumptions about sources and their locations constructed in a patient- specific finite element model (FEM) [70], [71]. Electric current dipoles are usually used as sources, provided that the regions of activations are relatively focused [28].

To employ the finite element method, the object space has to be segmented into finite elements. Inside each element a set-up function has to be chosen, according to

the model of interests. The resulting shape functions span a function space which includes the approximation of the solution. The grade of shape functions, the shape, the size, and the amount of used elements can vary widely which yields to a high flexibility.

Neuronal processes, such as cognition , coordination, perception are carried out via the propagation of electrical impulses through the brain. These impulses give rise to electromagnetic fields that can be measured extracranially by sensitive recording devices. These realistic model forward solutions can be generated using the boundary element method (BEM) or FEM. Both methods can represent the complex boundaries and the inhomogeneous regions of realistic models. The FEM has the additional advantage that it can capture an isotropic conductivities of the domain. The main idea behind the FEM is to reduce a continues problem with infinitely many unknown field values to a finite number of unknowns by discretizing the solution region into elements. The value at any point in the field can then be approximated by interpolation functions within the elements. These interpolation functions are specified in terms of the field values at the corners of the elements, points known as nodes. It can be noted that for linear interpolation potentials, the electric field is constant within an element.

The finite element analysis requires a defined physical property to present an actual representation of the brain and head model in a forward problem. This can not be an accurate representation in an inverse problem.

## 2.5 Other Modalities for Detection of Brain Sources

Functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) are used for detection of the locations of neural activity.

- Hemoglobin becomes demagnetized when oxygenated, and magnetized when

deoxygenated, This causes change in the magnetic resonance (MR) signal of blood depending on the level of oxygenation. These differential signals can be detected using an appropriate MR pulse sequence as Blood Oxygenation Level Dependent (BOLD) contrast.

- A positron is an anti-electron, and they are given off during the decay of the nuclei of specific radioisotopes. A radioactive fluorine is a positron emitter. When matter collides with its corresponding antimatter, both are annihilated. When a positron meets an electron, the collision produces two gamma rays having the same energy, but going in opposite directions. The gamma rays leave the patient's body and are detected by PET scanner.

A multi-modal technique can incorporate a combined EEG/fMRI or EEG/PET to explain both high temporal resolution of EEG and high spatial resolution of fMRI or PET.

# Chapter 3

# TOPOGRAPHIC LOCALIZATION OF BRAIN SIGNAL SOURCES

## 3.1 Separation

Independent component analysis (ICA) refers to a family of related algorithms [76, 79, 84] that exploit independence to perform blind source separation. EEG contains information about the brain and its functions.

The block diagram, shown in Figure 3.1 demonstrates the procedure for localization of the EEG sources based on back projection. Since the EEG's are the result of combining a number of signals generated by the neurons, it would be difficult to relate each mixture to a particular brain's functionality. The separation of independent sources from a superposition of mixed EEG signal is accomplished by taking into account the structure of the mixing process and by making assumptions about the sources. The major factors required for consideration would be the geometrical

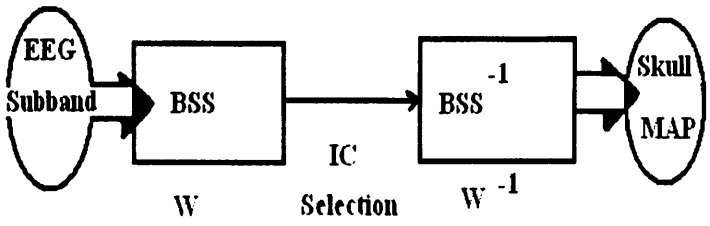



Figure 3.1: Topographic localization based on back projection

location of the sources within the brain and also the existing clinical information is some distinct frequency band. Since various brain disorders manifest themselves as abnormal rhythms in certain frequency band, a multi-resolution blind source separation (MRBSS) is adequate to efficiently discriminate the actual signal sources. The spherical head model is a simple reasonable model for calculating the location of EEG sources [63]. Although the realistic head model approximates the head as a set of homogeneous layers, typically the model comprises of three layers of scalp, the skull, and the brain [62], it causes localization of sources in the head to be inaccurate, due to different conductivity of the layers. The standard conductivity values for brain, skull and scalp are approximately 0.33, 0.0042 , $0.33(\Omega^{-1})$ respectively. Further investigation has shown that the skull inhomogeneities causes source location error [93]. The inhomogeneities of the head layers will be discussed in next chapter.

The separation process will not be effective unless the independent components are identical to the original sources. A source signal may be statistically independent from other signals but the number of statistically independent components can be limitless. The separation should be based on desired features linked to optimal number of sources and frame size that would separate the feature of interest.

The steps taken for topographical localization of the sources over the scalp are:

- **EEG recordings in Sub-bands**: Load the EEG signals and sub-divide them into related sub-bands (i.e. Delta, Theta, Alpha, Beta, Gamma) using bandpass, filtering to capture diagnostic features specified by the physician for certain abnormalities in the brain. The signal spectrum is often estimated through autoregressive moving average (ARMA) modelling.
- **BSS**: Apply a BSS algorithm such as JADE, SOBI, or any other suitable algorithm to separate the EEG mixture recordings into a number of statistically independent sources (i.e. IC's), normalizing the result and sort them with respect to their energy in a descending order.
- **Eliminate ICs, invert unmixnig matrix, and Back project** : Select one IC at the time, and eliminate the rest. Find the inverse of un-mixing matrix resulting from the selected IC. Back project the single IC to the topographic image using interpolation method.
- **Identify the dominant ICs**: Repeat the process of elimination and back projection for all of the ICs, and observe the location of the most dominant ICs where they are repeating for an specific location on the projection over the scalp.

### 3.1.1 Separation of the EEG source signals using iterative SOBI algorithm

In BSS $n$ number of source signals, normally comes down to identifying the mixing matrix $\mathbf{A}$ such that one can find a matrix $\mathbf{W}$ so that the matrix format of this model is shown in equations (1.1) and (1.2). $\mathbf{A}$ and $\mathbf{W}$ are respectively mixing and unmixing matrices.

Assuming the sources have unit variance, for independent sources

$$\mathbf{R_s}(0) \triangleq \mathbf{E}[\mathbf{s}(t).\mathbf{s}^*(t)] \tag{3.1}$$

where $(\cdot)^*$ denotes the conjugate operator and $(\cdot)^H$ denotes the hermitian operator. The covariance matrix for independent sources can be formulated as followings:

$$\mathbf{R_s}(\tau) = \boldsymbol{E}[s(t+\tau).s^*(t)] = \begin{bmatrix} \boldsymbol{E}[s_1(t+\tau).s_1^*(t)] & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \boldsymbol{E}[s_n(t+\tau).s_n^*(t)] \end{bmatrix} \tag{3.2}$$

The covariance matrix of the observation matrix:

$$\mathbf{R_x}(0) = \boldsymbol{E}[\mathbf{x}(t).\mathbf{x}^*(t)] = \mathbf{A}\mathbf{R_s}(0)\mathbf{A}^H + \sigma^2\mathbf{I} \tag{3.3}$$

$$\mathbf{R_x}(\tau) = \boldsymbol{E}[\mathbf{x}(t+\tau).\mathbf{x}^*(t)] = \mathbf{A}\mathbf{R_s}(\tau).\mathbf{A}^H,\ \tau \neq 0 \tag{3.4}$$

For the sources with unit variance

$$\mathbf{R_x}(0) = \boldsymbol{E}[\mathbf{x}(t).\mathbf{x}^*(t)] = \mathbf{A}\mathbf{I}\mathbf{A}^H + \sigma^2\mathbf{I} \tag{3.5}$$

Also note that $\mathbf{R_s}(\tau)$ is a diagonal matrix. The first step of separation process is the whitening step. This is done by applying a whitening matrix $\mathbf{W}$ to the output $\mathbf{y}(t)$:

$$\boldsymbol{E}\{\mathbf{W}\mathbf{y}(t)\mathbf{y}^*(t)\mathbf{W}^*\} = \mathbf{W}\mathbf{R_y}(0)\mathbf{W}^H = \mathbf{W}\mathbf{A}\mathbf{A}^H\mathbf{W}^H = \mathbf{I} \tag{3.6}$$

The above equation shows that if $\mathbf{W}$ is a whitening matrix, then

$$\mathbf{WA} = \mathbf{U} \tag{3.7}$$

where $\mathbf{U}$ is a $n \times n$ unitary matrix. Now the matrix $\mathbf{A}$ can be factorized as:

$$\mathbf{A} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{U} = \mathbf{W}^\dagger\mathbf{U} \tag{3.8}$$

where $\dagger$ denotes the Moore-Penrose pseudoinverse [40]. From equations (3.1) and (3.3) it is true that $\mathbf{A}\mathbf{A}^H = \sigma^2\mathbf{I}$. Also from equation (3.4) the whitening matrix $\mathbf{W}$ can be calculated from the covariance matrix of the observed signals, and the separation process is described as

$$\mathbf{z}(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{W}(\mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)) = \mathbf{U}\mathbf{s}(t) + \mathbf{W}\mathbf{n}(t) \tag{3.9}$$

The covariance matrix of z(t) is now

$$\mathbf{R}_z(\tau) = \mathbf{W}\mathbf{R}_\mathbf{x}(\tau)\mathbf{W} \tag{3.10}$$

From equation (3.3) and equation (3.8) we have:

$$\mathbf{R}_z(\tau) = \mathbf{U}\mathbf{R}_\mathbf{s}(\tau)\mathbf{U} \quad \forall_\tau \neq 0. \tag{3.11}$$

The joint diagonalization of several matrices is a way to define the "average eigen-structure" shared be the matrices [74]. The separation matrix can be found by finding the global minimum of a cost function $J(\mathbf{W})$, which provides a measure of independence of the estimated sources. Therefore the goal of the diagonalization algorithm is to find a $\mathbf{W}$ that will make the output covariance matrix $\mathbf{R}_y$ diagonal. Hence minimizing $J(\mathbf{W})$ will ensure that the estimated sources are as independent as possible. The covariance matrix $\mathbf{R}_y$ to be diagonalized, is given by

$$\mathbf{R}_\mathbf{y} = \mathbf{W}[\mathbf{R}_\mathbf{x} + \mathbf{R}_\mathbf{n}]\mathbf{W}^T \tag{3.12}$$

where $\mathbf{R_x}$ is the estimation of the covariance matrix of signal mixtures over the current block and $\mathbf{R_n}$ is the covariance matrix of the noise. Since it has been assumed that the noise is uncorrelated $\mathbf{R_n}$ will be a diagonal matrix [96].

$$\mathbf{R_x} - \mathbf{R_n} = \mathbf{A}\mathbf{R_s}\mathbf{A}^T \tag{3.13}$$

where $\mathbf{R_s}$ is a diagonal covariance matrix of the independent source signals. The LS estimate of $\mathbf{W}$ is

$$J(\mathbf{W}) = arg \min_{\mathbf{W}} \sum_{t=1}^{T_B} \|E(t)\|_F \tag{3.14}$$

where $\|\cdot\|_F$ is the Frobenius norm and $E$(t) is the error to be minimized between the covariance of the source signals $\mathbf{R}_S$ (diagonal due to independence) and the estimated sources $\mathbf{R}_Y$ and $T_B$ is the data block length. Therefore a suitable cost function is defined that minimizes the off diagonal elements, defines as

$$J_m(\mathbf{W}) = arg \min_{\mathbf{W}} J_M(\mathbf{W}) = arg \min_{\mathbf{W}} (offdiag(\mathbf{R}_y))^2 \tag{3.15}$$

where $offdiag[\cdot]$ means off-diagonal elements of a matrix, i.e. $(\mathbf{R}_y - diag[\mathbf{R}_y])$.
It has been found that estimating covariance matrices using SOBI is easier and computationally less expensive than estimating fourth order cumulants, hence it gained speeds up in estimating the covariance matrices. Since the signals in reality are not stationary, joint diagonalization of several matrices is a way to define the "average eigenstructure" shared by the matrices. Hence SOBI [138] or as in [122], can better cope with nonstationarity of the data. SOBI can separate functionally distinct neuronal signals from each other and from other noise sources under poor signal to noise ratio (SNR). SOBI also is able to recover components that are physiologically and neuroanatomically interpretable [138] [139]. On the other hand, the signals may be considered stationary within short segments of approximately 10 seconds

(or approximately 2000 samples). SOBI has been selected as the best BSS system for separation of the EEG signals and modified so $\mathbf{W}$ can be estimated iteratively by jointly diagonalizing the cross-correlation function for a number of lags. As described in [135] $\mathbf{y}(t)$ is the output of the estimated original sources in (1.2).

## 3.2 Filtering

The paradigm of EEG recordings is the result of combining a number of signals generated by the neurons performing various tasks, it would be difficult to relate each mixture to a particular brain's functionality without any mathematical analysis. The localization of ICs from a superposition of mixed EEG signal can be accomplished by consideration of the clinical information and the geometrical locations of the sources within distinct frequency bands. Hence it is required to subdivide the EEG signals into related sub-bands (i.e. Delta, Theta, Alpha, Beta, Gamma) for that particular geometrical location of the EEG recordings. This work is based on instantaneous assumption of the EEG signals. Therefore frequency subdivision or spatial filtering techniques such as Discrete Fourier transform (DFT) method or Auto-regressive (AR) or ARMA based estimation are only considered. AR estimation of the spectrum is often preferred to DFT to avoid windowing and noise effect.

Filtering of the EEG signals using DFT based spectrum, from a patient with Jakob-Creutsfeldt disease (CJD) is shown in Figure 3.3. The EEG recordings of a normal person moving his left finger is demonstrated in Figure 3.4. Using autoregressive (AR) modelling with order of 8, the space frequency representation of the multi-channel signals after filtering are represented in Figure 3.5. The amplitude peak location is approximately at 8 Hz (i.e. Alpha band).

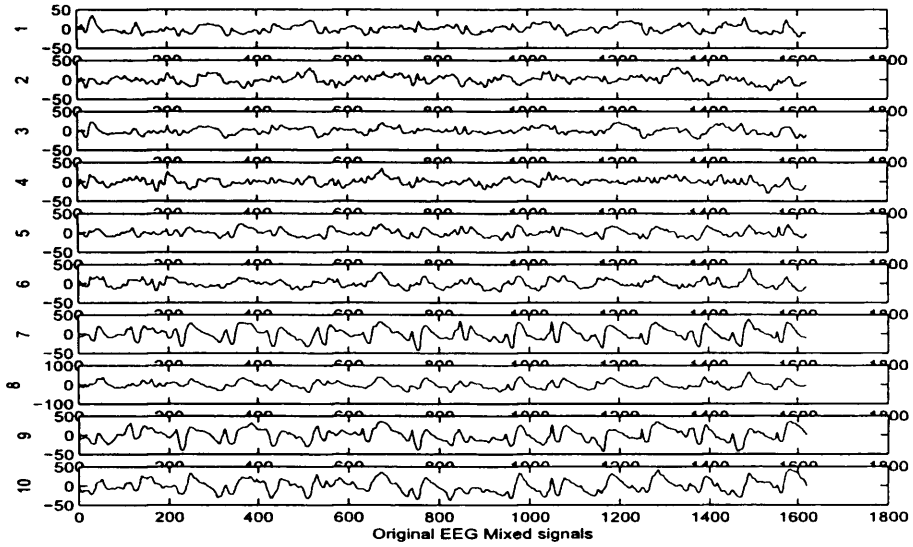



Figure 3.2: EEG signals from a patient with CJD symptoms

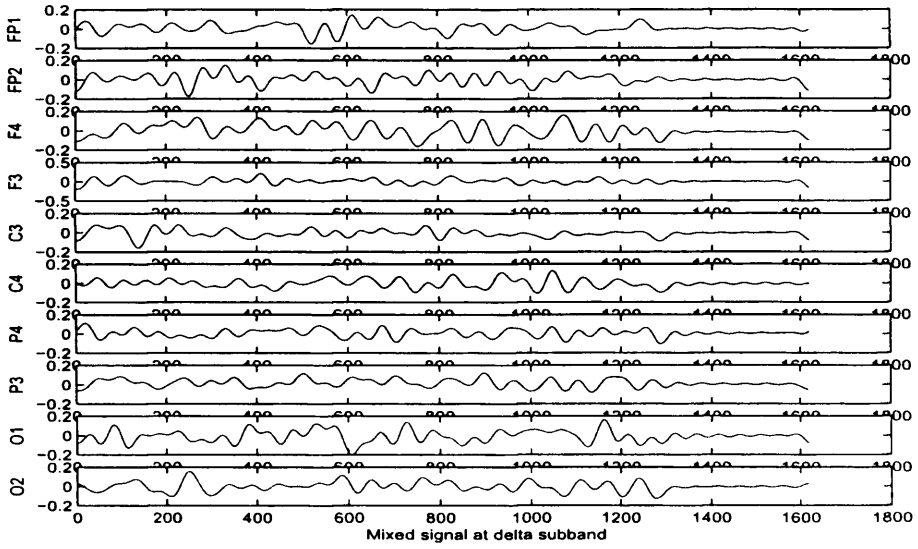



Figure 3.3: Filtered signals at Delta band, from a patient with CJD symptoms (DFT is used for transformation).

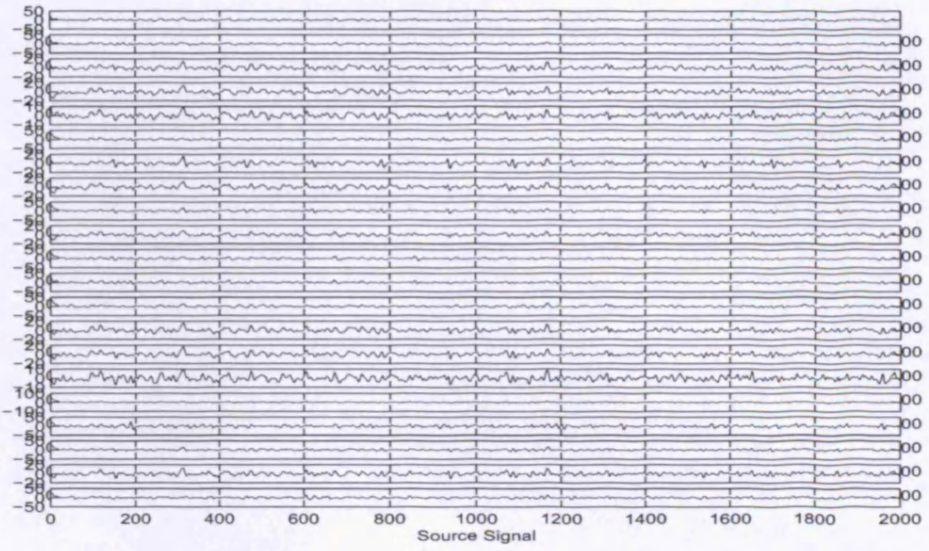



Figure 3.4: EEG recordings of a finger movement of a normal awake person.

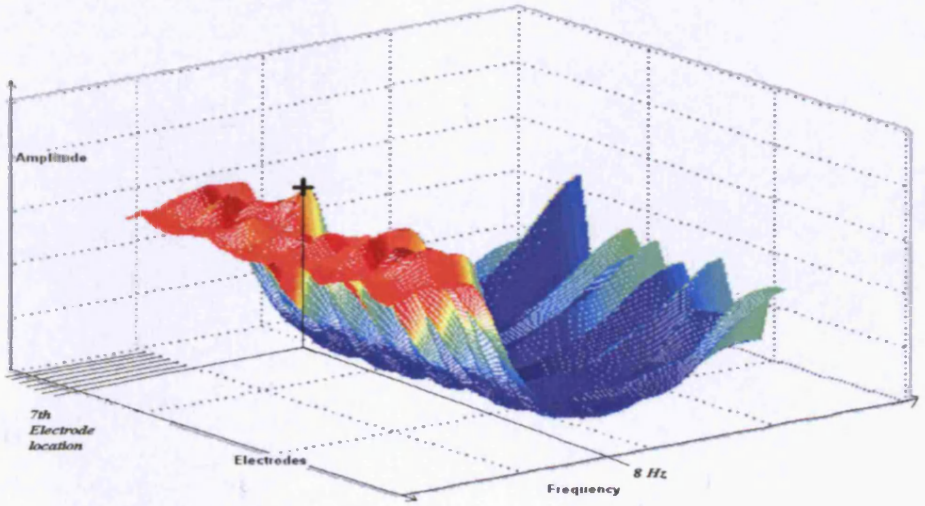



Figure 3.5: Space (electrodes)-frequency representation of filtered signals related to finger movement. AR with order of 8 is used for transformation. Hence, the highest peak amplitude at the electrode location seven in electrode axis with frequency of 8 Hz is considered as the known source.

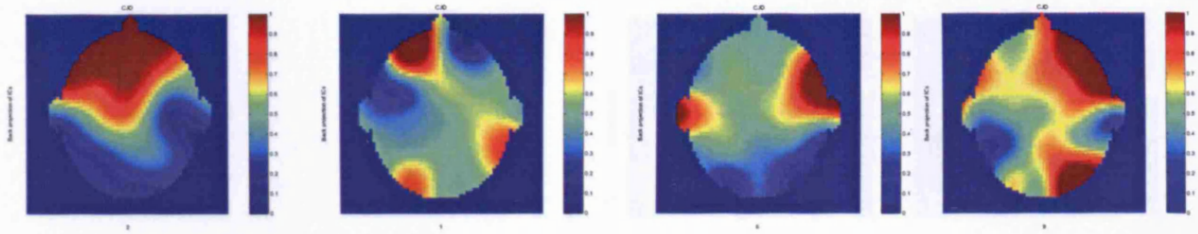
## 3.3 Localization Over the Scalp using Back Projection

A Matlab program was provided to localize a 2D geometrical location of the EEG recordings, another words to project the ICs to the scalp topograph. This program after applying BSS to the EEG recordings, selects one of the ICs, and sets the rest to zero, uses the invert of the mixing matrix (i.e. $\mathbf{W}^{-1}$), generates the mixture signal relating to the single IC, and plots the amplitudes of the selected ICs on each electrode as a pixel in an image and then interpolates to make a smooth image. Hence approximating the location of source corresponding to project the IC to the scalp topograph.

In figures 3.6(a) to 3.6(d), and 3.7(a) to 3.7(d), and 3.8(a) to 3.8(d), the dark red represents maximum and dark blue indicates the low amplitude. This can help to roughly estimate the geometrical location of the abnormal EEG activity from over the scalp without any help of an expert.

To carry out a simulation, EEG recordings of a patient with Creutzfeldt-Jakob disease (CJD) is considered. This is a spongiform encephalopathy disease that affects about one in each million inhabitants in most countries. Recently, a new variant of CJD has been linked to the epidemic of bovine spongiform encephalopathy. Therefore, vigilance concerning the disease's incidence has been increased. CJD affects the frontal electrodes specially in Delta band. The result of the simulations for a patient with CJD and back projection of ICs are shown in Figures: 3.6(a) to 3.6(d). The result indicates the sources are dissipating prominently in frontal lobe of the brain.

In the next simulation, the results of back projection of the ICs for a patient in state of drowsiness are shown in Figures: 3.7(a) to 3.7(d). The drowsiness normally dissipating from frontal and temporal lobes, and the results from back projection in



(a) (b) (c) (d)

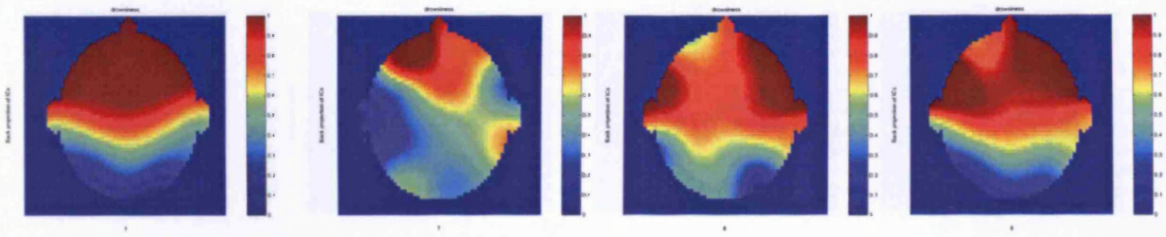
Figure 3.6: A 2D topological view of the brains ICs for a patient with CJD. The dark red are the highly active areas: (a) IC1 (b) IC2 (c) IC6 (d) IC9.

this simulation confirm this.

Further experiment has been carried out and the result of the simulations for a patient with focal seizure, with 12 EEG electrode connections to his head, and back projection of ICs are shown in Figures: 3.8(a) to 3.8(d). The focal seizures are sharp focal sources dissipating around the head. In this simple approach the results can be confirmed by clinical trial and examinations. Since the sources of most of the abnormalities are close to the cortex the scheme provides very valuable results.

## 3.4 Detection of the Number of Sources by implying Spatial and Frequency Constraints

Accurate estimation of location of the brain sources may be possible if the number of dipole sources is known *a priori*. Alternatively assumptions need to be made with regard to the number of dipole sources. Selecting the dominant eigenvalues and evaluation of the inconsistencies in their determination can be considered as an approach to identify the number of signal sources. Evaluating the eigenvalues could be sufficient for noiseless environment but localization of brain signals requires

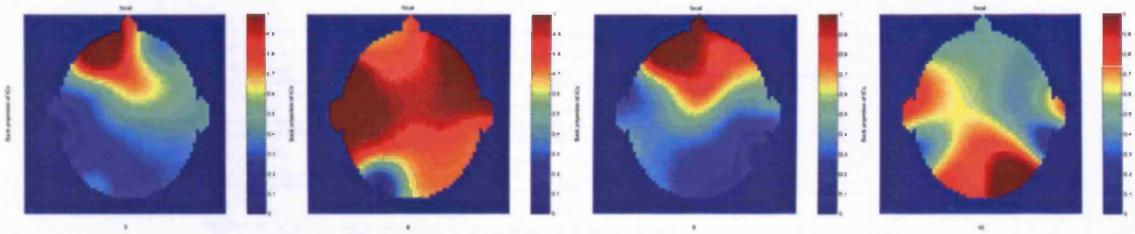


(a) (b) (c) (d)

Figure 3.7: A 2D topological view of the brains ICs for a patient in state of drowsiness. The dark red region represent high activity areas; frontal and temporal: (a) IC1 (b) IC7 (c) IC8 (d) IC9

estimation of the sources in a noisy condition. For simplification it is appropriate to mention that the deep source signals are very weak and not easy to detect. In addition the deep sources do not have considerable contribution to the distribution of either signal or noises, hence most of functional signals are generated from the sources close to cortex layer situated on the surface of the brain under the skull. Therefore it is expected to have a localized activity if the estimated sources are re-projected to the scalp one by one, and checked how peaky the topography is. The next step would be the frequency distribution constraint. The brain sources, particularly the desired sources are expected to be cyclic or semi-cyclic, hence by converting the estimated source signals into frequency domain, the peakedness of a signal in certain frequency ranges can be considered as the cycle frequency of the desired source. By a logical combination of large eigenvalues, peaky spatial topography domain, and peaky spectrum, the number of the sources, and the signal -noise subspaces can be identified.

This can be represented using a synthetic simulation of some source signals. Let us



(a) (b) (c) (d)

Figure 3.8: A 2D topological view of the brain's ICs for a patient with focal seizure. The locations of most dominant sources are (a) IC5, (b) IC8, (c) IC9, and (d) IC10.

assume there are four source signals combined together as shown in Figure 3.9.

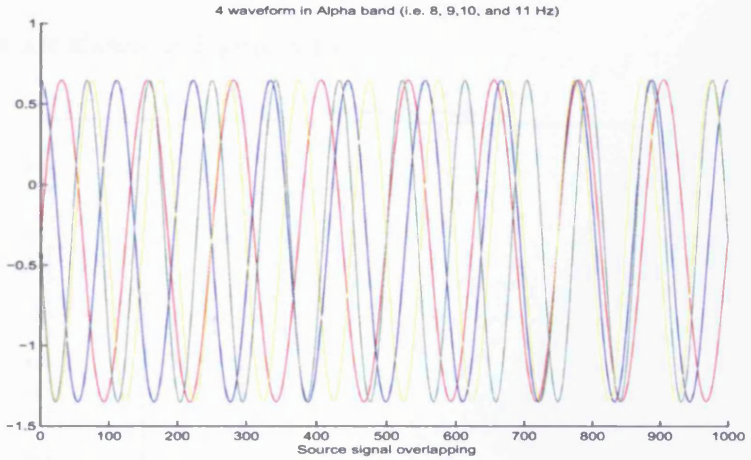



Figure 3.9: Four synthetic waveforms in 8, 9 ,10 , and 11 Hz representing the source signals in Alpha band

The signals are mixed together with four generated random noise signals as shown in Figure 3.10.

The eigenvalue of the mixture signals are shown in Figure 3.11. Considering the contribution of larger eigenvalues as the source signals, and smaller eigenvalues as the noise. The distributions of the sources are shown in Figure 3.12, and the noise

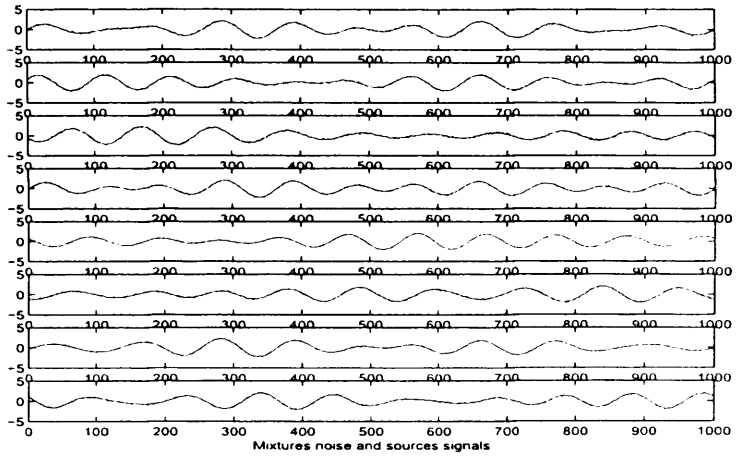



Figure 3.10: Mixtures of source and noise signals

distributions are shown in Figure 3.13.

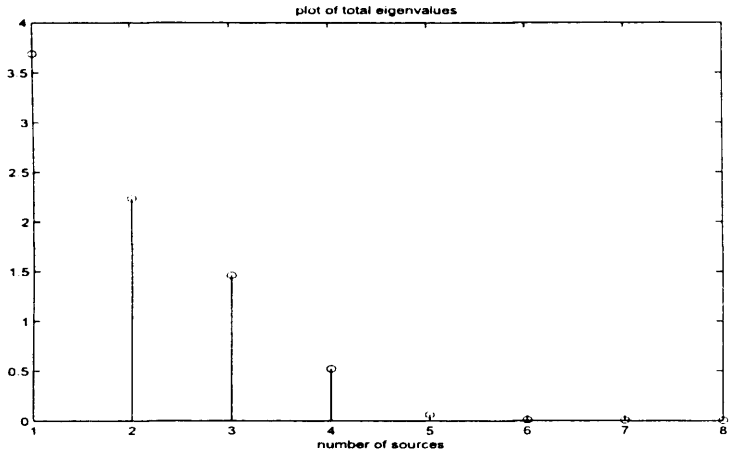



Figure 3.11: Plot of eigenvalues of the mixed signals

The distribution of source signal indicates super Gaussian, whereas the distribution of noise is Gaussian. The next step is to look into the topography of scalp and find the peak's strength. This can be considered as a spatial constraint for the estimated

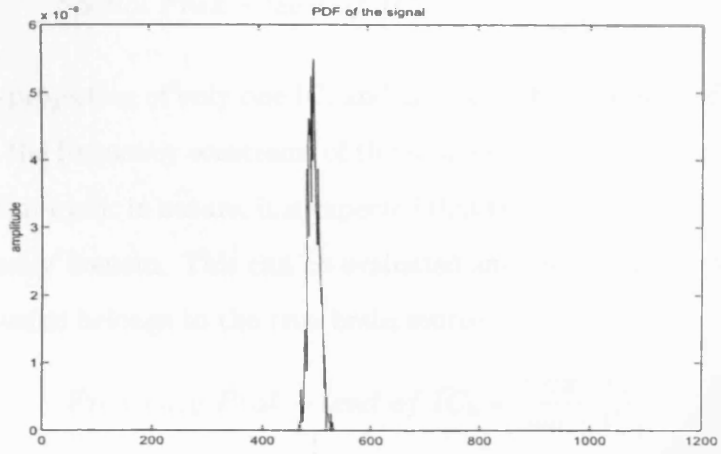



Figure 3.12: Distribution of the source signals

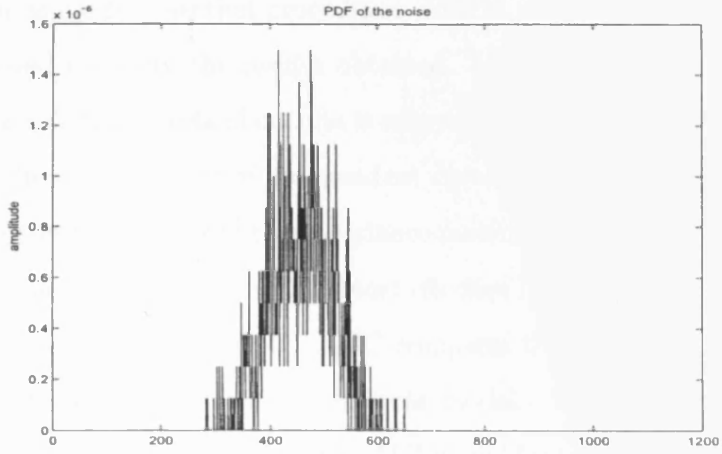



Figure 3.13: Distribution of the noise signals

source signals.

$$Spatial\ Peak - level\ of\ IC_k = \frac{max(max(\hat{X}))}{var(\hat{X})} \tag{3.16}$$

where $\hat{X}$ is re-projection of only one IC, and $var(\hat{X})$ is the variance of $\hat{X}$. The next step is to find the frequency constraint of the sources. Since the brain neural signals are cyclic or semi-cyclic in nature, it is expected that the true sources exhibit a strong peak in frequency domain. This can be evaluated and used to recheck whether the selected eigenvalue belongs to the true brain source.

$$Frequency\ Peak - level\ of\ IC_k = \frac{max(|\ Y_k|)}{var\ |\ Y_k\ |} \tag{3.17}$$

where $Y_k$ is discrete Fourier transform (DFT) or AR spectrum amplitude of the $k^{th}$ IC, and |.| denotes the absolute value.

Finally the results for real sources can be checked with a scheme for rationally estimating the number of ICs from EEG recordings. The information about the brain sources known as model selection procedures (MSPs), or goodness of fit procedures [142] can be used to verify the results obtained. MSPs are evaluated for different source and noise. The theoretical criteria is referred to as information criterion (ICr). In the ICr method the number of independent dipole sources can be determined by only analyzing the eigenvalues of the covariance matrix of the measured data, thus it avoids solving the inverse problem. The most effective ICr method described in [143] is Akaike information criterion (AIC). AIC compares the log-likelihood function of the total error to a penalty term for different model. The penalty term contains the number of parameters. Note that the AIC resembles the Likelihood ratio (LR), which evaluates if one model is more likely than another.

The ICr [42] for the determination of the number of ICs from EEG or MEG can be summarized to the following steps:

1. Calculating the covariance matrix **C** of the measured EEG recordings data matrix .

2. Using SVD to decompose the covariance matrix **C**, and get all eigenvalues such that, $\lambda_1 < \cdots < \lambda_m$, or in this simulation $\lambda_1 > \cdots > \lambda_m$.

3. Calculating the ICr value with eigenvalues of the covariance matrix **C**. The ICr can be calculated when the distribution of the noise eigenvalues is known and when it is unknown [43].

4. According to the rule of the ICr method, the number of sources with minimum information criterion ICr is selected as the estimated number of sources.

A simulation carried out proved that the correct number of sources even without *a prior* knowledge about the sources, has been estimated. Based on equation (12), and (13) from [44], the ICr generated as shown in Figure 3.14, and Figure 3.15, when the noise information is available and when the noise information is not available respectively.

### Application of the method to real EEG

The method was applied to a set of EEG recordings from a patient with focal epilepsy just before and during the seizure. Initially the dominant eigenvalues are identified, Therefore attempt was made to find the distribution of the sources and the noise signals before seizures are shown in Figures 3.16, and 3.17 respectively. The distribution of the sources and the noise signals during seizures are shown in Figures 3.18, and 3.19, respectively.

The next step is to find the spatial and frequency peaks based on the topographic projections and the frequency domain representation of the estimated sources. The result of the simulation indicate that the sources for the case of before seizure are

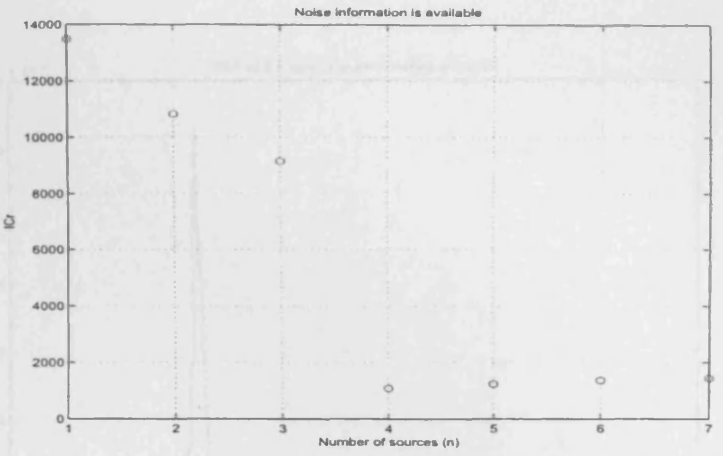



Figure 3.14: Akaike information criterion when the noise information is available. Plot of $n$ sources from $m$ mixtures against ICr, where the Criterion reaches to a minimum value.

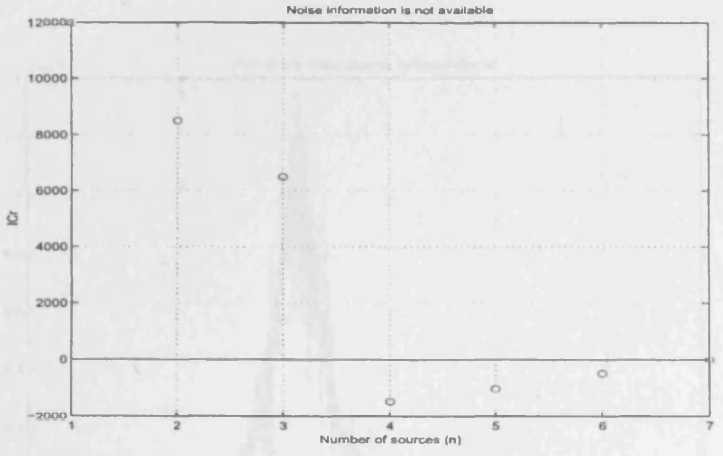



Figure 3.15: Akaike information criterion when the noise information is not available. Plot of $n$ sources from $m$ mixtures against ICr, where the Criterion reaches to a minimum value.

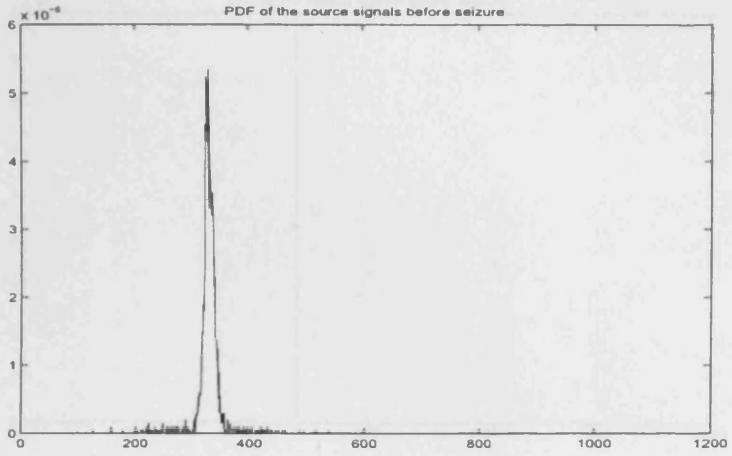



Figure 3.16: Distribution of the source signals in a patient just before focal seizure.

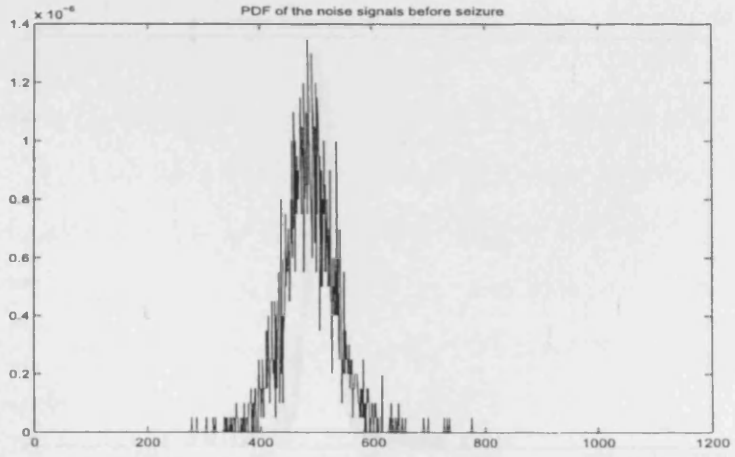



Figure 3.17: Distribution of the noise signals in a patient just before focal seizure.

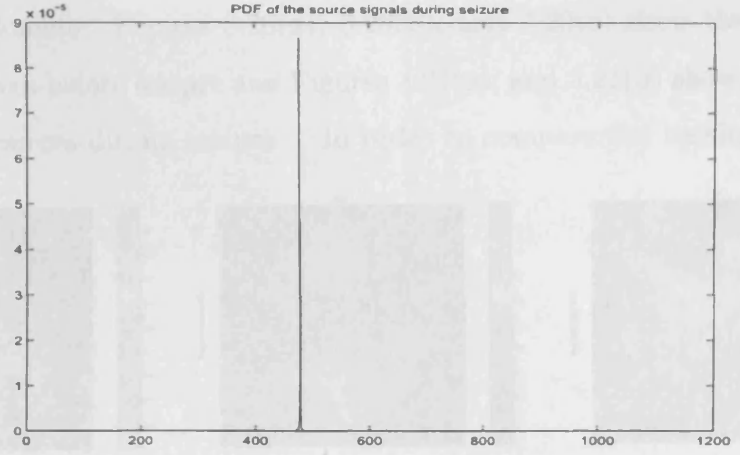



Figure 3.18: Distribution of the seizure related source signals in a patient during focal seizure.

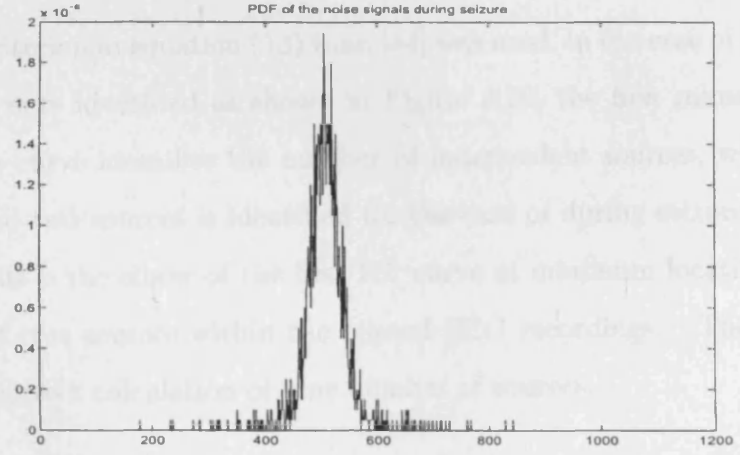
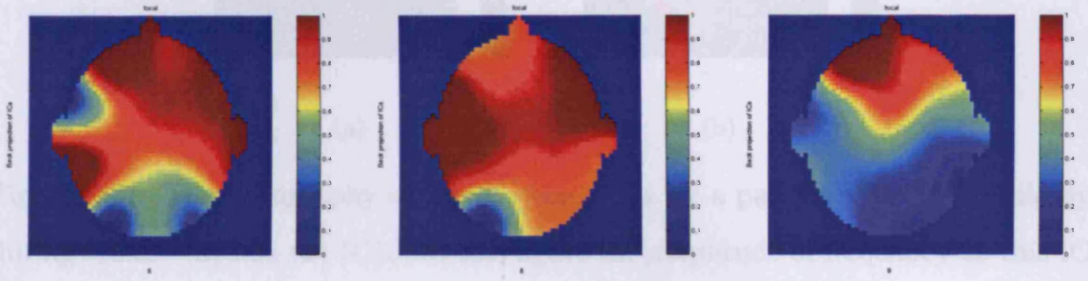



Figure 3.19: Distribution of the noise signals in a patient during focal seizure.

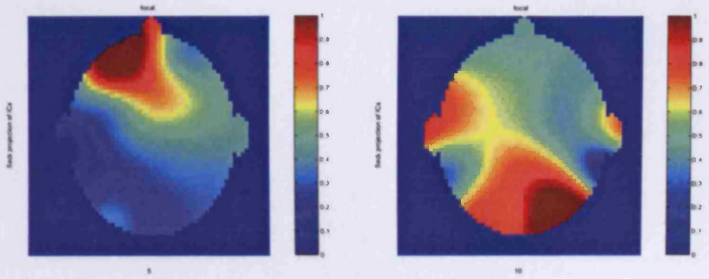
most likely to be IC8, IC9, and IC10, and for the case of during seizure are likely to be IC9, IC10, since they have the highest peaks in the tomography and strong peaks in frequency domain. Figures 3.20(a), 3.20(b), and 3.20(c) show the locations of dominant sources before seizure and Figures 3.21(a), and 3.21(b) show the locations of dominant sources during seizure. In order to compare the results, the Akaike



(a) (b) (c)

Figure 3.20: The topography of the projected ICs for a patient with focal epilepsy before seizure (a) IC8 (b) IC9 (c) IC10, in these figure the amplitude of the peak in frequency domain at this IC is at the highest peak.

information criterion in equation (13) from [44] was used, in the case of before seizure three sources were identified as shown in Figure 3.22, the first minimum value in ICr against $m$ curve identifies the number of independent sources, where ICr is at minimum. And two sources is identified for the case of during seizure, as shown in Figure 3.23 this is the elbow of the first ICr curve at minimum location, indicating the number of true sources within the related EEG recordings. The above result confirms the correct calculation of true number of sources.



(a) (b)

Figure 3.21: The topography of the projected ICs for a patient with focal epilepsy during seizure (a) IC5 (b) IC10, in this figure the amplitude of frequency at this IC is at the highest peak.

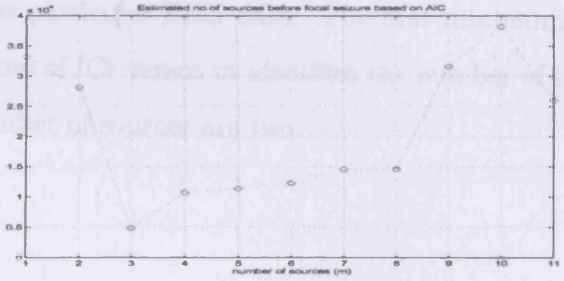



Figure 3.22: Plot of ICr against the number of sources before focal seizure. "o" indicates ICr of this particular EEG data. The first minimum value in the curve, which is the minimum of ICr versus $m$ identifies the number of independent sources. In this case the number of sources are three.

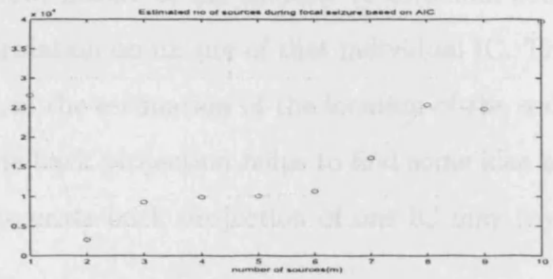



Figure 3.23: Plot of ICr against the number of sources during focal seizure. "o" indicates ICr of this particular EEG data. The first minimum value in the curve, which is the minimum of ICr versus $m$ identifies the number of independent sources. In this case the number of sources are two.

## 3.5 Conclusions

The BSS algorithm gives an estimation of dominant sources known as ICs, with some assumption about nature of the sources. A dominant source with higher amplitude gives a confirmation on nature of that individual IC. This is where the back projection can confirm the estimation of the location of the sources.

In this simulation the back projection helps to find some idea about the location of the sources. In some cases back projection of one IC may cover a large area over the head.

In order to obtain better results the number of sources can be reduced to the signals that are actually generating from the related neurons with their corresponding locations. This can be achieved by filtering the EEG recordings for the specified diagnostic abnormalities, i.e. CJD is dissipating more prominently in Delta band. After identifying the related frequency band, it is useful for identifying the regions of interest by constructing topographical plots of the power in a given frequency band. The topographical plots are important for assessing the location of sources. For diagnosis of an abnormality in the brain or pre-surgerical assessments it is very useful.

Considering spatial and frequency of the independent components within specific period in time can better solve the localization problem. The related number of ICs are estimated using the iterative criterion discussed earlier in this section.

# Chapter 4

# LOCALIZATION IN A 3D SPACE BASED ON THE CORRELATION MEASUREMENTS

A three dimensional space localization based on the correlation measurements of independent sources from a superposition of mixed EEG signal can be accomplished by considering the sparse structural condition of the mixed signal in form of an existing clinical information within specific distinct frequency band. Therefore the application of BSS and measurement of the correlation between ICs and the mixtures give the level of contribution of each estimated source to the electrode signals. In the separation process based on correlation, it is assumed that the ICs are identical to the original sources, and the separation should be based on desired features linked to optimal number of sources and frame size that would separate the features of interest.

After identifying the correlations the least-squares (LS) algorithm can be implemented to estimate the location of the corresponding sources. Hence to achieve a unique solution we need an extra point for the number of dimensions. So, since we are dealing with 3-D space the minimum number of electrodes for having a single solution is 4 [2].

The steps taken for localization of the sources based on correlation over the scalp:

- **EEG recordings**: Store the position of all the EEG readings obtained in a lined up manner, so each position is known after each separation process.

- **Sub-band**: Subdivide the EEG signals into related sub-bands (i.e. Delta, Theta, Alpha, Beta, Gamma) using bandpass, or auto-regressive moving average (ARMA) filtering to capture diagnostic features of abnormalities in the brain.

- **Whitening**, the mixture.

- **BSS**: Apply a BSS algorithms such as JADE, SOBI, or any other suitable algorithm to separate the EEG mixture recordings into number of statistically independent sources, normalizing the result and sort them in a descending order.

- **Correlation**: Since the separated data is sorted in a descending order of amplitude the most dominant signal will be on top of the list. By correlating the dominant signal of each independent component with the original mixed vector data $\mathbf{x}$. The largest correlation value

identifies the closeness of the source to the estimated location of the first most contributed EEG signal.

- **Eliminate and correlate**: Store the most contributed signal and eliminate the electrode location contributing to this effect. Now there will be one less electrode Signal. Repeat the correlating process with one less EEG electrode signal and store the result.

- **Threshold**: Repeat the procedure up to a threshold value or up to a level to ensure all the electrode locations corresponding to that specific geometrical feature are covered. If two sources are closed to the same electrode, the threshold set the final number of the most contributed EEG signals and their level of contribution.

- **Convert an undetermined condition to determined condition**: The most contributed values stored in each procedure correspond to the specific electrode location, which contains the information about the related diagnostic features. Hence it eliminates the repeated, or unwanted EEG signals. What is achieved here is to change an undetermined case into a determined case.

- **Intersection point of the three spheres using LS method**: Applying LS to find the intersection point of the three spheres [110] using the measure of correlation between the dominant IC, and the corresponding electrode signal. Assuming a homogeneous media for the head, the location of the source can be calculated by minimiz-

ing the sum of the square error between the fitted value and the measured data.

Before outlining the localization based on correlation method there are some experiments carried out as described below:

## 4.1 The Information Held by Mixing Matrix A

The mixing matrix **A** from equation (1.1) holds the information about the geometrical location of the sources. To check the effect of geometrical location of the sources on mixing matrix consider two signals $s_1(t)$ and $s_2(t)$ as sources shown in (4.1) desipating signals within a brain with a head radius of seven centimeters. Since the electrodes captures the sources which are attenuated by the head tissue, the mixing matrix **A** is expected to have its columns proportional to the inverse of the distances from the sources. We can consider the following synthetic sources:

$$\begin{aligned} s_1(t) &= \sin(40\pi t) \\ s_2(t) &= \sin(24\pi t) \\ \mathbf{s}(t) &= [s_1(t) s_2(t)]^T \end{aligned} \tag{4.1}$$

$$\mathbf{A} = \begin{pmatrix} 4.5 & 9.4 \\ 3.8 & 8.9 \\ 3.7 & 8.1 \\ 3.8 & 7.5 \\ 4.3 & 6.8 \\ 5.9 & 5.5 \\ 6.7 & 4.5 \\ 8.0 & 3.8 \\ 8.9 & 4.0 \\ 9.5 & 4.5 \end{pmatrix} \tag{4.2}$$

The source signals are shown in Figures (4.1) and (4.2). JADE is applied to

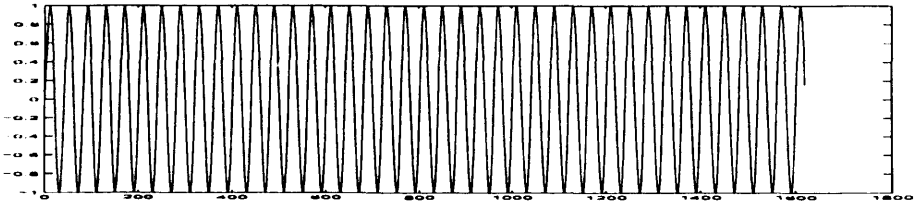


Figure 4.1: Signal: $s_1(t) = \sin(40\pi t)$

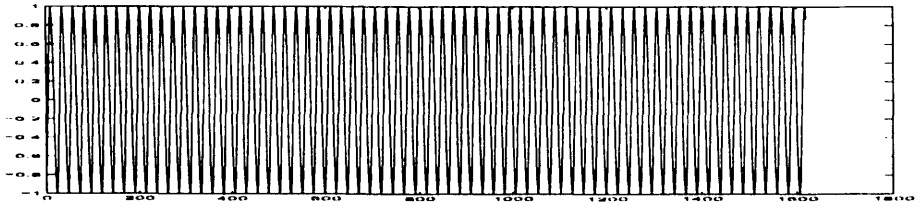


Figure 4.2: Signal: $s_2(t) = \sin(24\pi t)$

generated mixing matrix $\mathbf{x}$ in (1.1), and separated source signals are shown in Figure (4.3). Here $B^T$ representing the transpose of the un-mixing matrix.

$$\mathbf{B}^T = \begin{pmatrix} 0.0523 & 0.0282 \\ 0.0522 & 0.0307 \\ 0.0465 & 0.0262 \\ 0.0401 & 0.0204 \\ 0.0295 & 0.0096 \\ 0.0046 & -0.0172 \\ -0.0111 & -0.0335 \\ -0.0280 & -0.0525 \\ -0.0333 & -0.0602 \\ -0.0334 & -0.0624 \end{pmatrix} \tag{4.3}$$

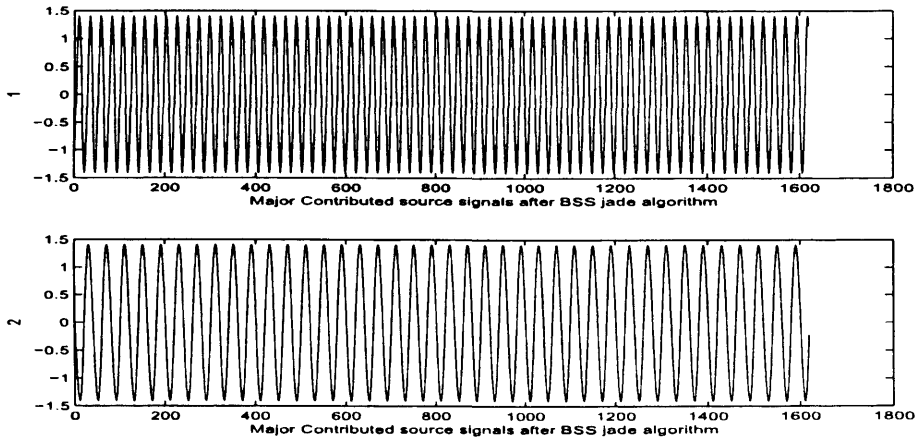



Figure 4.3: Recovered signals by using the described separation algorithm

Now to simulate the condition for the case when one source is detected via different geometrical location sensors. This can be represented as two sources emitting the

same signal, from different locations i.e. $s_1(t) = s_2(t) = \sin(40\pi t)$ with the mixing matrix shown in (4.2). After applying the JADE algorithm to the mixed signals $\mathbf{x}$ from equation (1.1), the unmixing matrix $\mathbf{B}$ can be observed as shown in (4.4).
In this occasion the un-mixing matrix $\mathbf{B}$ ignoring the scaling factor is defined as:

$$B^T = \begin{pmatrix} 3.8816 & 3.8816 \\ 1.4379 & 1.4379 \\ -0.3844 & -0.3844 \\ 1.7773 & 1.7773 \\ -1.9326 & -1.9326 \\ 1.0730 & 1.0730 \\ -4.3742 & -4.3742 \\ -0.3844 & -0.3844 \\ -1.2286 & -1.2286 \\ -0.6445 & -0.6445 \end{pmatrix} \tag{4.4}$$

It can be noted that the transpose of the un-mixing matrix is combination of similar values in both rows. When the elements of mixing matrix is altered the result of unmixing matrix displayed a new set of values but each row with a same values again.
In the next experiment, three signals were generated, two of them were the same and the third one was different, along with a three rows of mixing matrix . The result of unmixing matrix displayed two different rows. We can conclude that the JADE algorithm served its propose as for the two similar signals it separated accordingly. Further studies required to examine the unmixing matrix and their behavior to solve the localization problems.

## 4.2 PCA Rather Than ICA

In PCA the redundancy is measured by correlation between the data elements, while in ICA the concept of statistical independency is used. Also in ICA the reduction of the number data dimensionality of variables is less emphasized. Using only the correlation as in PCA has advantage that the analysis can be based on second-order statistics only. PCA is a useful pre and post processing step. Lets examine the difference between ICA and PCA with an experiment. In previous experiment the source signals are assumed to be those defined by $(s_1(t) = s_2(t) = \sin(40\pi t)$ with the mixing matrix shown in (4.2)), and they are configured as shown in (1.1) format. After running **E**=eig(cov(**x**')), where $\mathbf{x}'$ denotes transpose of **x** in Matlab, and observing the eignevalues shown in (4.5) it is indicated that there is only one nonzero eignevalue meaning there is only one source in this mixed signal. As far as ICA is concerned the two signals are two independent entities with similar values. In PCA the correlation result gives only one value, that means PCA only shows one source where we have actually two sources, therefore PCA fails to represent the actual sources.

$$E = \begin{pmatrix} -0.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 749.9507 \end{pmatrix} \tag{4.5}$$

Observing un-mixing matrix can help to estimate the characteristics of the sources. Furthermore, if we apply the geometrical constraint (i.e. priori knowledge about the appropriate location of a source within an specific frequency band and possibly the direction of arrival of the signals) it helps us to estimate the location of smaller number of sources.

## 4.3 Variance Test

A test to verify the result of the previous experiment i.e."PCA rather than ICA" is the variance test; finding the covariance and eigenvalues of the mixed signals using fourth order cumulant as shown in equation (2.14) and estimate a threshold corresponding to the maximum eigenvalues. The covariance matrix of the mixed signal $\mathbf{c} = \langle \mathbf{x}(k)\mathbf{x}(k)^T \rangle$ and the eigenvalues $\Lambda = eig(\mathbf{c})$; are used in BSS to evaluate the number of sources. Amongst the eigenvalues there are some values which are

extremely small to be considered as element of any source, and should not be used for evaluating the sources. This is where threshold is set. The threshold value may be optimized by using an iterative procedure to have the maximum peak to average value for different cross correlations. The threshold value can be set by $THR = 0.15\hat{\Lambda}$. where $\hat{\Lambda}$ are the eigenvalues. A threshold of 15 percent was set for this experiment, and the result on Figure (4.4) is obtained. The cross- correlations

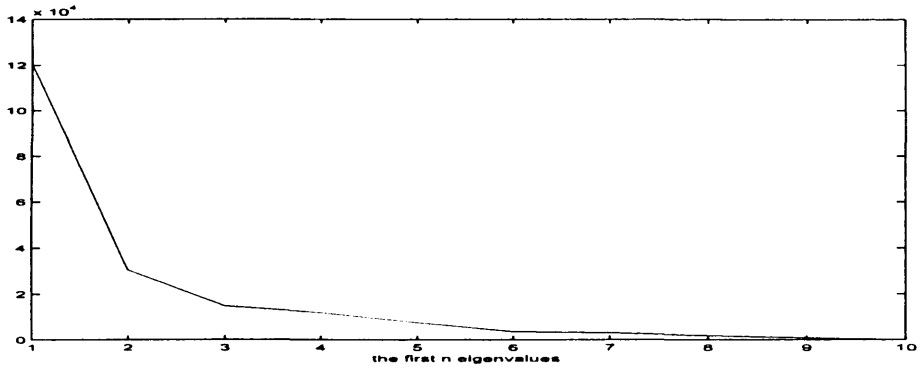



Figure 4.4: The first n number of Eigenvalues of the 4th order cumulant tensor

between the outputs and the mixtures produce the correlation charts, as shown in Figure(4.5). In the correlation charts one may justify the number of sources when for each output there will be a separate dominant peak; also there will not be two same size peaks in one chart. The value of threshold can be updated. The number of sources is proportional to the eigenvalues above the threshold levels.

The algorithm seems to localize the abnormalities and makes sure that the localized signals are actually EEG sources and not noise. If the output dominant signals were noise the result from the correlation with the original signal would not give a peak value, thus it proves that the resulting values are the actual sources which are to be localized.

In some occasions the threshold may result in localizing the abnormality. A three dimensional position of each electrode can be specified for the conventional $10-20$

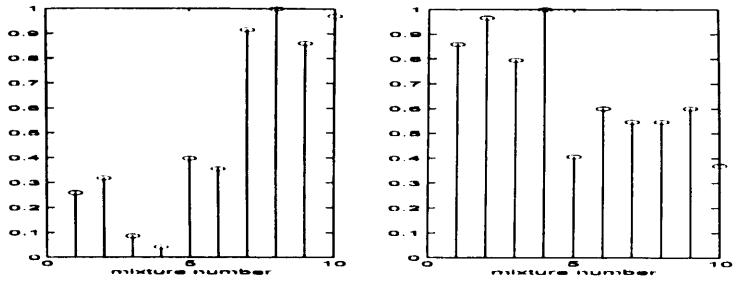


Figure 4.5: The cross-correlations between the estimated sources and the mixtures and the corresponding correlation charts

EEG electrode cap as a hemisphere. A priori knowledge of frequency band (i.e. Delta, Alpha, Beta) of the abnormalities helps to filter out the noise and unwanted signals, resulting in the error reduction in the final calculations. Knowing the radius of the head and locations of the electrodes, the source positions can be estimated.

## 4.4 Localization of the Brain Sources

Sources may be localized by

- Finding the permutation **R** and the mixing matrix **A**
- Separation and localization of the EEG signals based on correlation measurement

Which each are described as followings:

### 4.4.1 Finding permutation R and mixing matrix A

Following our discussions about ICA in previous chapter, it is required to determine the order and scales of the independent components

$\mathbf{x} = \mathbf{As} = (\mathbf{AD}^{-1}\mathbf{R}^{-1})\mathbf{RDs}$ where **D** is a diagonal scaling and **R** is a permutation

matrix.

Diagonal scaling or the variances (energies) $\mathbf{D}$ is unknown. This is because both s and $\mathbf{A}$ being unknown, any scalar multiplier in one of the sources $s_i$ could always be cancelled by dividing the corresponding column $\mathbf{a}_i$ of $\mathbf{A}$ by the same scalar, say $\alpha_i$:

$$x = \sum_i (\frac{1}{\alpha_i}\mathbf{a}_i)(s_i\alpha_i) \tag{4.6}$$

As a consequence, we may quite as well fix the magnitudes of the independent components. Since they are random variables, the most natural way to do this is to assume that each has unit variance: $E\{s_i^2\} = 1$. Then the matrix $\mathbf{A}$ will be adapted in the ICA solution methods to take into account this restriction. Note that this still leaves the ambiguity of the sign: we may multiply an independent component by $-1$ without affecting the model. This ambiguity is, fortunately, insignificant on the localization of the sources as it can be normalized to a single scale factor. But we can not find the order of independent components due to the permutation matrix $\mathbf{R}$. Formally, $\mathbf{R}$ or its inverse can be substituted in the model to give $\mathbf{A} = \mathbf{RDW}^{-1}$. Generally, if $\mathbf{R}$ is known then the mixing matrix, $\mathbf{A}$ will be obtained from the unmixing matrix, $\mathbf{W}$. We insist that incorporating some constraints such as locations of some known sources within the brain, restriction on the geometrical boundaries of the region in which the sources are, or time-frequency properties of the signals within different subbands may solve the permutation problem thereby accurate localization of the sources can be achieved.

## 4.5 Separation and Localization of the EEG Signals Based on Correlation Measurement

This experimental simulation represents the algorithm developed in order to separate the EEG in different frequency subbands.

In this experiment the conventional 10-20 electrode setup is been considered. Figure 4.6 shows a typical scalp electrode layout. There are ten electrode positions selected to be used in this experiment as shown below:

FP1 and FP2 : left and right fronto-polar

F3 and F4 : left and right superior frontal

C3 and C4 : left and right Central

P3 and P4 : left and right Parietal

O1 and O2 : left and right Occiptial

A1 and A2 : the referential ear electrodes.

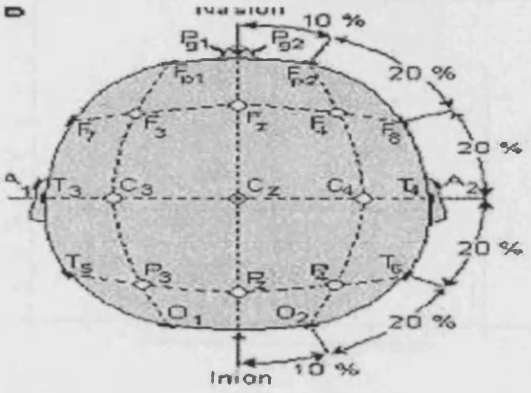



Figure 4.6: Convention 10/20

In this experiment the EEG data acquired in referential mode, meaning that all channels are derived with reference to a common left and right ear electrodes.
The program initially read the mixed signal in the form of EEG recordings obtained from a patient with Jakob-Creutsfeldt disease (CJD). CJD introduces periodic spikes in delta band. Therefore a function in Matlab is generated in order to filters the EEG recordings for each frequency band. From known diagnostic hypothesis the

CJD's EEG recordings were dissipated in Delta band, therefore a frequency range of $0.5 - 4$ Hz for Delta band is focused in order to concentrate on the neurons generating such source signals. Hence the number of sources is reduced to a smaller number. The original mixed signals and the filtered mixed signals for this example are shown in Figures 3.2, and 3.3, respectivley (at section "Filtering"). Next step JADE algorithm is used to separate the statistical characteristics of the EEG recordings (known as mixed signals) and then sort them in descending order (i.e. according to their eigne values). Therefore the first location in the sorted list signal would be the major contributed signal i.e. $correlation = \sum_{i=1}^{N}(Y_1 * x_i)$ where $Y_1$ is the most contributed signal after BSS, and $x_i$ is $i^{th}$ mixed signal. This can be seen in Figure (4.7). The rest of the procedure is as explained at the beginning of this

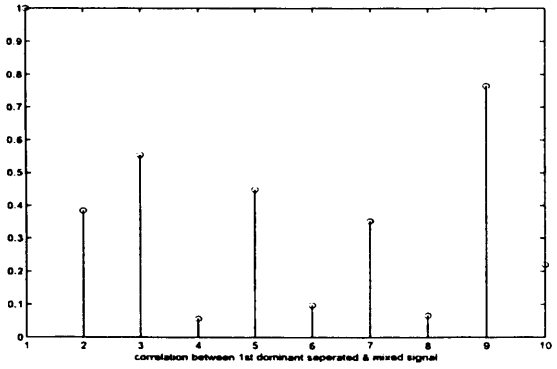



Figure 4.7: Correlation between 1st contributor and the mixed signals

chapter. Hence the result of correlation measurement gives several maximum values as shown in Figure 4.8.

The program stops and returns all of those major contributed signals obtained through the process, this is shown in Figure 4.9. Hence, the locations of these major contributed signals from the original EEG configuration help to find the closest sensors location to the position of the sources. In this experiment the localized sources are at locations O1, O2 (left, right Occiptial), and C3 (left central).

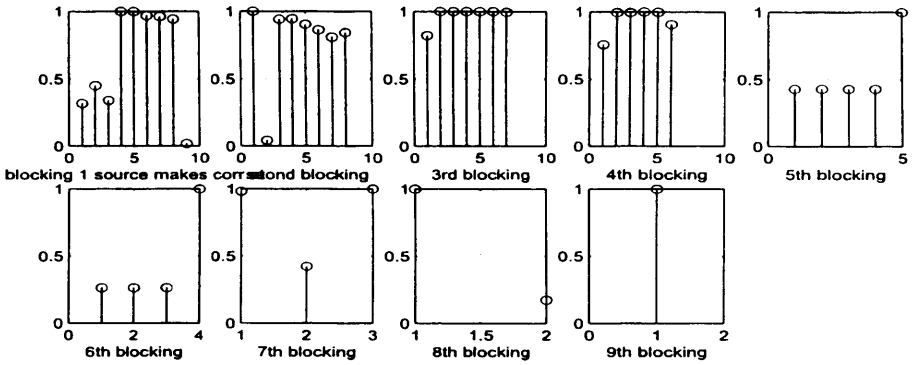



Figure 4.8: Blocking each contributor after correlation.

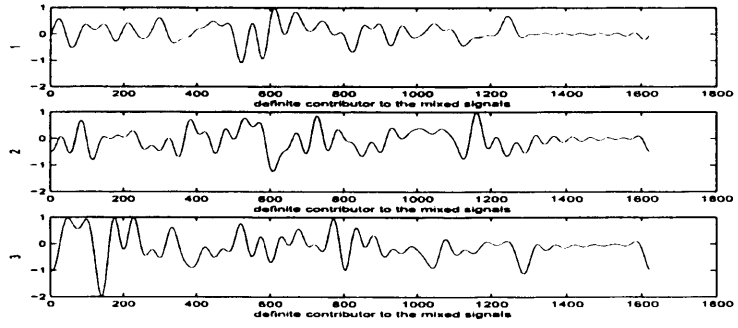



Figure 4.9: Definite contributors to the mixed signals.

### 4.5.1 Estimation of the point of intersection using the signals' contributions

The mixture signals received from electrodes contain information which can help to localize the abnormalities in a patient. This can be achieved by initially separating the independent components of the mixture signals, as shown in equation (1.2) and sort their eigenvalues in a descending order ($\mathbf{y} = [y_1, y_2, ..., y_N]$). The prim location of signal $y_1$, which has largest eigenvalues than other mixtures, represents the dominant signal from the separated signals after BSS. The next step is to correlate the dominant signal with the original set of the mixture signal $C_i =< y_1.x_i >$ for $i=1,\ldots,n$, the largest value of the correlation represents the closeness of the mixture to the source signal. This signal can be named as the most contributed signal (MCS). In each step the electrode location related to the MCS is eliminated in order to allocate further MCS signals. The process of elimination repeats until the resulting separated signals and their correlation with the dominant signal reaches a threshold value, which gives several maximum contributed values (i.e. the system doesn't give unique MCS any longer). The resulting most contributed values represent the closeness to the location of the source signals. These values are normalized with respect to the radius of the head. If we consider the location of each major contributed electrode to be the center of a sphere, a point where all the spheres (the radius of each sphere is proportional to the inverse of the measured correlations) intersect is the location of the original source. Consequently the problem of determining the points of intersection of $n$ spheres in $\mathbb{R}^n$ is applied [110].

Let us assume there are three electrode locations. The coordinate of electrodes are known and it is required to calculate the coordinates of an unknown point when the distances of the unknown point from the given points are known. This problem is clearly equivalent to finding the intersection point of the three spheres as shown in

Figure 4.11 [110]. Assuming a homogeneous media for the head, the location of the source can be calculated using the following equation:

$$||\mathbf{f}_s^k - \mathbf{a}_j||_2^2 = d_j^2 \qquad j = 1, 2, \ldots n \tag{4.7}$$

where $k = 1, \ldots, 4$ are the frequency sub-bands (i.e. Delta, Theta, Alpha, and Beta) of sources with their geometrical locations (i.e. let assume in a case they are three such as P3, C3, and O1), $\mathbf{f}_s^k$ are the three dimensional coordinates of the sources, and $\mathbf{a}_j(x, y, z)$, are the pre-calculated geometrical locations of the electrodes over the scalp shaped as shown in Figure (4.10) with the values shown in table 4.1. All the values in the table are in radians and the head radius is considered to be unity. Therefore for each individual case the radius should be calculated separately (*i.e.* $Circumference = 2.\pi.raduis$).

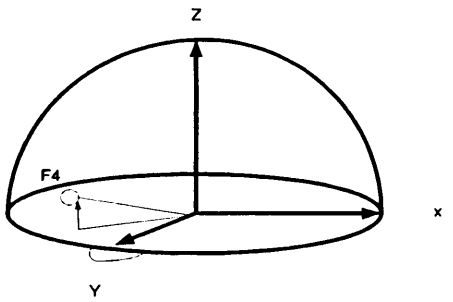



Figure 4.10: Electrode position within a Cartesian coordinate

The result for intersection of three spheres in a common point from the last simulation when only homogenous condition is considered illustrated in Figure 4.12.

| Electrodes | X | Y | Z |
|---|---|---|---|
| Fp1 | -0.29389 | 0.90451 | 0.30902 |
| FP2 | 0.29389 | 0.90451 | 0.30902 |
| F3 | -0.47553 | 0.58778 | 0.65451 |
| F4 | 0.47553 | 0.58778 | 0.65451 |
| C3 | -0.58778 | 0 | 0.80902 |
| C4 | 0.58778 | 0 | 0.80902 |
| P3 | -0.47553 | -0.58778 | 0.65451 |
| P4 | 0.47553 | -0.58778 | 0.65451 |
| O1 | -0.29389 | -0.90451 | 0.30902 |
| O2 | 0.29389 | -0.90451 | 0.30902 |
| F7 | -0.76942 | 0.55902 | 0.30902 |
| F8 | 0.76942 | 0.55902 | 0.30902 |
| T3 | -0.95106 | 0 | 0.30902 |
| T4 | 0.95106 | 0 | 0.30902 |
| T5 | -0.76942 | -0.55902 | 0.30902 |
| T6 | 0.76942 | -0.55902 | 0.30902 |
| Fz | 0 | 0.58778 | 0.80902 |
| Cz | 0 | 0 | 1 |
| Pz | 0 | -0.58778 | 0.80902 |

Table 4.1: Geometrical locations of the electrodes over the scalp within a cartesian coordinate

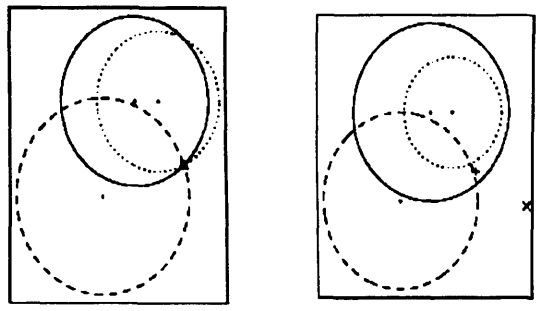


Figure 4.11: Cross-section of 3 spheres with an intersection point (left) and without a common intersection point (right) adopted from [110]

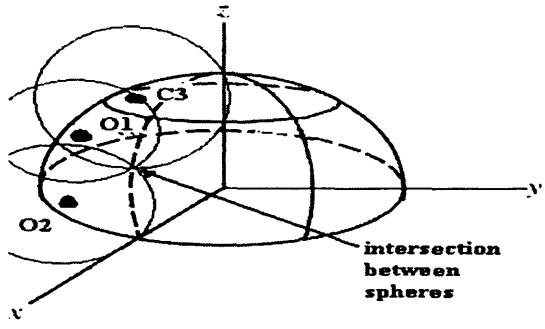



Figure 4.12: Intersection of three sphere in a common point represents location of a source

In equation (4.7) all the parameters on the left side of the equation are the source and mixture coordinates. The parameter $d_j$ is inversely proportional to the correlation between the M estimated sources and the three major contributing signals:

$$d_j \propto \frac{1}{< x_j(t).y_i(t) >} \quad for j = 1,2,3 \quad i = 1,2,\ldots \tag{4.8}$$

where $d_j$ is normalized with respect to the radius of the head. $x_1$ is mixed signal, and $y_1$ is correlated MCS with corresponding mix signal. Hence the source location can be identified, as shown in Figure 4.13.

However, head comprises three layers of brain, skull, and scalp with different conductivities. Therefore the transitions from one layer to another has to be incorporated in the above calculation. This can be performed either by nonlinearly normalizing the links $d_j s$ or by accurately considering the non-homogeneity of the brain. The estimated location of the sources in an inverse calculation solving for $f_i^k$ in equation (4.7). As a conclusion, the purpose of this experiment is to use an iterative BSS

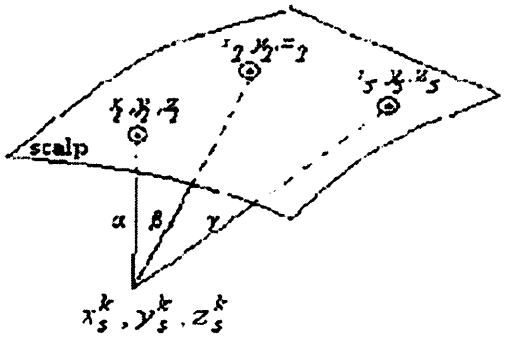



Figure 4.13: Part of scalp including the electrodes and the location of a source to be identified

method to detect the number of original contributed EEG signal sources, based on

correlation between ICs and the electrode signals. This method changes an undetermined system into a determined one. A priory knowledge of the geometrical location of these major contributed signals along with their related correlation values helps to localize the sources based on intersection of three spheres. We assume the medium is homogeneous and the source is located at a unique position. The outcome of the experiment with regards to localization of the CJD sources; the biopsy after postmortem has been indicated that the patient with CJD would have spongiform changes in cerebral cortex and cerebellum. The result of this experiment confirms the location of source in occipital lobe, where the cerebellum is situated.

## 4.6 Incorporating Non-homogeneity of the Head Tissues

There are various factors which violet the homogenous assumption of the dense media such as:

- Various magnetic properties of the layers: The hydrogen atoms alternately absorb and emit radio wave energy, generating non-homogenous environment between different medias of the head.

- Various resistivity (conductivity): The resistivity of different layers generate non-homogenous environment for penetration of the brain source signals to the sensors.

- Various noise: In general EEG signals are statistically non-stationary and corrupted by other human internal signals, such as eye blinking signal, heart beat, noise of the measurement system, environment noise, and interference from the adjacent electrode signals.

All above phenomenon creates a nonlinear conditions for the signal sources of brain for calculation.

### 4.6.1 Non-homogenity problem

With some indeterminacy in the result we can approximate the location of the sources within the brain. Unlike the methods in [118] and [119], which consider the sources as magnetic dipoles, we simply consider them as the sources of isotropic signal propagations. Therefore the head (mixing medium) model only mixes and attenuates the signals. The attenuation corresponds to the distance and the resistance of the medium between the sources and the fixed electrodes. Alpha waves are recorded from the occipital and parietal regions of the cerebral cortex. However, the Alpha waves from the occiptial area are prominent with higher amplitude [127], [134]. These sources generate reference signals within a small frequency band of $7 - 13$ Hz in healthy adults, without any attention, visual, or possibly auditory stimulation, and without dysfunction of the central nervous system (CNS). Since we can measure both the link weights and the energy of the mixtures within the selected bands we will be able to compensate for the nonhomogenity by finding a relationship between $\tilde{\mathbf{A}}$, found through measurement of the geometrical locations and $\mathbf{A}_g$, found through measurement of the energy of the signal(s) of the known source(s). The energy within the Alpha band is obtained by carefully bandpass filtering the EEGs around the peak in the range of Alpha frequencies. These amplitudes are then inverted to give the entries of the $k$ columns of $\tilde{\mathbf{A}}$.

On the other hand, the geometrical location of the known sources can be approximately determined off-line (denoted $\tilde{\mathbf{A}}_g$). It is clear that $\tilde{\mathbf{A}} = f(\tilde{\mathbf{A}}_g)$, where $f$ represents the nonhomogenity of the medium between the known sources and the electrodes. Instead of using the sources of normal brain rhythms as a known a pri-

ori, we may synthetically provide a number of sinusoidal sources in certain locations under the skull. This may be done by setting a number of electrodes under or close to the brain through the nose or mouth. No significant invasive surgical operation is needed for such purposes. In the second method, by using a set of sharp bandpass filters, projection of the sinusoidal waves to the electrodes can be easily evaluated. Therefore the entries of $\tilde{\mathbf{A}}$ will be accurately identified. Having more known sources, the positions of the sources as well as measuring the non-linearity resulting from the non-homogeneity of the head including the brain (white and gray tissues), the skull, and the scalp can be estimated more accurately. In a spherical model of the head we may consider three main layers; brain, skull, and scalp for which the thicknesses are known. To incorporate the non-homogeneity $f$ has to be completely identified for all the sources. In some simplified practical situations the column vectors of $\mathbf{A}$ are proportional to $1/\Gamma^{\alpha}$ [rather than $1/\Gamma$]. where $\alpha$ is non-linear variable, as described in Chapter 4, Non homogeneity effects.

Having more than one known source location, in order to extend the above nonlinear map to all the estimated source locations a simple means of extrapolation of the columns of the estimated mixing matrix would be adequate.

### 4.6.2 Non-homogeneity considered as Nonlinearity of the head

In blind source separation the mixing matrix corresponds to the link weight between the sources and the sensors. After separation the resulting independent components are approximating the original sources. The BSS only reproduces the statistical representation of the original source without consideration of the scaling and order of the original signal. This separation is valid for both case of homogenous and non-homogeneous case of the media. In the non-homogeneous case of the brain where

the source signals penetrate and pass the white and grey matters and blood fluid, skull, and scalp each of which with different conductivity or resistivity, until reaching to the electrodes causes attenuation to the signals received. The effect of this non-linearity or non-homogeneity can not be considered in BSS as the effect on scaling and rotation of the original signals is unaltered. It is required to generate a model that shows how the environmental noise are causing the attenuation of signals, hence this helps to create an inverse solution to the effect of non-homogeneity of different media in localization problem. In a homogeneous sphere model shown in Figure 4.14, the isotropic sources are attenuated and mixed while travelled to the scalp across brain in all directions. High amplitude signals measured at the electrodes could represent the closeness of those electrodes to the abnormal sources within an specific frequency band, and thus they can be considered a prior knowledge for location of some of the sources. In a model this can be represented as weighted link between sources and electrodes. Hence, a minimum of three weighted links between sources and electrodes are sufficient to estimate the location of a source. Therefore, the problem of localization of sources can be considered as to find the intersection of three spheres in a three dimensional space, when the radii and centers of the spheres (the electrode locations) are known [110].

When the three spheres do not intersect in a common point, e.g. due to error in the estimated distances caused by noise and other nonlinearity factors, LS minimization can be used to obtain an optimum solution. If a solution exist, the intersection of the spheres sometimes gives two different points, but the selection between them can be done considering physiological aspects, for example, if one of the solutions lies outside the brain, it can be discarded. Therefore one solution would be the estimated location of the source.

To achieve a unique solution we need an extra sphere. Therefore, since we are dealing with three dimensional space the minimum number of electrodes for having

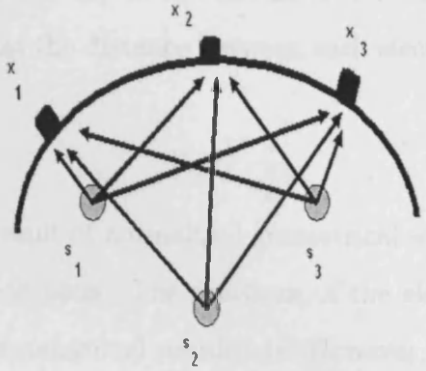



Figure 4.14: Homogeneous spherical model of the brain

a single solution is four.

In homogeneousspherical brain model, the largest value of the correlation represents the closeness of the mixture to the source signal. Therefore, the cross-correlation between the sources and the mixtures will be related to their distances. This assumption enables us to estimate the distances as a function of the cross-correlation. However, the amount of internal noise increases with the distance from the electrodes, assuming the noise is spatially uniformly distributed. Figure 5.4 illustrates the relationship between the electrodes and the mixing matrix **A**. The columns of the mixing matrix refer to the locations of brain sources within the brain. In a homogenous medium $a_{ij}$ are inversely proportional to the distances from the sources i.e. $a_{ij} \sim \frac{1}{\Gamma_{ij}^2}$. That is because, the signal is attenuated proportional to the square of the distance. On the other hand, in a non-homogenous medium it is not attenuate with the square of the distance but with some other power, which depends on the conductivities of the different layers, or nonlinearities caused by internal or external noise i.e. $a_{ij} \sim \frac{1}{g(\Gamma_{ij})}$ where $g(.)$ is a nonlinear function that describes the effect of the head medium including the scalp, skull, and the brain (white and grey) tissues. Assuming a homogenous head medium each element of the columns of the mixing

matrix is inversely proportional to the distances between the respective electrode and source. It means that the distance between each electrode and each source have the following relation:

$$\Gamma_{ij} \sim \frac{1}{a_{ij}^{0.5}} \tag{4.9}$$

Figure (5.4) shows the result of normalized geometrical source localization when the head is considered homogenous. The positions of the electrodes $e_i$ are depicted in Table 4.1 in cartesian (rectangular) coordinate. However, the homogeneity is a weak assumption. A more precise method can be achieved after the non-homogeneity of head is modelled and exploited [91] [92]. In a realistic environment where conductivity of the tissue layers and where the effect of noise (added uniformly to the signal sources) are involved, reasonable assumptions have to be made in construction of the head model and corresponding $g(\Gamma)$ has to be estimated.

The relationship between homogeneity and non-homogeneity of the head medium is illustrated as a graph of distance against cross-correlation for the case of homogenous media as the straight doted line and for the case of non-homogenous head model as an monotonically decaying curve in Figure 4.15. In reality the head has been considered to be a series of concentric spherical regions (brain, skull, and scalp each of which is considered to be homogenous), as illustrated in Figure 4.16 (Russ and Driscoll, 1969). In this model, the inner and outer radii of the skull are chosen to be 8 and 8.5cm, respectively, while the radius of the head is 9.2cm. For the brain and the scalp a resistivity of $2.22\Omega m$ is selected, whereas for the skull a resistivity of $80 \times 2.22\Omega m = 177\Omega m$ is assigned. These numerical values are given solely to indicate typical (mean) physiological quantities. Because of this symmetry, and simplicity, it is easy to construct a mathematical and computer model. It is also easy to perform calculations with a spherical geometry. Though this simple model does not consider the anisotropy and inhomogeneity of the brain tissue and cortical

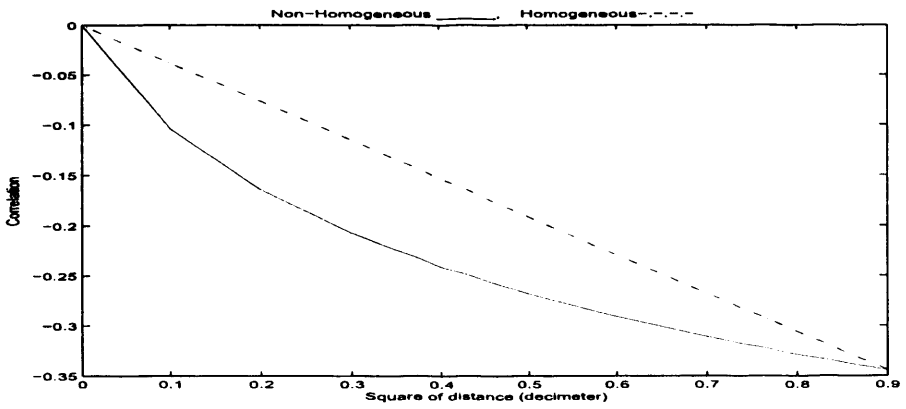



Figure 4.15: The graph of relationship between cross correlation of the source signal verses square of distance between the electrode at the scalp and sources inside the head in decimeters. The slope for Homogeneous spherical model of the brain is shown in doted line and an asymptotically decaying curve between the same cross correlation considering the non-homogenous case shown in solid line.

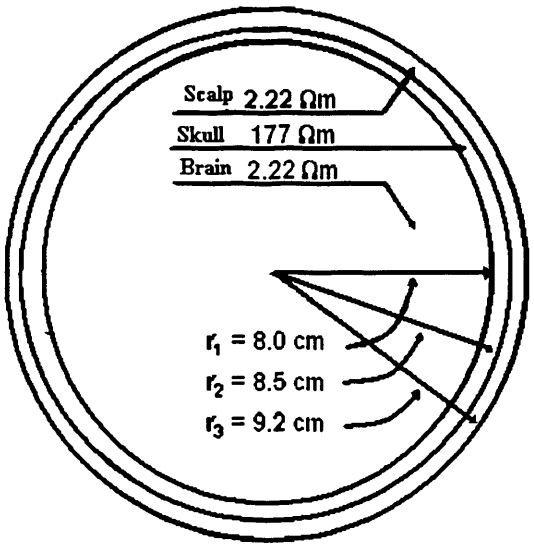



Figure 4.16: Concentric spherical head model by Rush and Driscoll (1969). The model contains a region for the brain, scalp and skull, each of which is considered to be homogeneous.

bone (Saha and Williams, 1992), it gives results that correspond reasonably well to the measurements.

Let assume there is no noise, no scaling and sign ambiguities caused by BSS, and no normalization is considered. The distance from electrode to the source is inversely proportional to the power-2 of the correlation between the source and the electrode signals.

$$\frac{1}{C_{ij}^{\alpha}} = \Gamma_{ij} \tag{4.10}$$

where $C$ is the correlation between the mixture $x_j$ and the estimated source $y_i$, $\alpha$ is the non-linearity factor, and $\Gamma$ is the distance between the source and the electrodes. Obviously the attenuation caused by the resistivity and conductivity of non-homogeneous brain layers makes no different in BSS results. Therefore the way of looking at the problem should be changed. Now we have to look at result of the BSS separated signal and rely our assumption on some other criterion. The most important and effective criterion is the noise. The BSS results are ICs which do not bear any information about locations of the sources. The noise can be a factor to indicate if the source is closer to the sensor or further away from it. By assuming the head containing large amount of noise coming from many directions, and our desire source signal is affected by such unwanted noises. There is no information about the nature of these noises and whether are additive or multiplicative. Consequently the sources closer to the cortex are less affected by noise than those deep inside the brain. According to this assumption the correlation and distance curve can be is shown in Figure 4.17, where from left of the figure straight line slop represents the nonlinear effect of noise caused by scalp and the line continuous in a parallel to horizontal line along with increase of distance between the sensor and the source this is the period when the signal passing through the skull, which shows no noise is affecting the correlation value, and finally as the signal passes through the brain a decaying

curve shows the effect of noise as the signal is getting closer to the sensor location.

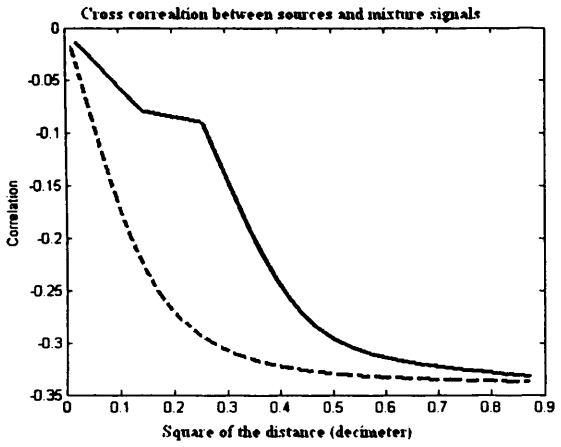



Figure 4.17: Curve of Non-Homogeneous spherical model of the brain considering the effect of noise shown as solid line, and for the homogenous case shown as doted line

If the attenuation and noise characteristics of the head is given, then $\alpha$ can be estimated in such a way to best describe the head model. Now, with such a non-homogenous model, we are required to generate a head model that can incorporate the estimated mixing matrix $\mathbf{A}$ to localize the actual location of the sources. Starting by considering the voltage at sensor $j$ due to source $i$:

$$e_j = \frac{K_1}{\Gamma_1^2}\frac{K_2}{\Gamma_2^2}\frac{K_3}{\Gamma_3^2} f_i \tag{4.11}$$

where $K_1, K_2, K_3$ model the resistivity of the three layers and $\Gamma_1, \Gamma_2, \Gamma_3$ are the segments of source-sensor intervals corresponding to the intervals within the three layers. Remembering that the thickness as well as resistivities of the layers are very different, we wish to find a relation similar to the one of the homogenous medium (4.9). The elements of the columns of the mixing matrix now represent the following:

$$a_{ij} = \frac{K_1}{\Gamma_{ij_1}^2} \frac{K_2}{\Gamma_{ij_2}^2} \frac{K_3}{\Gamma_{ij_3}^2} \tag{4.12}$$

which we wish to relate to the total distance $\Gamma_{ij} = \Gamma_{ij_1} + \Gamma_{ij_2} + \Gamma_{ij_3}$ as follows:

$$a_{ij} = \frac{K}{\Gamma_{ij}^{\alpha}} \tag{4.13}$$

where $\alpha$ is the non-linearity factor. The value of $\alpha$ can be iteratively computed to have the best match between the synthesized sources and the estimated ones. The achieved parameter is then applied to the localization of real EEG sources.

## 4.3. LS solution

As described in the pervious chapter, given the mixing matrix $\mathbf{A}$ one can employ LS minimization [110] to find the location of source $k$ using $\Gamma = \{\Gamma_{k1}\Gamma_{k2}...\Gamma_{km}\}$ where $m$ represents the number of sensors as: where $\mathbf{f}_k$ is the coordinates of the $k^{th}$ source and $E$ is the set of electrodes. The positions of the electrodes $e_j$ are depicted in Table I in cartesian (rectangular) coordinates. The $\Gamma_{ij}$ values are computed according to the non-homogenous model. The LS solution is equivalent to finding the point of intersection of $n$ spheres. To achieve a unique solution we need an extra sphere for the number of dimensions. So, since we are dealing with 3-d space the minimum number of electrodes for having a single solution is 4. Also, the larger the number of electrodes is the better the localization is achieved since errors in measurements or model errors in individual electrodes will be less important. Also, the scaling problem caused by the estimation of the mixing matrix in BSS will be of less significant since the LS algorithm will attempt to find the closest solution.

## 4.7 Conclusions

The work can be summarized into the following steps:

1. **Find A rather than W**: Since mixing matrix **A** does not fully represented by $\mathbf{W}^{-1}$, we need to find a forward model to estimate **A**. The method introduced in [89] is valid only if the mixing matrix is square (i.e. the number of sources are equal to the number of mixtures). In our research we are assuming that the system is undetermined, and the number of sources are unknown to us. Within a iterative coupled loop and a priori knowledge about some of column of matrix **A** the remaining elements of mixing matrix can be calculated.

2. Incorporate the non-homogeneity of the head into the definition of non-linear normalization of the correlation measures. Refractions of current densities cause wrong estimation of the locations due to resistivity and conductivity of inhomogenous layers of the brain. In order to obtain a better estimation of the sources, the thickness and resistivity of scalp, skull, and brain should be taken into account [91] [92]. Its found that the skull conductivity ranges from 0.0735 $Sm^{-1}$ to 0.00467 $Sm^{-1}$ The standard conductivity values for brain, skull and scalp are 0.33 $Sm^{-1}$, 0.0042 $Sm^{-1}$, 0.33 $Sm^{-1}$ respectively [93]($Sm^{-1}$ is siemens per meter( the German-born engineer Sir William Siemens (1823 – 83), the siemens is a unit of electrical conductance equal to $1/ohm$. i.e.= $1/[ohm - meter]$).

3. The elements of the columns of the mixing matrix can be represented as: $a_{ij} = \frac{K_1}{\Gamma^2_{ij_1}} \frac{K_2}{\Gamma^2_{ij_2}} \frac{K_3}{\Gamma^2_{ij_3}}$ and it is required to relate to the total distance $\Gamma_{ij} = \Gamma_{ij_1} + \Gamma_{ij_2} + \Gamma_{ij_3}$ shown:

$$a_{ij} = \frac{K}{\Gamma^{\alpha}_{ij}} \tag{4.14}$$

where $\alpha$ is the non-linearity factor. The value of $\alpha$ can be iteratively computed to have the best match between the synthesized sources and the estimated ones. The achieved parameter is then applied to the localizations of real EEG sources. Given the mixing matrix $\mathbf{A}$, one can employ LS minimization [110] solution is equivalent to finding the point of intersection of $n$ spheres. To achieve a unique solution we need an extra sphere for the number of dimensions. So, since we are dealing with 3-D space the minimum number of electrodes for having a single solution is 4. However for every individual source a better localization can be achieved by increasing the number of electrodes involved, due to the geometrical configuration of the head and the sensors. Furthermore, the scaling problem caused by the estimation of the mixing matrix in BSS will be of less significance since the LS algorithm will attempt to find the closest solution.

# Chapter 5

# PARTIALLY CONSTRAINED METHODS

## 5.1 Extension of the Method Based on the Correlation Measurements

An effective and simple algorithm for localization of abnormal sources of the EEG signals within the brain has been developed here. In this method the signals are separated first, then the estimated independent components are lowpass filtered and normalized. In the next stage the correlation values between the estimated sources and the electrode signals are measured. On the other hand the sources with known locations are separated offline using narrowband bandpass filters. Finally, as the main contribution of this section of the thesis the mixing matrix is estimated using the information about the known sources and the estimated sources. The locations of the unknown sources are then measured with respect to the columns of the mixing matrix and the geometrical properties of the head and electrode locations.

Localization of abnormal brain sources such as focal epilepsy has been an impor-

tant subject of research by both clinical and signal processing workers for at least two decades. Some source localization techniques have been introduced by the researchers in this area based on subspace estimation in beamforming and measurement of the direction of arrival (DOA) of the signals such as MUSIC and RAP-MUSIC [104] [119]. Although in general these methods perform well even at the presence of noise they do not exploit the nature of the sources and make the best use of the prior knowledge about the known sources. Therefore the outcome suffers lack of accuracy in the localization especially where the number of sources is high. An effective and simple update equation has been used here for the separation and localization of EEG sources by incorporating blind source separation (BSS). In this method it is assumed that the normal (Control) EEG sources and also the abnormal rhythms (Tasks) are independent. In localization of the sources we use the fact that the geometrical coordinates of the sensors and the correlation between each independent component (IC) and the observed EEGs can be measured and therefore will be known to us. Also we are able to filter out one or some of the known brain rhythms using carefully centred bandpass filters. The normal Alpha rhythm with a frequency of between 8 to 13 Hz, consistently exist for awaked adults. Also the location of the source is known to be in the hippocampus or the posterior brain lobes. This can be confirmed by correlating the filtered signal with all the electrode signals and localize the source (using the method explained in Section 3) with respect to the three electrode positions whose signals have the largest correlation with that source. After we estimate the location of all the sources (which are basically addressed by the columns of the mixing matrix) with some indeterminacy in the solution, we will be able to readjust the locations with reference to the location of the known sources. This is in fact equivalent to estimation of the permutation matrix. In the following sections the overall procedure is explained and the performance of the method for real EEG signals of a patient with epilepsy is given.

The Infomax BSS algorithm [108], [109], based on minimization of mutual information or maximization of the entropy has been used here to separate the sources from the EEG mixtures (iterative SOBI can also be used here in the same way). The update equation for estimation of the unmixing matrix, W, is defined as

$$\mathbf{W}(t+1) = \mathbf{W}(t) + \Delta\mathbf{W}(t) \tag{5.1}$$

where by considering the extension to the Natural gradient Algorithm (NGA) proposed by Amari [109] we have

$$\Delta\mathbf{W}(t) = \mu\frac{\partial J(\mathbf{W})}{\partial \mathbf{W}}\mathbf{W}^T\mathbf{W}$$

$$= \mu\left\{\gamma\mathbf{I} + (1 - \frac{2}{1+exp(\mathbf{Wx})})(\mathbf{Wx})^T\right\}\mathbf{W} \tag{5.2}$$

Here $J$ is the Infomax cost function, $\mu$ is the learning rate, $\gamma$ is a constant, $\mathbf{I}$ is a unitary matrix. $\mathbf{W}$ is initialized to $\mathbf{W}_{init} = \mathbf{I}$ and $\mu$ is calculated empirically via the following adaptive criterion:

$$\mu(t) = \mu_0\left\{\frac{\alpha}{\|\mathbf{R}_x\|_F^2} + \frac{\beta}{\zeta + \|\Delta J_C\|}\right\} \tag{5.3}$$

Where $\mu_0$, $\alpha$, $\beta$, *and* $\zeta$ are constants adjusted for adaptation.

It has also been demonstrated that the effect of noise on the performance of the update equation is minor since minimization of the main objective function does not depend on the noise. The separation, however, is subject to the inherent scaling and permutation of the estimated sources i.e. $\mathbf{A} = \mathbf{RDW}^{-1}$, where $\mathbf{R}$ and $\mathbf{D}$ are the permutation and scaling matrices.

The proposed algorithm was implemented for separation and localization of epileptic sources. Figure 5.1 shows a set of EEG signals for a patient suffering from focal epilepsy. Various methods as described in [106] and [107] using a constrained BSS

or based on an offline measure can be applied in advance to eliminate the effects of artifacts. Lowpass filtering seems to be very effective in removing the high frequency noise from the estimated ICs. The seizure effect may be observed in almost all the signals detected at the sensors. By looking at the spectrum of the signals we can easily see that the seizure signal has a prominent (slightly decreasing) frequency around $6-7$ Hz. The normal Alpha rhythm has also a frequency of around 11 Hz. Figure 5.2 shows the contribution of the mixtures to the desired rhythmic outputs and the dominant ICs are illustrated in Figure 5.3. Each estimated source is then localized based on the approach described in [110]. At this stage the solution to the following least square problem was obtained:

$$min\ S(\mathbf{f}_k), \mathbf{f}_k \in \mathbb{R}^n \tag{5.4}$$

where

$$S(\mathbf{f}_k) = \sum_{j=1}^{3} \{\|\mathbf{f}_k - \mathbf{a}_j\|_2 - d_j\}^2 \tag{5.5}$$

where $\mathbf{f}$ and $\mathbf{a}_j$ refer to the source and the electrode coordinates respectively, and $d_j$ are nonlinearly proportional to the inverse of the correlations between the estimated source $k$ and the electrode signals (the mixtures). $j = 1, 2, 3$ represents the electrode involved in calculation of the correlation values, and $k = 1, 2, \ldots, M$, shows the source number. In these equations all the variables except the source coordinates, $\mathbf{f}$, are known. The solution gives the approximate locations of all the sources. Unfortunately, there are ambiguities in both the accuracy and the number of solutions. In order to mitigate the ambiguities we need to use a priori information about the known sources. This is done in two steps: The estimated sources are scaled to make the energy of the known sources equal to those extracted by bandpass filtering. The permutation matrix, R, is estimated by minimizing $\|\mathbf{A}_{es}\text{-}\mathbf{RW}^{-1}\|_F^2$ , where $\mathbf{A}_{es}$ is an $n \times m$ matrix with a number of known columns proportional to the known sources and zeros for the rest, W is the estimated unmixing matrix using

BSS, and $\|.\|_F$ is Frobenius norm. The solution to this is simply achieved through linear programming. The mixing matrix A will be then determined as $\mathbf{A} = \mathbf{R}\mathbf{W}^{-1}$. Unlike the localization methods in [104] and [119], which rely on decorrelation of the sources, this method exploits the independency of the sources. Moreover, the computational cost is much lower. Using this method the seizure source was localized around the left tempro-lateral lobe segment. This has been verified by appearing white patches in the brain fMRI (BOLDs) in the same segment.

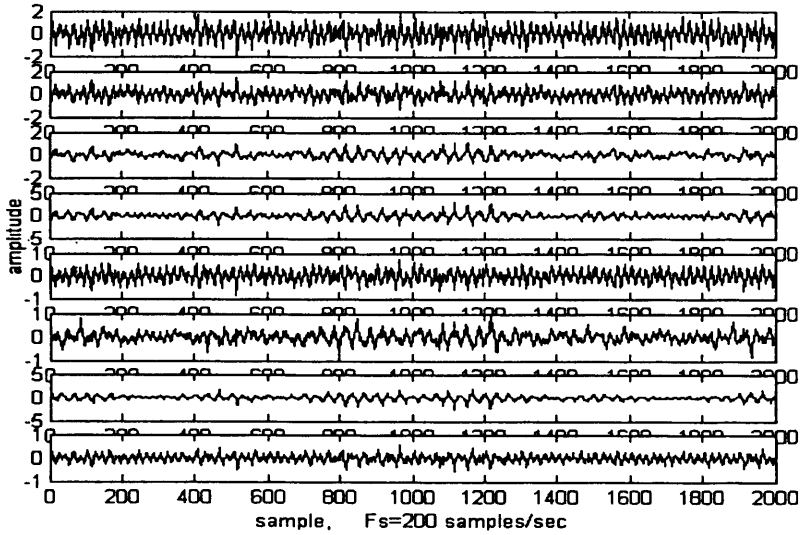



Figure 5.1: The original electrode signals during an ictal period.

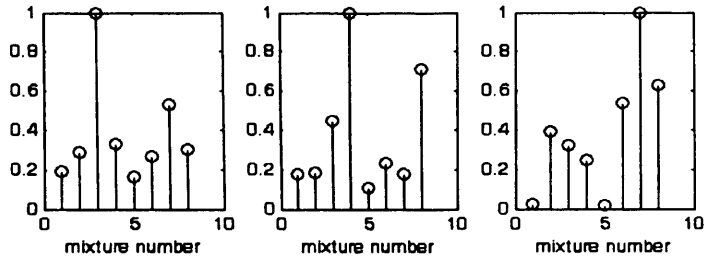



Figure 5.2: Representing the normalized correlation between the estimated sources and the mixtures (electrode signals), e.g the second bar in the third frame is the normalized correlation between the second mixture and the third output.

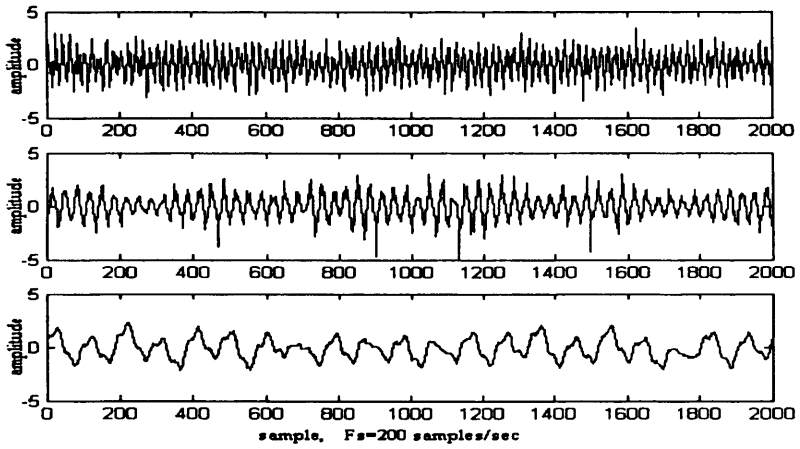



Figure 5.3: The separated dominant sources from the signals in Fig. 5.1; the top figure represents the normal Alpha rhythm. The middle one is clearly an epileptic seizure signal with a frequency of between 6-7 Hz. The bottom signal is a Delta rhythm.

## 5.2 Partially Constrained BSS for Brain Source Localization

The instantaneous BSS formulation in (1.1) and its corresponding estimation of the sources $\mathbf{y}$ in (1.2), requires a solution for finding the unknown column vectors of the mixing matrix $\mathbf{A}$ and the sources $\mathbf{s}$.

The separation matrix, $\mathbf{W}$, can be found by finding the global minima (or maxima) of a cost function, which provides a measure of independency of the estimated sources. Using ICA we can attempt to separate the signals into their independent components. The number of outputs may be approximated by one of the methods described in [119]. However, the separation is subject to the scaling and permutation of the sources i.e.

$$\mathbf{A} = \mathbf{DRW}^{-1} \tag{5.6}$$

Where $\mathbf{D}$ and $\mathbf{R}$ are the scaling and permutation matrices respectively, and $\mathbf{W}^{-1}$ is the pseudo inverse. The effect of $\mathbf{D}$ can be constrained by the size of the head and it can be generally disabled by normalization of the estimated separating matrix after each iteration. However, without solving the permutation problem no solution to the estimation of $\mathbf{A}$ will be possible. This means there will be no clue to finding a unique solution to the localization problem.

As an example, Figure 5.4 shows part of the scalp including three electrodes and two sources located inside the brain. In general, the distances between the sensors and

the sources are inversely proportional to the correlations between the corresponding estimated sources and the observed EEGs. To formulate the problem consider $k$ out of $m$ sources are known. This means that the scaled values of the $k$ columns of $\mathbf{A}$ are known. In the example of Figure 5.4 we may expand the generative model to

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{AS} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \cdot \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \tag{5.7}$$

and if source $s_1$ is known we define a new matrix as $\tilde{\mathbf{A}}$ such that its columns correspond to the known and the unknown sources the element of which are represented by $a_{ij}^{k}$ and $a_{ij}^{uk}$ respectively. In general case

$$\tilde{\mathbf{A}} = [A_k \vdots A_{uk}] = \begin{bmatrix} a_{11}^{k} \dots a_{1k}^{k} & a_{1k+1}^{uk} \dots a_{1m}^{uk} \\ \vdots & \vdots \\ a_{n1}^{k} \dots a_{nk}^{k} & a_{nk+1}^{uk} \dots a_{nm}^{uk} \end{bmatrix} \tag{5.8}$$

where $A_k$; $n \times k$, and $A_{uk}$; $n \times (m\text{-}k)$, are respectively the known and unknown sub-matrices.

In most of the BSS algorithms $\mathbf{W}$ is calculated iteratively in order to obtain the most statistically independent sources. Now, during the separation process we may simultaneously try to enforce the following constraint.

$$\mathbf{J}_c = \|\tilde{\mathbf{A}} - \mathbf{RW}^{-1}\|_F = trace([\tilde{\mathbf{A}} - \mathbf{RW}^{-1}][\tilde{\mathbf{A}} - \mathbf{RW}^{-1}]^T). \tag{5.9}$$

where $\|\cdot\|_F$ is the Frobenius norm. In this equation $\mathbf{D}$ is discarded since $\mathbf{W}$ is normalized after each iteration. This constraint is then incorporated into the main BSS cost function resulting in an unconstrained problem for finding $\mathbf{W}$. After each iteration the estimates of $\mathbf{R}$ and $A_{uk}$ are also updated based on the procedure described in the next section. In order to locate the sources more accurately the

non-homogeneity of the head region has to be exploited. In the proposed method a non-homogeneous head model has been used in order to compute the nonlinearity parameter and accurately estimate the geometrical locations of the sources from the estimated mixing matrix$\tilde{\mathbf{A}} = [A_k \vdots A_{uk}]$.

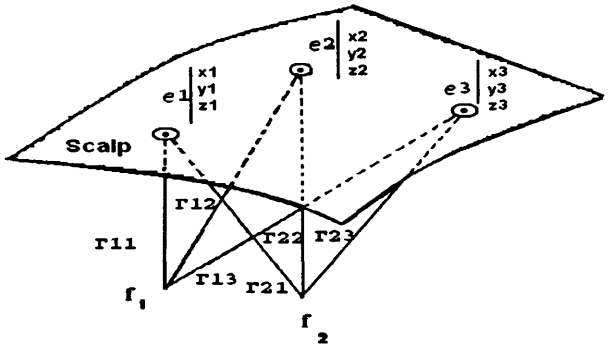



Figure 5.4: A section of the scalp including three electrodes and locations of the two sources

A number of recently developed techniques such as time-lagged second-order blind identification (SOBI) [138] or as in [122] can better cope with nonstationarity of the data. SOBI can separate functionally distinct neuronal signals from each other and from other noise sources under poor signal to noise ratio (SNR). SOBI is also able to recover those components that are physilogically and neuroanatomically interpretable [138] [139]. On the other hand, the signals may be considered stationary within short segments of about 10 seconds (or about 2000 samples). SOBI has been selected as the best BSS system and modified so $\mathbf{W}$ can be estimated iteratively by jointly diagonalizing the cross-correlation function for a number of lags. As described in [135] $\mathbf{y}(t)$ is the output of the estimated original sources in equation (1.2). The separation matrix can be found by finding the global minima of a cost function $J(\mathbf{W})$, which provides a measure of independence of the estimated sources. Hence minimizing $J(\mathbf{W})$ will make sure that the estimated sources are as independent as

possible. The covariance matrix $\mathbf{R}_Y$ to be diagonalized is given by

$$\mathbf{R}_Y = \mathbf{W}[\mathbf{R}_X + \mathbf{R}_V]\mathbf{W}^T \tag{5.10}$$

where $\mathbf{R}_X$ is the estimation of the covariance matrix of the signal mixtures over the current block and $\mathbf{R}_V$ is the covariance matrix of the noise. Since it has been assumed that the noise is uncorrelated $\mathbf{R}_V$ will be a diagonal matrix [96]. Following [96], the LS estimate is

$$J(\mathbf{W}) = arg \min_{\mathbf{W}} \sum_{t=1}^{T} \|E(t)\|_F \tag{5.11}$$

where $\|\cdot\|_F$ is the Frobenius norm and $E$(t) is the error to be minimized between the covariance of the source signals $\mathbf{R}_S$ (diagonal due to independence) and the estimated sources $\mathbf{R}_Y$. Therefore a suitable cost function is defined that minimizes the off diagonal elements, defines as

$$J_m(\mathbf{W}) = arg \min_{\mathbf{W}} J_M(\mathbf{W}) = arg \min_{\mathbf{W}} (offdiag(\mathbf{R}_y))^2 \tag{5.12}$$

where $offdiag[\cdot]$ means off-diagonal elements of a matrix. Therefore, to find $\mathbf{W}$ and $\mathbf{R}$ one can add a constraint to the main cost function and solve the following unconstrained problem:

$$\mathbf{J}(\mathbf{W}) = \mathbf{J}_m(\mathbf{W}) + \lambda \mathbf{J}_c(\mathbf{W}) \tag{5.13}$$

where $\mathbf{J}_m(\mathbf{W})$ is defined in (5.12) the main least-squares BSS cost function, $\mathbf{J}_c(\mathbf{W})$ is the constraint defined by equation (5.9), and $\lambda$ is the Lagrange multiplier. To minimize equation (5.13) the following update is used.

$$\mathbf{W}_{t+1} = \mathbf{W}_t - \mu \nabla_{\mathbf{w}} \mathbf{J} \tag{5.14}$$

where $\mathbf{J}$ is defined in (5.13), $\mu$ is the learning rate, and $\nabla$ denotes the gradient operator. Therefore we have

$$\mathbf{J}_c = tr(\tilde{\mathbf{A}} - \mathbf{R}_t \mathbf{W}_t^{-1})(\tilde{\mathbf{A}}^T \mathbf{R}_t^T \mathbf{W}_t^{-1^T}) \tag{5.15}$$

expanding the above

$$\mathbf{J}_c = tr(\tilde{\mathbf{A}}\tilde{\mathbf{A}}^T - \tilde{\mathbf{A}}\mathbf{R}_t^T\mathbf{W}_t^{-1^T} - \mathbf{R}_t\mathbf{W}_t^{-1}\tilde{\mathbf{A}}^T + \mathbf{R}_t\mathbf{W}_t^{-1}\mathbf{R}_t^T\mathbf{W}_t^{-1^T})$$

Let assume

$$\mathbf{U} \triangleq \mathbf{W}_t^{-1} \tag{5.16}$$

Using

$$\frac{d}{dX}(Y^{-1}) = -Y^{-1}\frac{d}{dX}(Y)Y^{-1} \quad and \quad Y = X$$

Therefore

$$\frac{\partial \mathbf{U}}{\partial \mathbf{W}_t} = -W_t^{-2} \tag{5.17}$$

Also from matrix manual

$$\frac{\partial \mathbf{J}_c}{\partial \mathbf{W}_t} = (\frac{\partial \mathbf{J}_c}{\partial \mathbf{U}})(\frac{\partial \mathbf{U}}{\partial \mathbf{W}_t}) \tag{5.18}$$

$$\frac{\partial \mathbf{J}_c}{\partial \mathbf{U}} = -\tilde{\mathbf{A}}\mathbf{R}^T - \tilde{\mathbf{A}}^T\mathbf{R} + \mathbf{R}^T\mathbf{U}^T\mathbf{R} + \mathbf{R}\mathbf{U}^T\mathbf{R}^T \tag{5.19}$$

replacing 5.16,5.17, and 5.19 into 5.18:

$$\nabla_{\mathbf{w}}\mathbf{J}_c = \tilde{\mathbf{A}}_t\mathbf{R}_t^T\mathbf{W}_t^{-2} - \tilde{\mathbf{A}}_t^T\mathbf{R}_t\mathbf{W}_t^{-2} - \mathbf{R}_t^T\mathbf{W}_t^{-1^T}\mathbf{R}_t\mathbf{W}_t^{-2} - \mathbf{R}_t\mathbf{W}_t^{-1^T}\mathbf{R}_t^T\mathbf{W}_t^{-2} \tag{5.20}$$

(when $m \neq n$, $\mathbf{W}^{-1}$ will be the pseudo inverse of $\mathbf{W}$) weighted by $\lambda$ and added to the gradient of $\mathbf{J}_m(\mathbf{W})$. Furthermore, within the same coupled iteration loop the permutation matrix $\mathbf{R}$ and $A_{uk}$ are updated through the following equations:

$$\mathbf{R}_{t+1} = \mathbf{R}_t - \gamma\nabla_R(\mathbf{J}_c) \tag{5.21}$$

where

$$\nabla_R(\mathbf{J}_c) = 2\mathbf{R}_t(\mathbf{W}_{t+1}^T\mathbf{W}_{t+1})^{-1} - 2\tilde{\mathbf{A}}_t(\mathbf{W}_{t+1}^{-1})^T \tag{5.22}$$

is the gradient of $\mathbf{J}_c$ with respect to $\mathbf{R}$, and

$$A_{uk_{t+1}} = A_{uk_t} - \zeta\nabla_{A_{uk}}(\mathbf{J}_c) \tag{5.23}$$

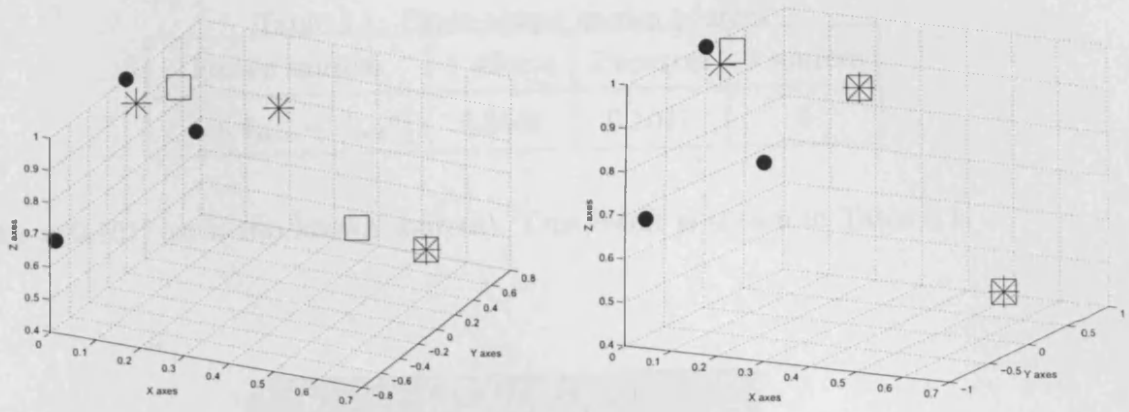
where

$$\nabla_{A_{uk}}(\mathbf{J}_c) = 2(A_{uk_t} - [\mathbf{R}_{t+1}(\mathbf{W}_{t+1}^{-1})]:, k+1:m) \tag{5.24}$$

is the gradient of $\mathbf{J}_c$ with respect to $A_{uk}$, and $\gamma$ and $\zeta$ are the learning rates (Matlab notation is used to denote the last $m-k$ columns in the right most term in equation (12)). After estimating $\mathbf{W}$ in each iteration the rotation matrix $\mathbf{R}$ and $A_{uk}$ are also iteratively calculated in a coupled loop. Consequently $\tilde{\mathbf{A}} = [A_k \vdots A_{uk}]$ is a good estimation of the mixing matrix and location of the sources. Sequential iteration of equations (5.14), (5.21), and (5.23) yields a robust solution to the ill-posed localization problem. Accurate selection of the learning rates $\mu, \gamma$, and $\zeta$, ensures simultaneous convergence of the algorithm. The stopping condition is governed by a proper threshold on the norm of $\mathbf{W}_{t+1} - \mathbf{W}_t$. After each iteration they are automatically updated based on $\mu_{t+1} = \mu_0\frac{norm(\mathbf{W}_{t+1}-\mathbf{W}_t)}{norm(\mathbf{W}_t)}$, $\gamma_{t+1} = \gamma_0\frac{norm(\mathbf{R}_{t+1}-\mathbf{R}_t)}{norm(\mathbf{R}_t)}$, and $\zeta_{t+1} = \zeta_0\frac{norm(\mathbf{A}\mathbf{u}_{t+1}-\mathbf{A}\mathbf{u}_t)}{norm(\mathbf{A}\mathbf{u}_t)}$, where $norm(\cdot)$ represents sum squared of the matrix entries. The proposed algorithm was implemented using second order blind identification (SOBI) [136] [124] , which has been recently validated for its ability to recover correlated neuronal sources [125], followed by localization of the EEG sources as described earlier.

## Localization of synthetic sinusoidal sources

A matrix of three signals containing three synthetic sinusoidal sources with specific geometrical locations were tested. The mixing matrix $\mathbf{A}$ is modelled once with three




(a) (b)

Figure 5.5: The estimated source locations when (a) there is only one out of three sources known, and (b) two sources are known. "•" represents the sensors locations, "∗" shows the actual locations of sources, and "□" represents the estimated locations of the sources. x, y, z are towards the front, lateral-right, and plannar views.

EEG electrodes via homogenous medium and another time via non-homogenous medium. Selection of matrix **A** was based on true geometry of the head model and the EEG 10-20 sensor positions. Accurate selection of the learning rates $\mu, \gamma$, and $\zeta$, ensures simultaneous convergence of the algorithm. The stopping condition is governed by a proper threshold on the norm of $\mathbf{W}_{t+1} - \mathbf{W}_t$. $\mu, \gamma$, and $\zeta$, are initilized to $\mu_0 = \gamma_0 = \zeta_0 = 0.001$ and after each iteration are automatically updated. The result of separation and localization were the same for both cases. The column vectors of the estimated mixing matrix, A, refer to the coordinates of the sources. The actual locations can be easily derived using the LS based sphere method [110]. In Figure 5.5(a), the original and simulated locations for one known source and, in Figure 5.5(b) with two known sources (for synthetic sources), are depicted. Using the SOBI BSS algorithm, the geometrical error $\epsilon$ ( $\epsilon = \|A_{new} - A_{old}\|_2^2$) is found to be less when the number of sources increases ($\epsilon = 0.8648$ with one known source

Table 5.1: Error verses known sources

| known sources | 1 source | 2 sources | 3 sources |
|---|---|---|---|
| $\epsilon = \|A_{new} - A_{old}\|_2^2$ | 0.8648 | 0.1017 | 0 |

and $\epsilon = 0.1017$ with two known sources). This result is shown in Table 5.1.

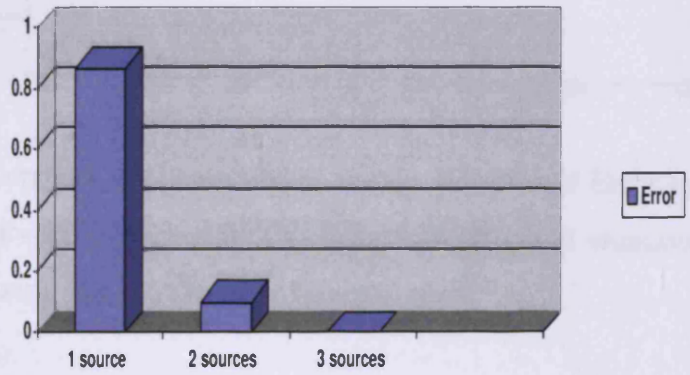



Figure 5.6: Error versus the number of known sources

The convergence plots for unmixing matrix **W**, is shown in Figure 5.7. Also, the LS convergence is convex and fast.

## Localization of the sources within a nonhomogeneous medium

We are required to generate a head model with effect of the degradation process as space-invariant or the distorted calibration caused by non-homogenity of the different layers between the sources and the sensors and to localize the actual location of the sources from the sensors.

To visualize this we have generated a head model as a three-layer sphere with three sources s1, s2, s3 and three sensors x1, x2, x3 in which each layer has its own

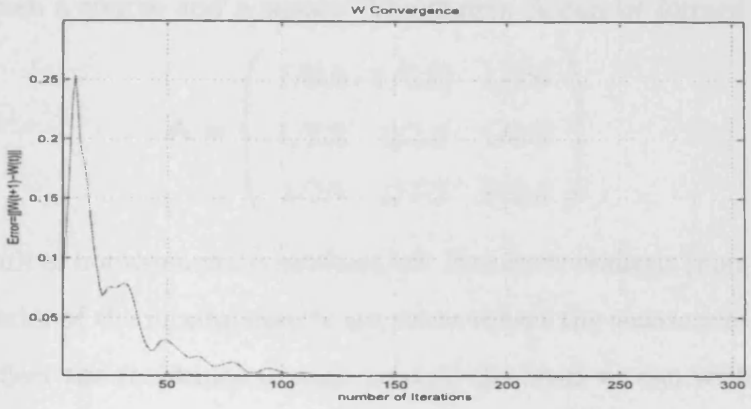



Figure 5.7: Convergence of the unmixing matrix $\mathbf{W}$; Plot of Error against number of iteration, Error=$\|\mathbf{W}(t+1) - \mathbf{W}(t)\|_F$, where $\mathbf{W}$(t) is old unmixing matrix, and $\mathbf{W}$(t+1) is new unmixing matrix after each iteration.

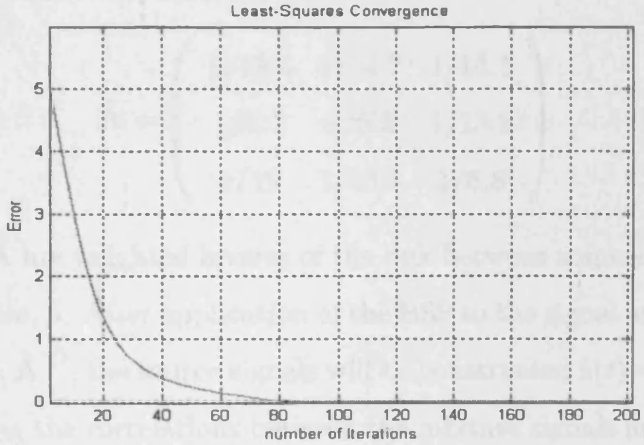



Figure 5.8: Least sqaures convergence plot, Error = $\|\mathbf{f} - \mathbf{d}\|_2^2$, where $\mathbf{f}$ is the source and $\mathbf{d}$ is inverse of correlations between the estimated source and electrode position.

resistivity. Each element of $\mathbf{A}$, the mixing matrix, is inversely proportional to the distance between a source and a sensor. The matrix $\mathbf{A}$ can be formed as

$$\mathbf{A} = \begin{pmatrix} 1/8.9 & 1/1.93 & 1/9.6 \\ 1/2.2 & 1/2.9 & 1/6.9 \\ 1/10 & 1/7.2 & 1/2.5 \end{pmatrix} \tag{5.25}$$

This is the result of homogeneity consideration. In a more realistic (non-homogenous) model, the entries of the mixing matrix are set to reflect the resistances of the layers. In order to reflect the resistance of each layer of the head we can write:

$$R = \ell_1\rho_1 + \ell_2\rho_2 + \ell_3\rho_3 \tag{5.26}$$

where $\rho_1$ is the resistivity of brain, $\rho_2$ is the resistivity of skull, $\rho_3$ is the resistivity of scalp, and $\ell_1, \ell_2, \ell_3$ are respectively thicknesses of brain, skull, and scalp. For simplicity, let assume $\rho_1 = \rho_3 = 1$ and $\rho_2 = 10$ so, according to concentric head model of Rush and Discoll (1969) shown in Figure 4.16. One can write the mixing matrix $\mathbf{A}$ with the effect of resistivity as:

$$\hat{\mathbf{A}} = \begin{pmatrix} 1/13.4 & 1/14.7 & 1/14.1 \\ 1/6.7 & 1/8.3 & 1/13.2 \\ 1/19 & 1/13.5 & 1/8.8 \end{pmatrix} \tag{5.27}$$

The elements of $\hat{\mathbf{A}}$ are weighted inverse of the link between sources and sensors with effect of resistivities, $\rho$. After application of the BSS to the signal and iterative calculation of the $\mathbf{W} = \hat{\mathbf{A}}^{-1}$, the source signals will be constructed $\hat{\mathbf{s}}(t) = \mathbf{y}(t) = \mathbf{W} * \mathbf{x}(t)$. Now by calculating the correlations between the mixture signals and the major contributed independent components the distance between source signals and the signals observed at the sensors can be calculated. Hence the locations of the sources can be obtained by LS algorithm.

Here the distance between the original location of sources and the location of sources with effect of the nonlinear factor $\alpha$ described in equation (4.10) as $error - distance$. In a Matlab program for $\alpha$ of values range between $0.1 - 2$ and related $error - distance$ between the original location of sources and the location of sources with the effect of $\alpha$ were calculated. The experiment tries to find the best value of $\alpha$, which gives the minimum $error - distances$. The result was that $\alpha = 0.5$ gives minimum error distance value as shown in Figure 5.9. The distance between sources and

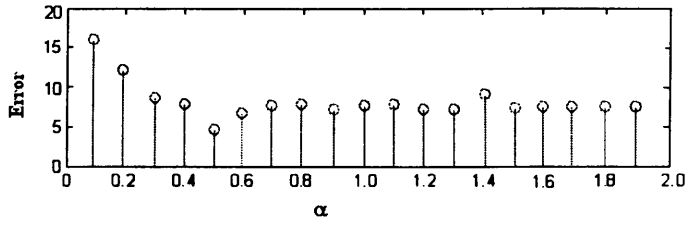



Figure 5.9: $error - distance$ plot against values of $\alpha$ in (4.10) described in section "Non-homogeneity considered as Nonlinearity of the head" for homogeneous spherical model of the brain

electrodes are inversely proportional to the signals' correlations (i.e. $d = 1/C_{ij}$). A non-homogeneous model can be generated by adding uniformly distributed noise to the some synthetic signals, shown in Figure 5.11(a). After application of BSS and re-scaling and re-permuting the original condition of the sources as shown in Figure 5.11(b), one can find the correlation and related contributor to each location and hence with the help of LS find the location of the source. At this stage there are the original location of the sources and the new locations after BSS application with the effect of the noise. The experiment requires to find a value for $\alpha$ which gives minimum $error - distance$ between the original source location and the displacement caused by nonlinearity of the noise after BSS application.

| Alpha | Error | Alpha | Error |
|---|---|---|---|
| 0.1 | 16.1196 | 1.1 | 7.9847 |
| 0.2 | 12.0628 | 1.2 | 8.0418 |
| 0.3 | 8.5218 | 1.3 | 8.5264 |
| 0.4 | 6.9677 | 1.4 | 8.6331 |
| 0.5 | 6.0998 | 1.5 | 8.8181 |
| 0.6 | 7.5027 | 1.6 | 8.984 |
| 0.7 | 7.5995 | 1.7 | 9.1018 |
| 0.8 | 7.6507 | 1.8 | 8.3089 |
| 0.9 | 7.7414 | 1.9 | 8.5124 |
| 1.0 | 7.8787 | 2.0 | 8.9145 |

Table 5.2: Error values for a set of $\alpha$ for homogeneous spherical model of the brain

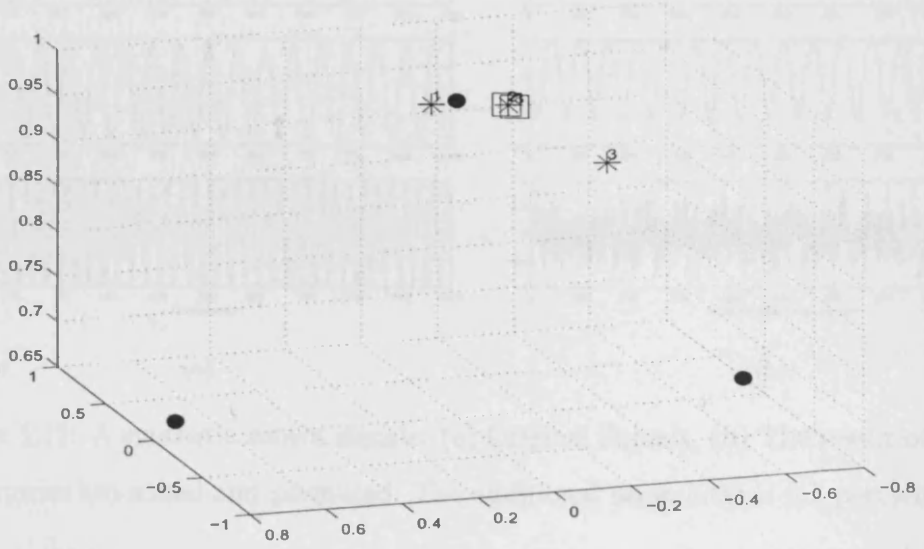
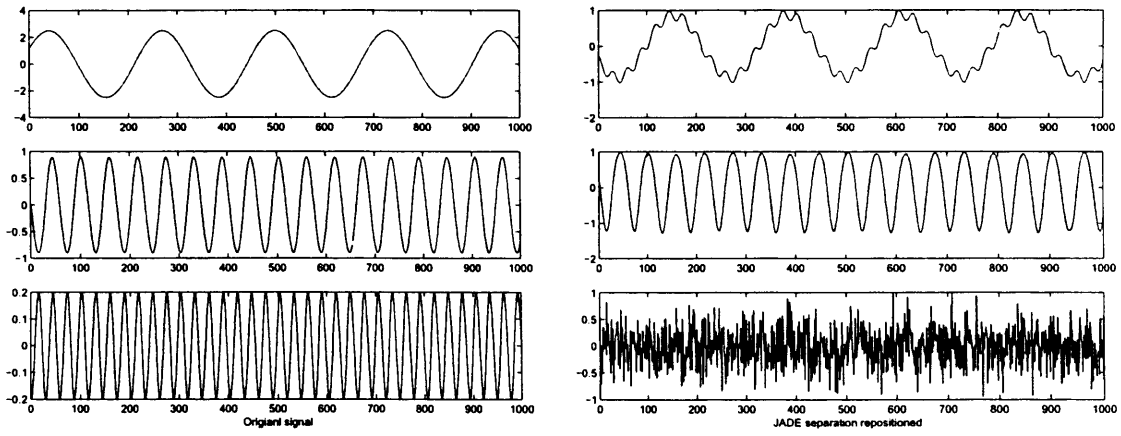



Figure 5.10: The estimated sources using BSS are shown in by " □", which are overlapping the " ∗" representing the actual sources, and "•" represent the locations of the sensors for a homogeneous spherical model of the brain

Figure 5.10 shows the estimated location of the sources without the effect of nonlinearity, where in Figure 5.13 and related *error – distance* shown in Figure 5.12 by setting the value of $\alpha$ *to* 0.7 a better source estimation can be observed.
The *error – distance* between each estimated source location for different values of $\alpha$ ranging from 0.1 to 1.4 with the interval of 0.1 with added noise level of 2 percent are shown in Figure 5.12. Figure 5.14 shows similar results for the added noise level of 5 percent. In Figure 5.14 an increase in overall *error – distance* values can be observed, in comparison with overall *error – distance* value shown in Figure 5.12, which are caused by an increase in the correlations between the mixed signals and the noise.




(a) (b)

Figure 5.11: A synthetic source signals: (a) Original Signals, (b) The result of BSS; the sources are scaled and permuted. The additional noise level is 0.1 percent.

By increasing the noise level in the mixture signal the *error – distance* increases too. Consequently, given the knowledge about the nonlinearity parameter $\alpha$ and noise level the distance error can be obtained.

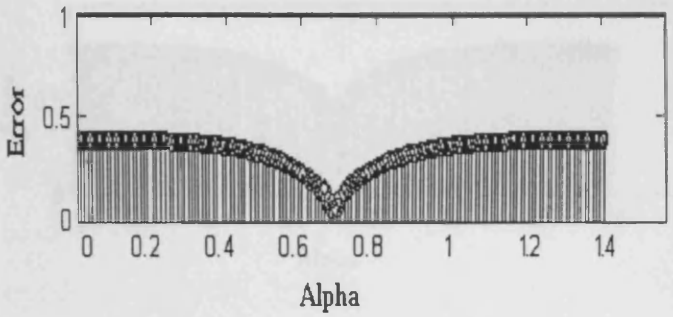



Figure 5.12: $error - distance$ between the original and estimated source locations with the effect of 2 percent noise added to the mixture signals. An average value of $\alpha = 0.7$ gives a minimum $error - distance$.

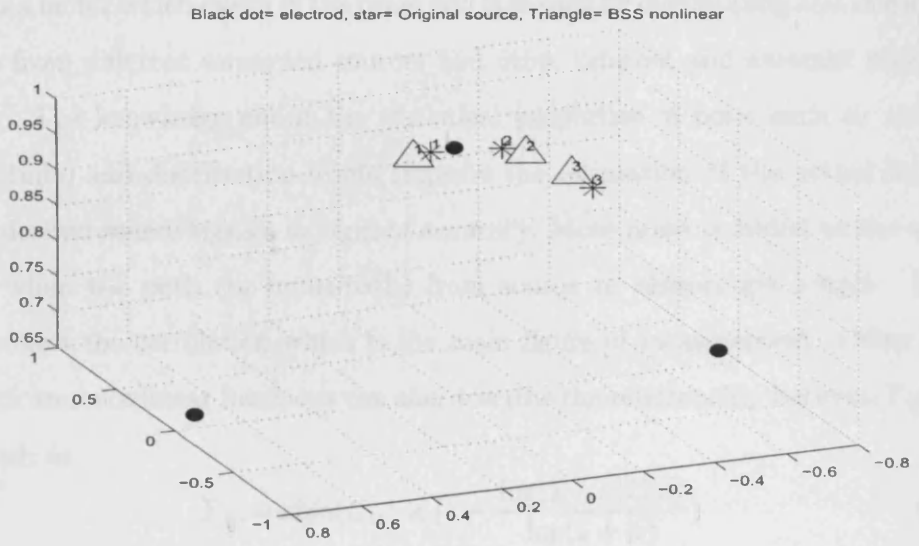



Figure 5.13: The effect of nonlinearity parameter $\alpha$ and $error-distance$ of estimated sources: the location of the original source is shown in star, the triangles indicating the noisy estimation of the sources with nonlinear parameter value of $\alpha = 0.7$. The black points are the electrode locations. The numbers shown on each location indicates the original order of each signal.

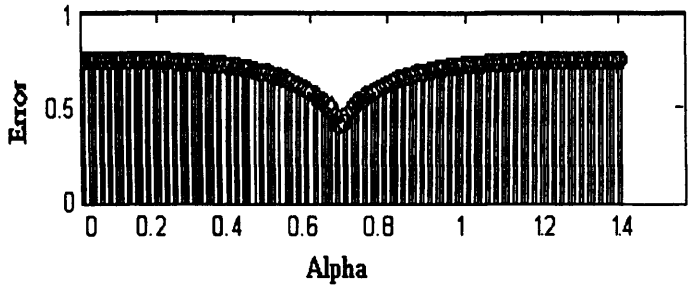



Figure 5.14: Distance error between original and estimated first, second, and third source locations respectively, with the effect of 5 percent noise to the mixture signals. An average $\alpha = 0.7$ gives a minimum $error - distance$ between the original source and the estimated one.

Noise is a factor which exists in the brain and is caused by overlapping and mixing the signals from different unwanted sources and other internal and external undesired signals. The knowledge about the statistical properties of noise such as variance (magnitude) and distribution would improve the estimation of the actual location of the desired source signals in term of accuracy. More noise is added to the source signal when the path (or multi-path) from source to sensors get longer. Hence this reduces the correlation which is the main figure of measurement. Other more sophisticated nonlinear functions can also describe the relationship between $\Gamma_{ij}$s and $C_{ij}$s such as:

$$\Gamma_{ij} = sign(C_{ij}) \times (1 - \frac{\log(1 + \alpha|C_{ij}|)}{\log(1 + \alpha)}) \tag{5.28}$$

or

$$\Gamma_{ij} = 1 - \exp(\alpha \times C_{ij}) \tag{5.29}$$

These models were also implemented but by using $\Gamma_{ij} = \frac{1}{C_{ij}^{\alpha}}$ and setting $\alpha = 0.7$, the best localization result was achieved.

## Localization of real EEG signals

For localization, the EEG signals for focal epilepsy with confirmed epileptic foci and normal EEG with both Delta and Alpha rhythms were used. The electrodes were set up according to [137]. The electrode location of the known source with its peak frequency is shown in Figure 5.15. After the artifact removal technique used in [135]. The estimated location of unknown sources can be observed in Figure 5.16(a).

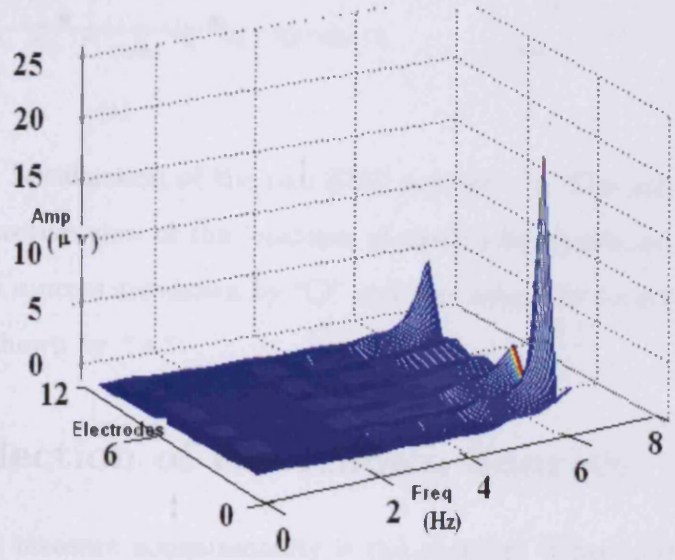
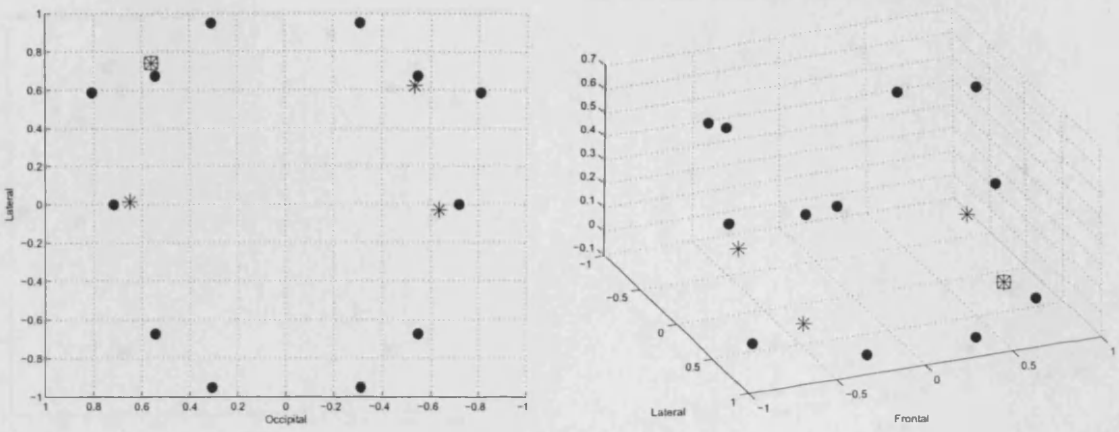



Figure 5.15: Selection of highest amplitude level in Alpha rhythm as for the known source

The estimated location of unknown sources based on calculation of nonlinear parameter $\alpha = 0.7$, can be observed in Figure 5.17(a) and its lateral view shown in Figure 5.17(b) for better observation of three dimensional clarity. The result of the dipole method, as in Figure 5.18, represents dipoles in different colors. Hence illustrates that the proposed method can localize most of the source.




(a) (b)

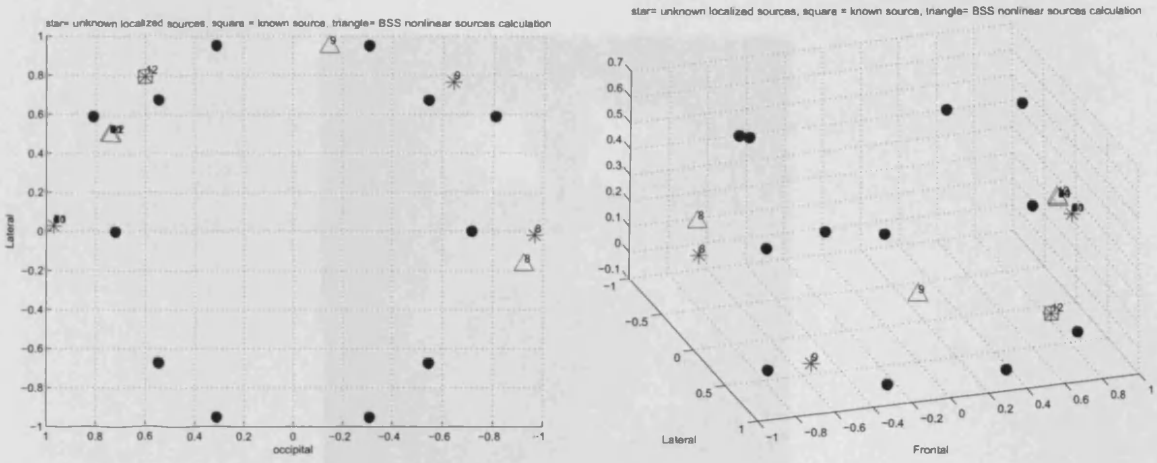
Figure 5.16: Localization of the real EEG sources (a) The top view, and (b) The lateral perspective view of the locations of electrodes shown by " • ", the locations of the known sources are shown by "□" and the estimated locations of the unknown sources are shown by " ∗ "

## 5.3 Selection of the Known Sources

A method to measure nongaussianity is the absolute value of the fourth-order cumulant at origin or kurtosis, hence it also can help to localize the location of the sources on the surface of the head. The kurtosis of a random signal $y$ shown in equation (7.1):

$$Kurt(y) = E\{y^4\} - 3(E\{y^2\})^2 \tag{5.30}$$

where $E\{.\}$ denotes statistical expectation and, it is assumed that the random variable has zero mean. The absolute value of the kurtosis is zero for Gaussian variable and greater than zero for nongaussian variables. By projection of the kurtosis measured from sensors, to the scalp, some of the brain source signals appear to have high peak in certain frequency bands. These can be considered as known sources. Initially the ICs are filtered in Delta, Theta, Alpha, and Beta bands and their kur-




(a) (b)

Figure 5.17: Localization of the real EEG sources with the effect of non-homogeneity; (a) the top view, and (b) the lateral perspective view of the locations of electrodes shown by " • ", the location of the known source shown by "□" and the estimated locations of the unknown sources shown by " ∗ ". The calculated nonhomogeneous (nonlinear) effect on establishing an estimated location for the unknown sources are shown in " △ ". Each location is marked with a number and some of the sources found are overlapping

toses in frequency domain were measured. In the next step the kurtoses of the frequency domain signals are back projected to the scalp in each band to provide the topographic map of these variables as shown in Figure 5.19. In Figure 5.19, the main periodic source starts at lower Theta (higher Delta). The harmonics of the wave appear at the Alpha and Beta bands, ignoring the above harmonics in Alpha band, there are two prominent Alpha sources which can be considered as the known sources.

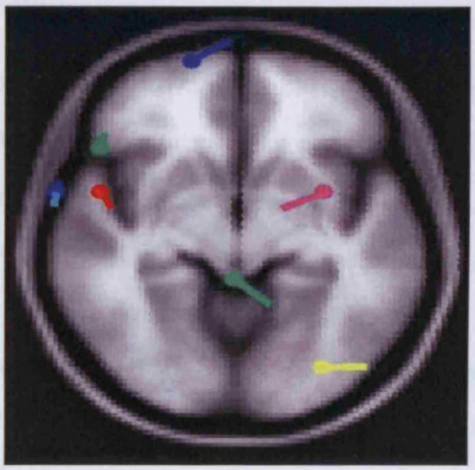


Figure 5.18: Implementation of the dipole fitting method from the EEG lab, which confirms the location of some of the sources computed using the proposed partially constrained method.

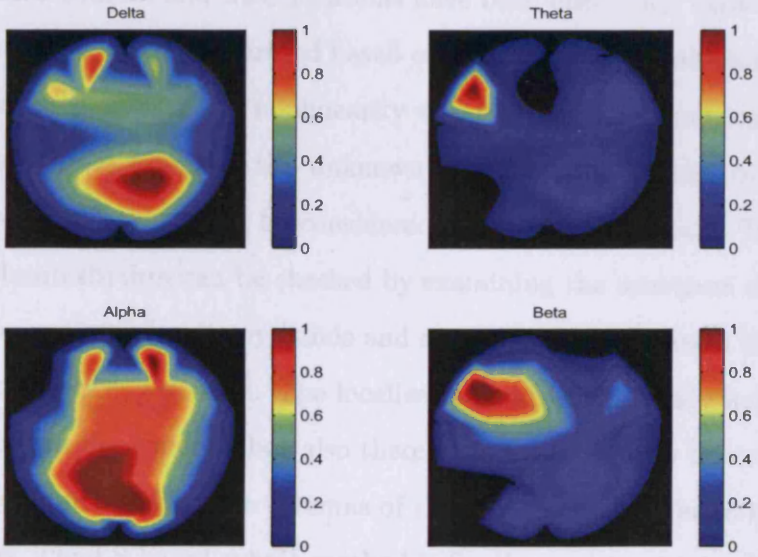



Figure 5.19: Projection of kurtosis to the scalp topography. The red patch in Beta range is the harmonic of the signals in Theta range.

### 5.3.1 Conclusions

In this chapter the work was to develop a partially constrained BSS algorithm to exploit the location of the known sources to separate and localize the unknown ones. Here we iteratively calculated the Rotational matrices $\mathbf{R}$, and $\mathbf{W}$ and $\mathbf{A}$ simultaneously. The reason for selecting SOBI as the separation algorithm was because: SOBI [138] [122] can better cope with nonstationarity of the data. SOBI can separate functionally distinct neuronal signals from each other and from other noise sources under poor signal to noise ratio (SNR). SOBI is also able to recover the components that were physilogically and neuroanatomically interpretable [138] [139]. On the other hand, the signals may be considered stationary within short segments of about 10 seconds (or about 2000 samples). The SOBI has been selected as the best BSS approach and modified so that $\mathbf{W}$ can be estimated iteratively by jointly diagonalizing the cross-correlation function for a number of lags. Then, a priori knowledge about the known EEG sources and their locations have been effectively exploited in localization of the other sources separated based on a partially-constrained BSS method. This method also exploits the nonlinearity effect of head conductivity and noise, and estimates the locations of the unknown sources. The normal brain rhythms with given source locations can be considered as the known sources. The existence of a normal brain rhythm can be checked by examining the spectrum of the signals in various conventional frequency bands and checking the kurtoses of the estimated sources in the frequency domain. The localized sources are not only independent of the pre-calculated head model, but also there is no need to know how deep or near to skull sources are located. The columns of the estimated $\mathbf{A}$, refer to the locations of the sources. The LS based sphere method is finally used to obtain the estimated geometrical locations. The accuracy of the results increases with increase in the number of known sources. The approximation error in respect to source location

is reduced by considering at least four intersecting spheres in the three-dimensional space.

Any prior information about the statistics and nature of the noise would benefit the estimation of locations of the desired sources. More noise is added to the source signal when the path (or multi-path) from source to sensors get longer. Consequently this reduces the correlation which is the main figure of measurement in this approach.

# Chapter 6

# SUMMARY, CONCLUSIONS, AND FUTURE WORK

## 6.1 Summary

Localization of the brain signal sources using the scalp EEG signal can be accomplished by mathematical analysis and consideration of physiological information. To achieve this it is also required to exploit sufficient information from the EEG signals in-order to select the right solutions when the problem is ill-posed. An ill-posed problem refers to infinite solutions, or no unique solution from the mixture signals received at the electrode. The ill-posed problem can be solved by reducing the number of solutions, by converting an undetermined system to a determined number of mixtures. Hence specifying some priori knowledge to the system. In order to classify the known sources as priori knowledge, some specification about the characteristic of the sources within the brain such as filtering within a certain frequency band and expected geometrical location of certain sources can be taken into account.

After identifying the related frequency band, and geometrical hypothesis location

of the sources, it is useful to identify the regions of interest by constructing topographical plots of the power in a given frequency band. The topographical plots are important for assessing the locations of the sources.

An initial development of the algorithms incorporating an iterative application of the algorithm using high order statistics (HOS) was able to indicate successfully the number of MCS from the EEG sources within the brain by exploiting the geometrical information about the electrodes and separate them into their constituent components. In this work it was assumed that the system is overdetermined but with an unknown number of sources. Decomposition of the signals into four conventional frequency bands, not only ensured that the number of sources remained smaller than the number of sensors but in an overall term the process mitigated the permutation of the independent components. In other words, priori knowledge about the geometrical and physiological aspects of the acquired EEG recordings, helped to impose some constraints on the mixing/unmixing matrix. Thus, it is expected to estimate the entries of the mixing/unmixing matrices in a way to increase the accuracy of the results. Also, we resolved the inherent scaling and permutation ambiguities of the BSS algorithm through incorporating such constraints. Localization based on instantaneous conditions of sources, usually has higher localization errors on independent sources in the presence of noise. On the other hand, considering only spatial inverse methods may also give error in the location of the sources when continuous EEG recordings are processed. Implementing the SOBI algorithm, with its time lags diagonalization characteristics for its update equation helped to overcome this problem.

In the next step of this work, a priori knowledge about the known EEG sources and their locations was effectively exploited in localization of the other sources separated

based on a partially-constrained BSS method. Unlike dipole based methods the proposed algorithm is computationally cost effective and is insensitive to noise (noise is either separated or cancelled out due to the inherent properties of BSS). The normal brain rhythms with given sources locations can be considered as the known sources. The known sources may also be generated synthetically. The existence of a normal brain rhythm can be checked by examining the spectrum of the signals in various conventional frequency bands. In the algorithm the unknown sources as well as the permutation matrix and the unmixing matrix were determined. The columns of the estimated $\mathbf{A}$, then, refer to the locations of the sources. The LS based sphere method was finally used to obtain the estimated geometrical locations. The accuracy of the results increased with the number of known sources.

Application of the overall procedure leads to both detection and localization of the comprising normal and abnormal EEG sources for each EEG sub-band. However from the experiments carried out above the developed algorithm effectively localized the sources in an non-homogeneous environment.

## 6.2 Conclusions

Research here was an attempt to localize the brain sources using BSS. The main drawbacks of the approach are due to:

1. The nature of the brain sources; unfortunately, there hasn't been any established findings in whether the sources are uncorrelated or independent; are they stationary or non-stationary, synchronous or not, etc. Therefore, here we tried to focus on separation and localization of the abnormal sources, given some knowledge about the known sources.

2. The number of the sources; many criteria for determination of the number of sources fails here mainly due to the right noise level and nonstationarity of the signals. Implementation of the new iterative minimization of the Akaike criterion [42] plus testing the behavior of the selected sources in both frequency and space gave some promising conclusions.

3. The effect of noise; this is probably the most troublesome phenomenon in the localization process based on the BSS. In this work we were hopeful that by exactly estimating the number of sources we could pave the path for efficiently separating the signal and noise subspaces using the corresponding independent components. Although some statistics of noise such as distribution can be identified using this method, but a robust solution to this problem is still under question.

## 6.3 Future work

Although the algorithms developed here were able to separate and localize some abnormalities from the EEG, future developments are:

1. Further exploitation of information from the EEG signals is required, in-order to distinguish the signals originated from inside the brain and those from outside (i.e. undesired signals) with the aid of a priori knowledge i.e. specification / characteristics of the sources within the brain. By applying nonlinear neurons in Neural Networking, the neuron is trained to learn the specification and characteristics of the sources of the brain. After training, the neuron is capable of classifying the undesired signals.

2. A time-frequency approach further enhances the performance of the localization since it incorporates the physiological aspects and characteristics of the

signals into the separation and localization algorithm. In subband BSS

- The system working with each band is less likely to be under-determined, since the sources of normal and abnormal rhythms are represented in distinct time-frequency domain.
- Implement BSS in combination with a wavelet transform.

3. Incorporate more robust method for the non-homogeneity of the head within the definition of non-linear normalization of the correlations. Furthermore, it is known that even the thickness of layers of the head are not uniform. Consequently, a true model of the head will be required to be constructed to precisely define the non-linear parameter, in this case a measure of independence of the source rather than correlations between sources may result in better localization.

4. Sensitivity distribution of EEG electrodes: Locating the electrodes closer and closer to each other causes the current between electrodes to flow more and more at the surface of the skin region, without passing through the skull, and finally into the brain region, decreasing the sensitivity to the brain region [97]. Therefore, the number of electrodes related to specific requirement in each diagnostic case should be studied in order to give better results.

5. A hybrid system including the methods based on dipole assumptions, and constrained BSS may be an optimum solution to the localization problem.

6. Fusion of fMRI into the EEG-based neuro-imaging requires further research in order to enhance the localization performance. Such a multi-modal system can optimize both space and time resolutions to enhance the performance of the localization algorithm.

# Chapter 7

# APPENDICES

## Appendix A

### Information theory

There are three different approaches, unified by the information theory;(a)Minimization of mutual information,(b)Maximum likelihood estimation (MLE) which is the same as Infomax, and (c)Maximization of Negentropy.

#### Mutual information

Entropy is a measure of uncertainty in the signal and is expressed in terms of bits in digital domain. It is worth emphasizing here that entropy is not a measure of the "disorder" or "mixed-upness" of a system, though those terms are often used (rather loosely at times) to describe entropy due in part to this statistical interpretation.

$$\mathbf{H(x)} = -\int p(\mathbf{x}).\log[p(\mathbf{x})].dx = -E\{\log[p(\mathbf{x})]\}$$
$$\mathbf{H(x)} = -\sum_i p_i(\mathbf{x}).\log[p_i(\mathbf{x})] \qquad (A.1)$$

where p(x) is probability distribution of x, and joint entropy is defined as:

$$\mathbf{H}(y_1,\ldots,y_N) = H(y_1) + H(y_2) + \cdots + H(y_N) - I(y_1, y_2, \ldots y_N)$$
$$\Longrightarrow I(y_{(1,\cdots,y)}) : \ mutual \ \ information \quad (A.2)$$

Mutual information is the amount of information that a random variable, x, contains about another random variable, y( and vice-versa). I(x,y)=0, if x,y are independent, otherwise $I(x,y) > 0$. Minimizing the mutual information is equivalent to maximizing the joint entropy.

**MLE ( such as Infomax)**

These algorithms are based on stochastic gradient ascent of the contrast function.

$$\mathbf{u}(t) = \mathbf{W}\mathbf{x}(t) \qquad (A.3)$$

where $\mathbf{W}$ is to be estimated in order to separate $\mathbf{x}$ into a set of independent component signals $\mathbf{u}$. $\mathbf{u}_i$s are independent if

$$\mathbf{P_U(U)} = \prod_i \mathbf{P}_i(\mathbf{u}_i); independency \qquad (A.4)$$

where $p_{\mathbf{U}}(\mathbf{U})$ and $p_i(\mathbf{u}_i)$ are respectively joint probably density function (pdf) and marginal pdf of set $\mathbf{U}$ and independent component $\mathbf{u}_i$.

Log-likelihood function is defined as: $\mathbf{L(u,W)} = log|\det(\mathbf{W})| + \sum_i \log[\mathbf{P}_i(\mathbf{u}_i)]$. Consequently, the gradient of this function is given as:

$$\frac{\partial}{\partial \mathbf{W}}\mathbf{L}(\mathbf{U},\mathbf{W}) = (\mathbf{W}^T)^{-1} + \sum \frac{\frac{d\mathbf{p}_i(\mathbf{u}_i)}{d\mathbf{u}_i}}{\mathbf{p}_i(\mathbf{u}_i)}.\mathbf{x} \qquad (A.5)$$

A natural gradient approach (i.e. the gradient is multiplied by $\mathbf{W}^T\mathbf{W}$) for optimization then, results in the following update [26]:

$$\Delta\mathbf{W} = (\mathbf{I} - \frac{\frac{d\mathbf{p}_i(\mathbf{u}_i)}{d\mathbf{u}_i}}{\mathbf{p}_i(\mathbf{u}_i)}.\mathbf{U}^T).\mathbf{W} \qquad (A.6)$$

## Negentropy

The whitening facilitates the separation of the underlying independent signals. The observed data are made uncorrelated and unit-variance.

The whitening may be accomplished by PCA projection: $\mathbf{z}(t) = \mathbf{V}\mathbf{x}(t)$, with $E\{\mathbf{z}(t)\mathbf{z}^T(t)\} = \mathbf{I}$. The whitening matrix $\mathbf{V}$ is given by

$$\mathbf{V} = \Lambda^{-\frac{1}{2}}\mathbf{U}^T \qquad (A.7)$$

where $\Lambda = diag[\lambda(1),\ldots,\lambda(m)]$ is a diagonal matrix with the eigenvalues of the data covariance matrix $E\{\mathbf{x}(t)\mathbf{x}^T(t)\}$, and $\mathbf{U}$ is a matrix with the corresponding eigenvectors as its columns. The transformed vectors $\mathbf{z}(t)$ are called white because all directions have equal unit variance. The term $\mathbf{z}(t)$ in equation (1.1) can be written as:

$$\mathbf{z}(t) = \mathbf{V}\mathbf{A}\mathbf{s}(t) \qquad (A.8)$$

and to show orthogonal matrix $\mathbf{W} = \mathbf{V}\mathbf{A}$, Therefore, the solution is:

$$\hat{\mathbf{s}}(t) = \mathbf{W}^T\mathbf{z}(t) \qquad (A.9)$$

There are different ways to measure non-Gaussianity. One of them is the absolute value of the fourth-order cumulant or kurtosis. The kurtosis of a random signal $y$ is given as:

$$Kurt(y) = E\{y^4\} - 3(E\{y^2\})^2 \qquad (A.10)$$

where it is assumed that the random signal has zero mean. The absolute value of the kurtosis is zero for Gaussian variable and greater than zero for non-Gaussian signals. For a Gaussian signal $y$, the fourth moment equals $3(E\{y^2\})^2$. The kurtosis is negative for source signals whose amplitude has sub-Gaussian probability densities ( i.e. more uniformly distributed), and positive for super-Gaussian (sharper than Gaussian). Maximizing the norm of the kurtosis leads to the identification of non-Gaussian sources. "Negentropy" is used as a measure of distance to normality. Consider a signal with a certain distribution, if the signal is Gaussian, the signal is said to have a normal distribution. A measure of non-Gaussianity and an important concept in finding out how independent component is separated can be referred to as negentropy [86] [87]. Negentropy is defined as:

$$J(y) = H(y_{Gaussian}) - H(y) \qquad (A.11)$$

where $H(y)$ represents the differential entropy of the random variable $y$, and $H(y_{Gaussian})$ is the entropy of a Gaussian random variable with the same covariance as $y$.
It is known from information theory that a Gaussian variable has the largest entropy among all random variables of unit variance. This means that entropy could be used as a measure of non-Gaussianity. In fact, this shows that Gaussian distribution is the "most random" or the least structured of all distributions. Entropy is small for distributions that are clearly concentrated on certain values, i.e. when the variable is clearly clustered, or has a pdf that is very "spiky". This property can be generalized to arbitrary variances, and what is more important, to multidimensional

spaces: The Gaussian distribution has maximum entropy among all distributions with a given covariance matrix. Therefore the maximality property shows that entropy could be used to define a measure of non-Gaussianity. Hence, in practice, by fixing the variance of the random variables to unity, the differential entropy can be used as a measure of non-Gaussianity. "Negnetropy" defines the concept of negative entropy, in information theory.

The disadvantage of kurtosis based approach is that kurtosis is sensitive to outliers. Although negentropy is, in some sense, the optimal estimator of non-Gaussianity, it is computationally expensive, since it requires the estimation of the probability density function. Therefore, only approximations of negentropy are used in practice.

# Appendix B

# Least squares estimation

In statistics, the Gauss-Markov theorem [77] states that, in a linear model in which the errors have expectation zero and are uncorrelated and have equal variances, the best linear unbiased estimators of the coefficients are the least-squares estimators. In the basic linear least-squares method equation (B.1) is the assumed model:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \qquad (B.1)$$

where $\mathbf{s}$ is $n$-dimensional unknown source vector, $\mathbf{x}$ is $m$-dimensional measurement vector. $\mathbf{A}$ is $m \times n$ real, typically a full column rank mixing matrix, and $t$ is time index. If the number of sources is more than the number of measurements (i.e. $m < n$) the system is called underdetermined, and for the case of $n > m$ the system is known as overdetermined, for determined case the number of sources are equal to number of measurements (i.e. $m = n$). For a determined system of

equations there may be a unique solution for $\mathbf{s}$, and infinite number of solutions or no exact solution in the case of underdetermined system.
Now in a system of linear equations consider

$$\mathbf{As}(t) = \mathbf{x}(t) - \mathbf{e}(t) = \mathbf{x}_{real}(t) \qquad (B.2)$$

where $\mathbf{e} = \mathbb{R}^m$ is the vector of unknown measurement errors, $\mathbf{x}_{real} = \mathbb{R}^m$ is the vector of real but unknown values. In practice it is required to find a solution for the source $\mathbf{s}$ such that the approximation to be minimal norm and as close as possible to the its real value. This can be formulated as an optimization system in order to find $\hat{\mathbf{s}}$ (where $\hat{\cdot}$ is an estimation of source $\mathbf{s}$) that minimizes the following cost function:

$$\mathbf{J}_p(\mathbf{s}) = \|\mathbf{x} - \mathbf{As}\|_p = \|\mathbf{e}(s)\|_p \qquad (B.3)$$

for $p \geq 1$, where the error vector $\mathbf{e}$ has the component $e(s) = [e_1(s), \ldots, e_m(s)]^T$ for a given vector $\hat{\mathbf{s}}$.
$\|\mathbf{e}\|_p$ is the $p - norm$ of the vector $\mathbf{e}$. Therefore $\hat{\mathbf{s}}$ minimizes the $p - norm$ of the error vector $\mathbf{e}$ with respect to vector $\mathbf{s}$.

$$\|\mathbf{x} - \mathbf{A}\hat{\mathbf{s}}\|_p \leq \|\mathbf{x} - \mathbf{As}\|_p \quad \forall \mathbf{s} \in \mathbb{R}^n. \qquad (B.4)$$

For the above optimization, there are three special cases to note:

- (a) $p$=1, norm used for Laplacian or sparse error distribution problem.
- (b) $p$=2, is called the $2-norm$, Euclidean, or linear least-squares (LS) problem, used for normal distribution.
- (c) $p$=$\infty$, is referred to as the Chebychev norm or minimax ( minimizing the maximum possible error) problem.

It should be noted that, in the special case of zero noise and square matrix $\mathbf{A}$ (i.e. invertible $\mathbf{A}$) the result of calculating error $\mathbf{e}$ is similar for all the three cases

mentioned above (i.e. $p$=1,2,$\infty$) Frobenous norm is used because it is an upper bound of any other norm, otherwise infinity norm ($p$=$\infty$) is the most robust measurement. The linear least-square problem stated in equation (B.3) can be formulated as cost function described in equation (B.5).

$$\mathbf{J(s)} = \frac{1}{2}\|\mathbf{x} - \mathbf{As}\|_2^2 = \frac{1}{2}(\mathbf{x} - \mathbf{As})^T(\mathbf{x} - \mathbf{As}) = \frac{1}{2}\|\mathbf{e}^T\mathbf{e}\| = \frac{1}{2}\sum_{i=1}^{m} e_i^2, \qquad (B.5)$$

where

$$e_i(\mathbf{s}) = x_i - \mathbf{A}_i^T\mathbf{s} = x_i - \sum_{j=1}^{n} A_{ij}s_j \qquad (B.6)$$

The cost function converges to the global minimum when the gradient is zero:

$$\nabla\mathbf{J(s)} = \mathbf{A}^T(\mathbf{x} - \mathbf{As}) = \mathbf{0} \qquad (B.7)$$

Consequently $\hat{\mathbf{s}}$ can have solution for the above mentioned three categories:

- (i) $\mathbf{A} \in \mathbb{R}^{n\times n}$ determined case:

$$\hat{\mathbf{s}} = \mathbf{A}^{-1}\mathbf{x} \quad for\ \mathbf{J}(\hat{\mathbf{s}}) = 0.$$

- (ii) $\mathbf{A} \in \mathbb{R}^{m\times n}$ Over-determined case (i.e. $n < m$)

$$\hat{\mathbf{s}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{x} = \mathbf{A}^+\mathbf{x}$$

  when

$$\mathbf{J}(\hat{\mathbf{s}}) = \frac{1}{2}\mathbf{x}^T(\mathbf{I} - \mathbf{AA})^+\mathbf{x} \geq \mathbf{0}$$

  where $\mathbf{A}^+$ is the pseudo-inverse of $\mathbf{A}$,

- (iii) $\mathbf{A} \in \mathbb{R}^{m\times n}$ Under-determined case (i.e. $m < n$): the solution is not unique, but the LS problem can give the minimum 2-norm $\|\mathbf{s}\|_2^2$ unique solution:

$$\hat{\mathbf{s}}(t) = \mathbf{A}^T(\mathbf{AA}^T)^{-1}\mathbf{x}(t) = \mathbf{A}^+\mathbf{x}(t)$$

with

$$\mathbf{J}\hat{\mathbf{s}}(t) = \mathbf{0}$$

therefore the result value of the norm leads to :

$$\|\hat{s}\|_2^2 = \mathbf{x}^T(t)(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{x}(t).$$

# Bibliography


[1] M A Latif, S Sanei, J Chambers, and L Shoker: 'Localization of abnormal EEG sources using blind source separation partially constrained by the locations of known sources',*IEEE Signal Processing Letter, vol. 13, no. 3, 2006*.

[2] 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',*The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology, 2006*.

[3] 'http://web-us.com/brain/aboutthebrain.htm'.

[4] 'http://www.sfu.ca/Ndkimura/articles/britan.htm'.

[5] E Niedermeyer, and F Lopes De Silva: 'Electroencephalography basic principles clinical applications and related fields', *Forth Edition, 1999*.

[6] P L Nunez: 'Electrical fields of the Brain: The Neurophysics of EEG',*Oxford Univ. Press, New York, 1981*.

[7] D Haines: 'Fundamental Neuroscience',*Churchill Livingstone, 1997*.

[8] A L Hodgkin, and A F Huxley: 'A quantiative description of membrane current and its application to conduction and excitation in nerve',*J Physiol. (Lond.) 117: 500-44 (1952)*.

[9] C S Herrmann, and A Mecklinger: 'Magnetoencephalographic response to illusory figures: early evoked gamma is affected by processing of stimulus feathres.', *Int. J. Psychophysiol. 1:38(3), pp. 265-281, 2000.*

[10] H Berger: 'Uber das Elektroenkephalogramm des Menschen',*Psychiatrie und Nervenkrankheiten, vol. 87, pp. 527-570, 1929*

[11] E D Adrian: 'Olfactory reactions in the hedgehog',*J Physiol* (Lond) *vol. 100, pp. 459-473.*

[12] P L William, and R Warwick: 'Gray's Anatomy,37th ed.',*Churchill Livingstone, Edinburgh, pp. 1598, 1989.*

[13] W Penfield, and T Rasmussen: 'The Cerebral Cortex of Man: A Clinical Study of Localization of Function',*Macmillan, New York, pp. 248, 1950.*

[14] J Malmivuo, and R Plonsey: 'Bioelectromagnetism, principle and applications of bioelectric and biomagnetic fields', *Oxford Univ. Press, New York, 1995.*

[15] C W Hesse, and C J James: 'The fastICA algorithm with spatial constraints', *IEEE Signal processing letters,vol. 12, no. 11, pp. 792-794, 2005.*

[16] S Amari, and A Cichocki: 'Adaptive blind signal processing - neural network approaches', *Proceedings-F vol. 140, pp. 363–3370, 1993.*

[17] S Baillet, R M Leahy, M Singh, D W Shattuck, and J C Mosher: 'Supplementary motor area activation preceding voluntary finger movements as Evidenced by Magnetoencephalography and fMRI ', *Int. J. of Bioelectromag. vol. 3, no. 1, 2001.*

[18] K Jerbi, S Baillet, J C Mosher, G Nolte, L Garnero, and R M Leahy: 'Localization of realistic cortical activit in MEG using current multiploes', *NeuroImage 22, pp. 779-793, 2004*.

[19] C Phillips, Michael D Rugg, and Karl J Friston: 'Anatomically informed basis functions for EEG source localization: Combining functional and Anatomical constraints', *NeuroImage 16, pp. 678-695, 2002*.

[20] T P Jung, S Makeig, M J McKeown, A J Bell, T W Lee, and T J Sejnowski: 'Imaging brain dynamics using independent component analysis',*Proc. of IEEE Signal Processing, vol. 89, no. 7, 2001*.

[21] G Darmois: 'Analyse Général des Liaisons Stochastiques',*Rev. Inst. Internat. Stat., vol. 21, pp. 2-8, 1953*.

[22] J Karhunen: 'Neural approaches to independent component analysis and source separation',*proc. Of the Europ. Sympo. on Art. Neural Networks, pp. 249-266, 1996*.

[23] J Malmivuo and R Plonsey: 'Bioelectromagnetism-principle and application of bioelectric and biomagnetic fields,' *Oxford Univ. Press, New York, 1995*.

[24] S Haykin: 'Unsupervised adaptive filters volume I blind source separation', *John Wiley and Sons,vol. I, 2000*.

[25] A Hyvarinen, and E Oja: 'Independent component analysis: A tutorial', *Helsinki Univ. of Tech., Lab. of Com. and Inf. Science Finland, aapo.hyvarinen@hut.fi, 2000*.

[26] A Hyvarinen, and E Oja: 'Independent component analysis',*John Wiley and Sons, 2001*.

[27] T P Jung, S Makeig, T W Lee, M J McKeown, G Brown, A J Bell, and T J Sejnowski: 'Independnet Component Analysis of Biomedical Signals', *2nd Int'l Workshop on ICA and Signal Separation, pp. 633-44, 2000.*

[28] Z J Koles: 'The quantitive extraction and topographic mapping of abnormal components in the clinical EEG', *Electroenceph. clin. Neurophysiol 79, pp. 440-7, 1991.*

[29] J F Cardoso, and A Souloumiac: 'Blind beamforming for non Gaussian signals', *IEEE-Proceedings-F, vol. 140, no. 6, pp. 362-370, 1993.*

[30] Z J Koles: ' Trends in EEG source localization',*Electroenceph. clin. Neurophysiol., vol. 106, pp. 127-137, 1998.*

[31] J C Maxwell: 'A Dynamical theory of the Electromagnetic File',*Philosophical transactions of the royal society of London 115, pp. 459-512, 1865.*

[32] S J Williamson, G L Romani, L Kaufman, and I Modena: 'Biomagnetism: An Interdisciplinary Appraoch',*NATO ASI seris, Series A: Life Sciences, vol. 66, pp. 706, Plenum Press, New York, 1983.*

[33] R M Gulrajani: 'Electroencephalography and Biomagnetism', *Bioelectricity and Biomagnetism John Wiley and Sons, 1998.*

[34] A Ioannides, J Bolton, and C Clarke: 'Continous probabilitistic solutions to the biomagnetic inverse porblem',*Inverse problems no. 6, pp. 523-542, 1990.*

[35] A P Gibson, J Riley, M Schweiger, J C Hebden, S R Arridge, and D T Delpy: 'A method for generating patient-specific finite element meshes for head modeling',*Institute of phys. Med. Bio. 48, pp. 481-495, 2003.*

[36] C Roy, and C Sherrington: 'C Roy and C Sherrington on the regulation of the blood-supply of the brain',*J. Physiol. no. 11, pp. 85-108, 1890.*

[37] P Jezzard, S Smith, and P M Mathews: 'Brain energy metabilism and the physiological basis of the haemodynamic response, in:Functional MRI: an introduction to methods',*Oxford University press, Oxford, pp. 37-66, 2001.*

[38] B Van Everbroeck, and P Pals, et al: 'Retrospectvie sudy of Creutzfeldt-Jakob disease in Belgium: neuropathological findings',*Acta Neuropathol (Berl) 99(4):358-64, 2000*

[39] *http://www.ninds.nih.gov/disorders/epilepsy/detail epilepsy.htm.*

[40] R Penrose: 'A Generalized Inverse for Matrices', *Proc. Cambridge Phil. Soc. no. 51 pp. 406-413, 1955.*

[41] Epilepsy foundation: *www.epilepsyfoundation.org.*

[42] B He, and J Lian: ' Electrophysiological Neuroimaging', *In B. He, Ed.: Neural Engineering, Kluwer Academic/Plenum Publishers, 2005*

[43] T Knosche, E Berends, H Jagers, and M Peters: ' Determining the number of independent sources of the EEG: A simulation study on information criteria', *Brain Topography, vol. 11, pp. 111-124, 1998.*

[44] X Bai, B He: ' Estimation of number of independent brain electric sources from the scalp EEGs', *The supercomputing institute and the biomedical engineering institute of the university of Minnesota. TBME-00378-2005.R1, 2006.*

[45] A Cichocki, J Karhunen, W Kasprzak, and R Vigario: ' Neural network for blind separation with unknown number of sources',*Neurocomputing, vol. 24, pp. 55-93, 1999.*

[46] D Weinsteint, L Zhukov, and G Potts: 'Localization of multiple deep epileptic sources in a realistic head model via independent component analysis',textit2nd International Symposium on Noninvasive Functional Source Imaging within the Human Brain and Heart (Zagreb, Croatia, 1999).

[47] D B Smith, R D Sidman, H Flanigin, J Henke, and D Labiner: 'A reliable method for localizing deep intracranial sources of the EEG', *Neurology, 35, pp. 1702-1707, 1985.*

[48] T B J Wiederin, K H Chiappa, T Krings, B N Cuffin, D B Hoch, A J Cole, and G R Cosgrove: 'The Utility of dipole source analysis of seizure onsets in the localization of epileptogenic zones as assessed by postsurgical outcome',*J. of contem. Neuro.: From the Clin. Neurophysiol. Lab. and the Epilepsy Service of the Department of Neurology, Medical Center, Boston, MA, USA. vol. 1999 Article 1A, 1999.*

[49] I F Gorodnitsky,J S George, and B D Rao: 'Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm',*Electroence. clin. Neurophysiol. 95, pp 231-251, 1995.*

[50] 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 trans. on Biomed. Eng., vol. 52, no. 10 pp. 1681-1691, 2005.*

[51] J G Taylor, A. Ioannides, and Hans-Wilhem Muller-Gartner: 'Mathematical analysis of lead field expansions', *IEEE transaction on medical imaging vol. 18, no. 2, 1999.*

[52] R D Pascual-Marqui, C M Michel, and D Lehnam: 'Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain',*Int. J. Psychophysiol. pp. 49-65 vol. 18, 1994*.

[53] R D Pascual-Marqui: 'Review of methods for solving the EEG inverse problem',*Int. J. Bioelectromagn., vol. 1, pp. 75-86, 1999*.

[54] R D Pascual-Marqui: 'Standardized Low resolution brain electromagnetic tomography (sLORETA): Technical details',*Meth. Findings Exp. Clin. Pharmacol., vol. 24D, pp. 5-12, 2002*.

[55] P H Schimph, H Liu, X Gao, and F Yang: 'Efficinet electromagnetic source imaging with adaptive standardize LORETA/FOCUSS', *BioMed. Eng. vol. 52, no. 5, pp. 547-552, 2005*

[56] I F Gorodnitsky, J S George, and B D Rao: 'Neuromagnetic source imaging with FOCUSS: A recursive weighted minimum norm algorithm', *J. Clinc. Neurophysiol., vol. 16, no. 3, pp. 265-295, 1999.*

[57] R D Pascual-Marqui: 'Standardized low resolution brain electromagnetic tomography (sLORETA): Technical details', *Method Findings Exp. Clin. Pharmacol., vol. 24D, pp. 5-12 , 2002*.

[58] H Liu, 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 Trans. on Biomed. Eng. vol. 52, no. 10, 2005*.

[59] S Supek, and C J Aine: 'Simulation studies of multiple dipole neuromagnetic source localization: model order and limits of source resolution',*IEEE Transaction Biomed Eng. 40 pp. 354-361, 1993*.

[60] A Achim and F Richer and Saint-Hilaire: ' Methods for separating temporally overlapping sources of neuroelectric data',*Brain topography 1 pp. 22-28, 1998.*

[61] M Scherg and D von Cramon: ' Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model', *Electroenceph. Clin. Neurophysiol. 62 pp. 290-299 ,1985.*

[62] J C Mosher, P S Lewis and R M Leahy: 'Multiple dipole modeling and localization from spatio-temporal MEG data', *IEEE Trans. Biomedical Engineering 1992, 39, pp. 541-557, 1992.*

[63] J C Mosher, and R M Leahy: 'Recursive MUSIC: a framework for EEG and MEG source localization', *IEEE Trans. Biomed. Eng., vol. 45, no. 11, pp. 1342-1354, 1998.*

[64] H Liu, and P H Schimpf: 'Efficient localization of synchronous EEG source activites using a modified RAP-MUSIC algorithm', *IEEE trans. Biomed. Eng. vol. 53, no. 4, 2006.*

[65] M Fuchs , M Wagner, Kohler T, and H Wischmann 'Linear and non-linear current density erconstructions', *J. Clin. Neurophysiol., vol. 16, no. 3, pp. 267-295, 1999.*

[66] D Contreras, and R Llinas: 'Voltage-sensitive dye imaging of neocortical spatiotemporal dynamics to afferent activation frequency', *J. Neurosci., 21(23): pp. 9403-13, 2001.*

[67] S Makeig: 'Response: event-related brain dynamics-unifying brain electrophsiology', *Trends Neurosci., 2002, 25(8): pp. 390, 2002.*

[68] W J Freeman: 'Neurodynamics: an exploration in mesoscopic brain dynamics', *Perspectives in neural computing 2000, London: New York: Springer. ix. 397 2000.*

[69] P Fries: 'A mechanism for cognitive dynamics: neuronal communication through neuronal coherence', *Trends in Cognitive Sciences, 9(10): pp. 474-480, 2005.*

[70] D Weinstein, L Zhukov, and G Potts: ' Localization of multiple deep epileptic sources in a realistic head model via independent component analysis', *CSCI Department of Computer Science University of Utah Salt Lake City UT 84112 USA UUCS-2000-004, 2000.*

[71] D Weinstein, l Zhukov, and C Johnson: 'Lead-field bases for EEG source imaging', *Annals of Biomedical Eng, vol. 28 no. 9 pp. 1059-1164, 2000.*

[72] D T Pham, P Garrat, C Jutten: 'Separation of a mixture of independent sources through a maximum likelihood approach', *Proc EUSIPCO 1992, pp. 771-774, 1992.*

[73] A Papoulis: 'Probability random variables and stochastic processes', *McGraw-hill 3rd edition, 1991.*

[74] Jean-François Cardoso: ' Perturbation of joint diagonalizers', *Ref# 94D027, 1994.*

[75] J Herault, and C Jutten: 'Space or time adaptive signal processing by neural network models', *presented at the Neural Networks for Computing AIP Conf EUSIPCO 1992.*

[76] P Comon: 'Independent component analysis, a new concept?', *Signal Processing, 36 (3), pp. 287-314, 1994.*

[77] *'http://en.wikipedia.org/wiki/Gauss-Markov theorem'*.

[78] R O Duda, P E Hart, and D G Stork: 'Pattern Classification - second edition', *John Wiley and Sons, Inc., 2001* .

[79] A J Bell, and T J Sejnowski: 'An information maximization approach to blind separation and blind deconvolution', *Neural Computing, 1995, 7, pp. 1129-1159, 1995* .

[80] M D Rugg: 'Functional neuroimaging in cognitive neuro-science.', *The Neurocognition of Language(C M Brown and P Hagoort, Eds.), pp 15-36. Oxford Univ. Press, Oxford 1999.*

[81] K Jerbi, S Baillet, J C Mosher, G Nolte, L Garnero, and R M Leahy: 'Localization of realisitic cortical activity in MEG using current multipoles', *NeuroImage 22 pp. 779-793, 2004.*

[82] A N Tikhonov, and V Y Arsenin: 'Solution of ill-posed problems, scripta series in mathematics', *Wiley, New York, 1977.*

[83] K Matsuura, and Y Okabe: 'Selective minimum-norm solution of the biomagnetic invrese problem', *IEEE trans. Biomedic. Eng. vol. 42, issue 6, pp. 608-615, 1995.*

[84] J F Cardoso, and B H Laheld: 'Equivariant adaptive source separation', *IEEE Trans. Signal Processing, 44, pp. 3017–3030, 1996.*

[85] T W Lee and M Girolami and and T J Sejnowski: 'Independnet component analysis using an extended infomax algorithm for mix sub Gaussian and super Gaussian sources', *Neural Computing, 1999, vol. 11 pp. 417-441.*

[86] J M Mendel: 'Tutorial on higher order statistics (spectra) in signal processing and system theory', *theoretical results and some applications Proceedings of the IEEE, 1991, 79(3) pp. 278-305, 1991.*

[87] S Amari: 'Natural gradient works efficiently in learning', *Neural Computing :1998, vol. 10, pp. 251-276, 1998.*

[88] J F Cardoso and A Souloumiac: 'Blind beamforming for non-Gaussian signals', *IEE Proceedings-F :December 1993, 140*(6), pp. 362-370, 1993.

[89] L Parra, and C Spence: 'Convolutive bind source separation based on multiple decorrealtion ', *IEEE Neural Networks and Signal Processing Workshop, Cambridge, UK, September 1998, presented at "Machines that Learn" Workshop, Snowbird.*

[90] C E Shannon: 'A mathematical theory of commuincation', *The bell system technical Journal, vol. 27, pp. 379-423, 623-659, 1948.*

[91] J J Ermer, J C Mosher, S Baillet and R M Leahy: 'Rapidly recomputable EEG forward modles for realistic head shapes', *Physics in Medicine and Biology : 2001, 46, pp. 1256-1281, 2001.*

[92] M X Huang, J C Mosher, and R M Leahy: 'A sensor weighted overlapping sphere head model and exhaustive head model comaprisonfor MEG', *Physics in Medicine and Biology: 2001, 44, pp. 423-440, 2001.*

[93] J Ollikainen, M Vauhkonen, P A Karjalainen, P J Ronkanen and J P Kaipio: 'Effect of skull inhomogeneities on EEG localization accuracy', *Prc. 19th. Int. Conf. IEEE EMBS 1997, Chicago, USA, pp. 2120-2123, 1997.*

[94] P Jiruška, J Prokš, O Drbal, P Sovka, P Marusič, P Nareš: 'Comparison of different mehtod of time shift measurement in EEG', *Physiol. Res. 54, pp. 495-465, 2005*.

[95] S Sanei, and J Chambers: 'EEG Neuro-Image reconstruction based on conditional chaos', *IEE signal processing professioanl network, Medical applications of signal processing, Seminar reference number 02/110, pp. 18/1-18/5, 2002*.

[96] A Belouchrani, K Abed-Meraim, J F Cardoso, and E Moulines: 'A blind source separation technique using second order statistics', *IEEE Trans. Signal Processing, vol. 45 pp. 434-444, 1997* .

[97] J A Malmivuo, and V E Suihko: 'Effect of skull resistivity on the spatial resolutions of EEG and MEG', *IEEE Transactions on Biomedical Engineering vol. 51 no. 7, pp. 1276-1280, 2004*.

[98] A Hyvarinen, J Karhunen, and E Oja: 'Independnet component analysis', *John Wiley and son Inc. 2001*.

[99] A Papoulis and S U Pillai: 'Probability, random variables and stockastic process', *McGraw-Hill, 2001*.

[100] T M Cover, and J A Thomas: 'Elements of informaion theory', *Wiley Interscience, 1991*.

[101] A Hyvarinen and J Karhunen and E Oja: 'Independnet component analysis: Algorithm and applications,', *Neural networks, vol 13 pp. 411-430, 2000*.

[102] E Oja, and A Hyvarinen, and J Karhunen: 'Independnet component analysis', *Thomson Learning Inc., 2001*.

[103] A J Bell and T J Sejnowski: 'An information-maximization approach to blind sepration and blind deconvolution', *Neural Computation, vol. 7, pp. 1004-1034, 1995.*

[104] J C Mosher, and R M Leahy: 'Source localization using recursively applied and projected (RAP) MUSIC,',*IEEE Trans. Signal Processing, pp. 332-340, 1999.*

[105] R M Leahy et al: 'A study of dipole localization accuracy for MEG and EEG using a human skull phantom', *Electroencephalography and Clinical Neurophysiology, vol. 107, pp. 159-173, 1998.*

[106] L Shoker, S Sanei, and J Chambers: 'Artifact removal from electroencephalograms using a hybrid BSS-SVM algorithm, ', *IEEE Signal Processing Letters*, vol. 13(3), pp. 117-120, 2005.

[107] L Shoker, S Sanei, W Wang, and J Chambers: 'Removal of eye blinking artifact from EEG incorporating a new constrained BSS algorithm,', *IEE Journal of Medical and Biological Engineering and Computing, vol. 43, pp. 290-295, 2005.*

[108] J F Cardoso: 'Infomax and maximum likelihood for blind source separation', *IEEE Signal Processing Letter, 4, 1997, pp. 109 - 111, 1997.*

[109] A Cichocki, and S Amari: 'Adaptive blind signal and image processing', *Book publisher Wiley, England 2002.*

[110] I D Coope: 'Reliable computation of the points of intersection of n spheres in $\mathbb{R}^n$', *ANZIAM J., 42(E), pp C461-C477, 2000.*

[111] J Cardoso, and A Souloumiac: 'An efficient technique for blind separation of complex sources',*In Proc. IEEE SP Workshop on Higher-Order Stat., Lake Tahoe, USA, pp 275-279, 1993.*

[112] M Joho, and K Rahbar: 'Joint diagonalization of correlation matrices by using newton methods with application to blind signal separation', *SAM 2002, 4-6, Rosslyn, VA, USA, pp 403-407, 2002.*

[113] R O Schmidt: 'Multiple emitter location and signal parameter estimation', *Proc. (RADC) spectrum estimation workshop, Griffiss (AFB), N.Y., pp. 243-258, 1979.*

[114] D Gutiérrez, A Nehorai: 'Estimating brain conductivities and dipole source signals with EEG arrays', *IEEE Trans. Biomed. Eng., vol. 51, no. 12, 2004.*

[115] C H Muravchik, A Nehorai: 'EEG/MEG error bounds for a static dipole source with a realistic head model', *IEEE Trans. Signal Proc., vol. 49, no. 3, 2001.*

[116] B. D. Van Veen, W. van Drongelen, M. Yuchtman, A. Suzuki: 'Localization of brain electrical activity via linearly constrained minimum variance spatial filtering', *IEEE Trans. Biomed. Eng., vol. 44, no. 9, pp. 867-880, 1997.*

[117] M Scherg, and D Von Cramon: 'Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model', *Electroenceph. Clin. Neurophysiol. 62, pp. 290-299, 1985.*

[118] J C Mosher, and R M Leahy: 'Source localization using recursivley applied and projected (RAP) MUSIC', *IEEE Trans. signal proc., pp. 332-340, 1999.*

[119] R M Leahy: 'A study of dipole localization accuracy for MEG and EEG using a human skull phantom', *Electroenceph. Clin. Neurophysiol., vol. 107 pp. 159-173, 1998.*

[120] R D Pascual-Marqui, C M Michel, and D Lehmann: 'Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain', *Inter. Jour. Psychophysiol., vol. 18, pp. 49-65, 1995.*

[121] J Muscat, R Grech, K P Camilleri, S G Fabri, C J James: 'An improvement in EEG forward model computational efficiency', *2nd Int. Conf. Adv. Biomed. Signal info. proc., pp. 99-103, 2004*.

[122] W Wang, J A Chambers, and S Sanei: 'A joint diagonalization method for convolutive blind separation of nonstatioanry sources in the frequency domain', *Proc. of ICA2003 Nara Japan pp. 939-944, 2003*.

[123] A J van der Veen: 'Joint diagonalization via subspace fitting techniques', *In proc. of IEEE ICASSP 2001, Salk Lake, USA 2001*.

[124] M Joho, H Mathis: 'Joint diagonalization of correlation matrices by using gradient method with application to BSS', *IEEE, SAM2002, August 4-6 Rosslyn, VA., USA, pp 273-277, 2002*.

[125] A C Tang, M T Sutherland, and Y Wang: 'Contrasting single-trial ERPs between experimental manipulations: Improving differentiability by blind source separation', *NeuroImage 29(2006), pp. 335-346, 2006*.

[126] A Ziehe, P Laskov, G Nolte, and K Müller: 'A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation', *Journal of Machine learning research 5 pp. 777-800, 2004*.

[127] California University: *http://www.csun.edu/ vcpsy00i/dissfa01/EEG lesson.html*.

[128] R D Pascaul-Marqui, M Esslen, K Kochi, and D Lehmann: 'Functional imaging with low resolution brain eletromagnetic tomography (LORETA): a review', *Methods and Findgings in Exper. and Clin. Pharmacology 2002, 24C: pp. 91-95, 2002*.

[129] J Yao, and J P A Dewald: 'Evaluation of different cortical source localization methods using simulated and experimental EEG data', *NeuroImage 25, pp. 369-382, 2005*

[130] M Adjouadi, N Mirkovic, M Cabrerizo, and M Ayala: 'An inverse solution to 3-D source localization of epileptic foci using integrated multimodal neuroimaging approach', *Inverse Problem, Design and Optimization Symposium Rio de Janeiro, Brazil, 2004*.

[131] S Sanei, L Spyrou, W Wang, J Chambers, P Carlos, and P Alberto: 'Localization of P300 sources in schizophrenia patients using constrained BSS, *Independent Component Analysis and Blind Signal Separation Conference, Garanda, Spain, vol. 3195, pp. 177-184, 2004*.

[132] B W Jervis, M R Saatchi, E M Allen, N R Hudson, S Oke, and M Grimsley: ' Pilot study of computerised differentiation of Huntington's disease, schizophernic, and Parkinson's diseas patients using the Contingent Negative Varitation', *Medical and Biological Engineering and Computing, vol.31, pp. 31-38, 1993*

[133] E Niedermeyer: 'Alpha rhythm as physiological and abnormal phenomena', *Int. J. Psychophysiol. 1997 Jun; 26(1-3):31-49, 1997*

[134] P J Marshall, Y Bar-Haim, N A Fox: 'Development of the EEG from 5 months to 4 years of age', *Clin. Neurophysiol 113, pp. 1199-1208, 2002*.

[135] S Sanei, L Shoker, and M A Latif: 'A new constrained BSS algorithm for separation of EEG signals with eye-blinking artifact ', *3rd IEEE Sensor Array And Multichannel signal processing workshop, SAM2004 Barcelona Spain.*

[136] L Shoker, S Sanei, and M A Latif: 'Removal of eye blinking artifacts from EEG incorporating a new constrained BSS algorithm', *3rd IEEE Sensor Array And Multichannel signal processing workshop, SAM2004 Barcelona Spain 2004*.

[137] J H Margerison, C D Binnie, and I R McCaul: 'Electroencephalographic Signs employed in the Location of Reptured Intracranial Arterial Aneurysms', *Electroencephalogr. and Clin. Neurophsiol. pp. 296-306, 1969*.

[138] A C Tang, M T Sutherland, and C J McKinney: 'Validation of SOBI components from high-density EEG', *NeuroImage 25, pp. 539-553, 2004*.

[139] A C Tang, B A Pearlmutter, N A Malaszneko, D B Phung, and B C Reeb: 'Independnet components of magnetoencephalography: localization', *Neural Computation, vol. 14, pp. 1827-1858, 2002*.

[140] C Philips, M D Rugg, and K J Friston: 'Systematic regularization of linear inverse solutions of the EEG source localization problem', *NeuroImage 17, pp. 287-301, 2002*.

[141] B He, and J Lian, 'Electrophysiological Neuroimaging', *In B. He, Ed.: Neural Engineering, Kluwer Academic/Plenum Publishers, 2005*.

[142] L J Waldorp, H M Huizenga, R P P P Grasman, K B E Bcker, J C de Munck, and P C M MolenaarC : 'Model Selection in Electromagnetic Source Analysis With an Application to VEFs', *IEEE Trans. on Biomed. Eng. vol. 49 no. 10, 2002*.

[143] L J Waldorp, H M Huizenga, A Nehorai, R P P P Grasman, and P C M Molenaar: 'Model Selection in Spatio-Temporal Electromagnetic Source Analysis', *IEEE Trans. on Biomed. Eng. vol. 52 no. 3 2005*.