KU LEUVEN

An evaluation of independent component analyses with an application to resting-state Electrocorticography

Marimuthu Ananthavelu

Thesis submitted for the degree of Master of Science in Artificial Intelligence, option Engineering and Computer Science

Thesis supervisor:
Prof.dr. Ir. Marc Van Hulle

Assessors:
Ir. Mansoureh Fahimi Hnazaee
dr. Ir. Ghumare Eshwar

Mentor:
Ir. Mansoureh Fahimi Hnazaee

Academic year 2018 – 2019

Preface

I would first like to thank my thesis advisor Mansoureh Fahimi Hnazaee of the Faculty of Medicine, Computational Neuroscience Group, for providing with knowledge, resources, guidance, help and most of all the motivation to continue learning and researching on this thesis work.

Besides my advisor, I would like to thank Prof. Marc Van Hulle for his encouragement, suggestions during the intermediate presentation.

I thank all the teammates in the Computational Neuroscience group for accommodating me during the time for completing this work.

Finally, I must express gratitude to my Parents, my Brother, my Sister, my Wife for their support and encouragement throughout my years of study.

Marimuthu Ananthavelu

Contents

Abstract

In this thesis work, a different number of Independent Components Analysis(ICA) algorithms are applied to resting state Electrocorticography (ECoG) data on five different human subjects. At first, the total number of reliable independent components or the so called cortical sources are validated with the help of an already established method in which the randomization of initial parameters in combination with bootstrapping of the samples followed by clustering was used. Once the total number of cortical sources are evaluated, the comparison has been made between FastICA, InfoMax, Jade, OGWE, Pearson, and EFICA algorithms for the quality of the estimated components using the measures of independence which are, the reduction in the mutual information for statistical independence and Spearman correlation coefficient for linear independence between the components. The estimated independent components from each algorithm, are ordered for the importance of non-gaussianity with the help of measures namely, Entropy and Kurtosis. The results are discussed while taking into account the individual functioning of each ICA algorithm and its specific goal of optimization.

List of Figures

List of Tables

List of Abbreviations and Symbols

Abbreviations

MI	Mutual Information
ECoG	Electrocorticography
ICA	Independent Components Analysis
CCA	CurviLinear Component Analysis
SL,CL,AL	Single Linkage, Complete Linkage, Average Linkage clustering
EEG	Electroencephalography
PCA	Principal Components Analysis
CCA	Curvilinear Components Analysis

Symbols

N	Number of samples
$\overline{x}$	Mean of time series, x
$\overline{y}$	Mean of time series, y
I_q	Cluster Quqlity Index
C_m	Set of components which belong to Mth cluster
H_x	Entropy of random variable x.
$H(x, y)$	Joint Entropy between random variables x and y.
$I(x, y)$	Mutual Information between random variables x and y.

Chapter 1

Introduction

1.1 Objective

The goal in this work is to evaluate how many numbers of reliable cortical sources shall be obtained using Independent Components Analysis (ICA) and to evaluate how different types of algorithms perform on the resting state Electrocorticography (ECoG) data recorded from the human brain. The evaluation is done on two levels. In the first level, the goal is to evaluate the total number of reliable independent components from the original number of observations, which is one of the important issues in ICA. In the second level, there exists the number of ICA algorithms and enough care needs to get accounted while choosing the best one for specific application for effectively extracting the independent components with quality for statistical independence. [11]. Considering the above, the goal of this work is;

- *To evaluate the number of reliable independent components* obtained from the Human brain resting state Electrocorticography recordings for five different subjects using FastICA algorithm[11].
- *To evaluate the effectiveness of different ICA algorithms* with the help of a measure namely mutual information for statistical independence and Spearman coefficient for linear independence between the components.
- Once evaluated for the number of independent components and the effectiveness of each ICA algorithms, the estimated independent components *are ordered for its importance* with the help of measures namely Entropy and Kurtosis and compared.
- The results are discussed while taking into account for each ICA algorithms optimization objective.

1.2 Importance

Applications of ICA are enormous and the availability of a large number of algorithms gives more alternatives for different datasets belonging to domains of interest. One issue where the ICA algorithms often fail to address is its ability to order the importance of the estimated components. This requires additional evaluation techniques to arrive at the reliable number of sources of independent components which are at most times less than of those originally observed.

In addition, different ICA algorithms work at the expense of specificity at which times some work better than others for an application. This requires caution in applying techniques such as ICA in which there is the number of optimization objectives the group of algorithms work upon. This needs an evaluation of each specific algorithm to a particular application where the advantages and disadvantages are quantified. This work is focussed on the set of ICA algorithms applied on the resting state Electrocorticography(ECoG) which uses an invasive procedure and is one of the most important techniques used in assessing the neuronal activity of the human brain. While there is the number of similar studies for the non-invasive recordings such as Electroencephalogram (EEG) and Magnetoencephalography (MEG), the need for a similar assessment on an invasive recording such as ECoG, is of importance considering its widespread applications in the field of computational neuroscience.

1.3 Relevance

This work is situated in the field of computational neuroscience where the neuronal behaviour is empirically evaluated with the help of electrical activity recorded using the different levels of invasiveness. The processing of these recordings is done using the statistical techniques which are mainly unsupervised in nature. In this work, the focus has been on Independent Component Analysis applied to rest state invasive Electrocorticography(ECoG).

1.4 Outline

This thesis work is organized as follows;

1. chapter 1 provides an introduction to the objective of the work, importance and relevance in the field of computational neuroscience with a focus on an Independent Component Analysis.

2. chapter 2 provides the background information needed to read this thesis work which includes the concepts of neuronal recordings, impact of invasive recordings, Electrocorticography(ECoG), relevance of ICA to neuronal sources, concepts behind ICA, working principle, measures of independence, relevance to the non-gaussianity, Kurtosis, Entropy, Joint Entropy and their significance in achieving statistical independence. In addition, the literature of the existing work has been reviewed with reference to this work.

3. chapter 3 discusses briefly on the different ICA algorithms which are used, detailed step by step procedure which was carried out in evaluating the reliable number of independent sources and in evaluating the effectiveness of the components obtained from different ICA algorithms.

4. chapter 4 discusses the results of the work stated in the methods and infers the observations with the help of Clustering techniques, Mutual Information, Spearman correlation coefficient, Kurtosis, Entropy, Joint Entropy measures using plots, distributions and tables.

5. chapter 5 concludes the thesis work with the main results achieved on the reliable estimates and on the chosen algorithms for ECoG recordings on the Human brain.

1.5 Motivation

Proposed Solution differs in nature from the previous works for the following reasons;

- *Comparing* the ICA algorithms for ECoG recordings whereas the existing works were *primarily* focussed on the non-invasive types of recordings.
- Considering the number of measures for statistical independence instead of relying upon the single one.

1.6 Conclusion

Aligning with an introduction presented on the scope, importance and relevance, the background and the review on the related literature are presented in the next chapter.

Chapter 2

Background and Literature review

2.1 What is Electrocorticography (ECoG)?

The non-invasive techniques such as EEG and MEG are distorted by skull, meninges, and skin. Invasive human brain techniques have become more useful due to its proximity of recordings with the underlying neuronal activities. *Electrocorticography (ECoG) is a technique which is used to record the neuronal activity invasively in the human brain [9].* ECoG is done with the help of electrodes placed on the cortical surface of the brain which is believed to be closer to the generators of the neuronal sources[9]. This proximity provides higher spatial resolution than scalp electrodes and offers highly valuable information about the underlying sources. In specific, the application to an epilepsy surgery, which enables further understanding of some important challenges such as epileptogenic zones and mapping functional areas effectively [9].
Figures 2.1 shows the channels location for ECoG.

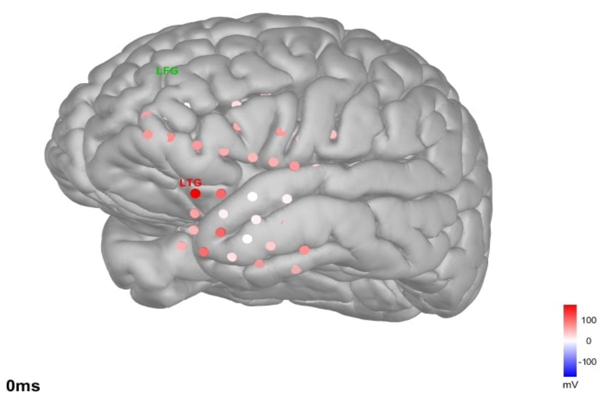


Figure 2.1: ECoG Channels on Cortex

2.2 What is Independent Component Analysis(ICA)?

Independent component analysis (ICA) is one of the most widely used Blind Source Separation (BSS) techniques for revealing hidden factors that underlie sets of random variables, measurements, or signals[25]. These random variables are of observed data points with or without time domain. This method can be used in which the raw measurements are used to estimate the sources which are assumed to be *statistically independent* from one another. It is also assumed that the recorded measurements are the linear combination of the true sources. Unlike PCA, ICA is based upon the higher order statistics which carries more information and works only when the given of random variables which follow the Gaussian distribution does not exceed one.[25]

2.2.1 How it works?

Suppose we have **n** independent signals, $s_i(t)$, whereas $i = 1, 2, 3..., n$ and s, indicates the source signals.The measured signals $x_i(t)$,i =1$, ..., N$ is of those which resulted from mixing.
The objective of an ICA algorithm is to deconstruct the recorded signals into the source signals and demixing matrix.

$$x(t) = As(t) \tag{2.1}$$

whereas;
$x(t)$ - recorded measurements
$A(t)$ - mixing matrix
$s(t)$ - source signals

It can also be written as ;

$$s_t = \frac{x_t}{A} = x_t W \tag{2.2}$$

whereas;
W=unmixing matrix

This deconstruction process can be viewed as below;

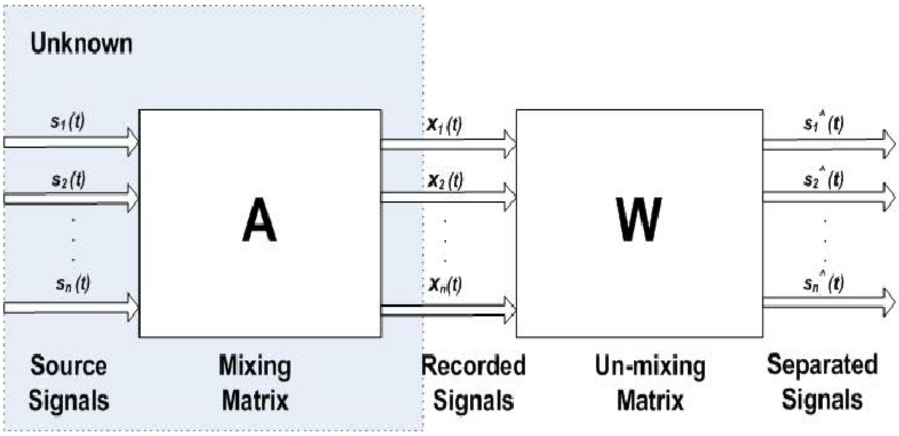


Figure 2.2: Souce Signals, Mixing Matrix, Recorded signals, Un-mixing matrix, Separated signals [25]

An example can be illustrated by considering the following two original sources which are generated using two randomly created *sine* waves namely A and B;

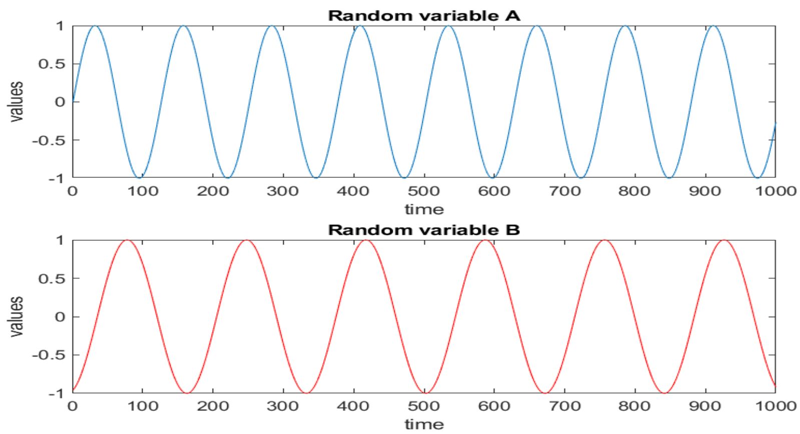


Figure 2.3: Original signal sources

M1 = A - 2.5*B and M2 = 1.65*A+3.7*B, the observed signals shall be as below;

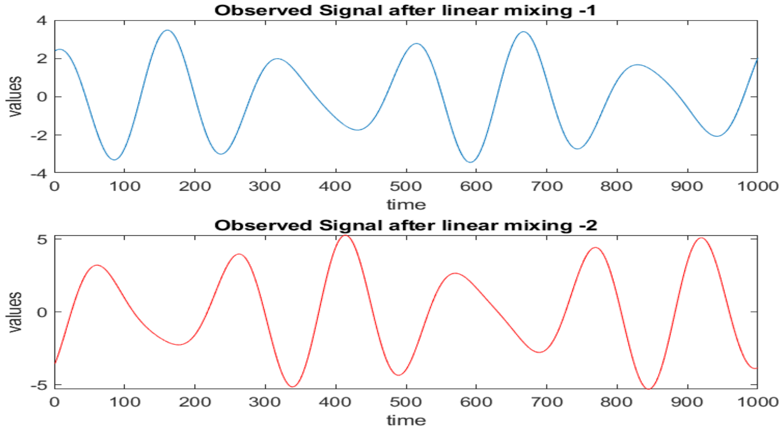


Figure 2.4: Sources mixed with random noise

The estimated recovered signals with the help of FastICA algorithm minimizing the non-gausianity, shall be as below;

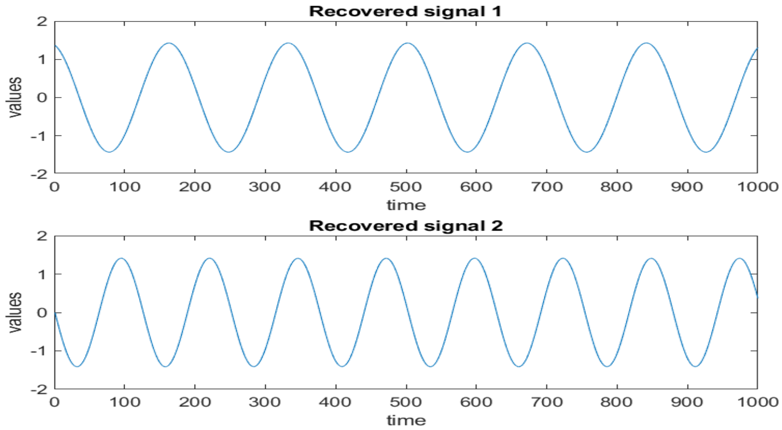


Figure 2.5: Recovered/estimated sources

It can be seen from the resulting independent components that the ICA algorithm is able to recover the true sources approximately. There are a number of ICA algorithms with several of them sharing the common objective function maximizing the non-gaussianity or minimizing the mutual information.

2.2.2 Statistical independence

To estimate the independent components, the statistical independence of the involved random variables needs to get achieved. *Statistical independence implies the relationship between two random variables and indicates one's dependency by the presence of another random variable. When the presence of a particular random variable affects* *another random variable, then both the random variables are said to be dependent [14].* Statistical independence can be described with the help of probability density function *(pdf)* in which each source is characterized by one.

$$p(s) \tag{2.3}$$

whereas;
$p(s)$ - probability density function of source s.
The joint probability density function for two sources, is defined as;

$$p(s_1, s_2) = p_{s1}p_{s2} \tag{2.4}$$

whereas;
$p(s_1, s_2)$ - joint probability density function
$p(s_1)$ - probability density function for Source s_1
$p(s_2)$ - probability density function for Source s_2

For the random variables s_1,s_2 with the following distributions, the joint probability distribution shall be statistically independent as given below i.e given one random variable, it is not possible to estimate the other random variable.

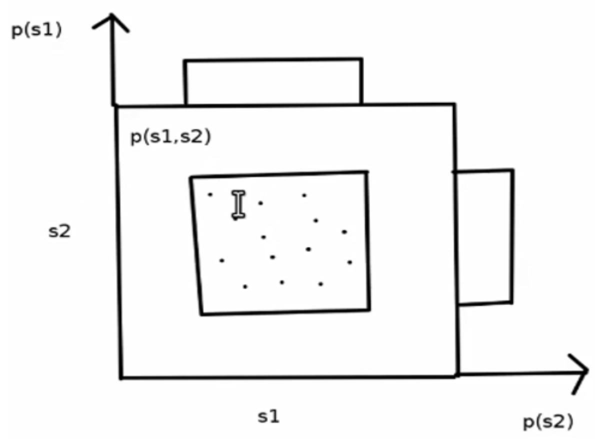


Figure 2.6: Statistically independent components[33]

From the above two distributions, there is no way s_1 can be estimated with other variable s_2 and vice versa.

However, in the following case of distributions, the random variables are not statistically independent.

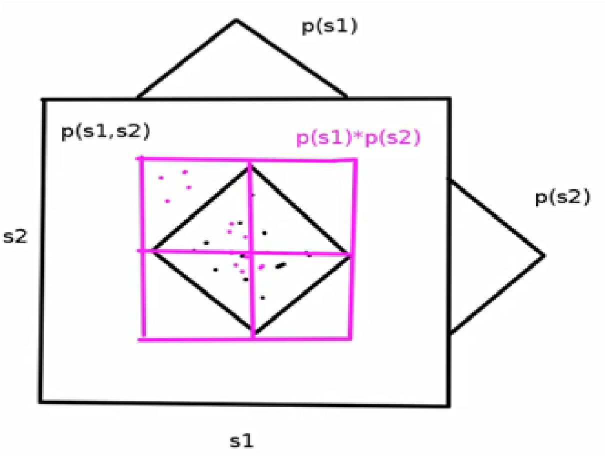


Figure 2.7: Statistically not independent components[33]

In case of the above two distributions, knowing one random variable s_1 gives an information about s_2.

2.2.3 Linear Independence

Linear dependence is a weaker form of Statistical dependence. The dependence between two random variables shall be of linear in nature or non-linear. While the linear dependence can be quantified with the help of correlation coefficients such as Pearson, Spearman, they are not necessarily the measure of statistical dependence. There exist non-linear dependencies which are not quantified with the help of above-mentioned correlation coefficients. Thus, uncorrelatedness is a weaker form of statistical independence which implies that it does not measure the complete statistical independence though it can be used as a weak indication about the type of dependency between random variables.

When it comes to estimating the correlation between two random variables, Spearman Correlation Coefficient is a measure to indicate the strength and direction of the linear relationship and it is calculated between two random variables as below;

$$r = \frac{\sum_{t=1}^{n} (\mathrm{x(t)} - \overline{\mathrm{x}})(\mathrm{y(t)} - \overline{\mathrm{y}})}{\sqrt{\sum_{t=1}^{n} (\mathrm{x(t)} - \overline{\mathrm{x}}^2)(\mathrm{y(t)} - \overline{\mathrm{y}}^2)}} \tag{2.5}$$

whereas;
$x(t)$ and $y(t)$ - random variables x and y which are of time series,
$\overline{x}$ and $\overline{y}$ - Means of random variables x and y

The reason for not using the Pearson correlation coefficient being, the variables are applied as they are in the above equation whereas, in Spearman correlation coefficient calculation, the measured variables are at first transformed to the corresponding 'rank' values and then applied on the above equation. As an example, the frequencies values of 5,8,3,4,15,115 Hz's are transformed to their rank in an ascending order i.e 3,4,1,2,5,6 (it can be seen that the lowest original number gets 1 and the highest gets 6 in the transformed data). By doing so, the impact of the outliers don't exist in case of Spearman correlation[7].

2.2.4 Measure of Statistical independence

The independence between random variables can be measured either in terms of Non-gaussianity or in terms of the Mutual Information(MI) they share in between. By definition, and as per Central limit theorem, the independent variables are either supergaussian or subgaussian[7]. Thus, this statistical independence can be achieved either by;

- by maximizing the non-gaussianity

- by minimizing the Mutual Information

2.2.5 Non-Gaussianity

The quantification of Non-gaussianity measures to what extent the random variable differs from Gaussianity.

Kurtosis

One of the important measures of non-gaussianity is 'Kurtosis'. *This is a fourth order cumulant which is used to signify whether a particular random variable follows Gaussian distribution or not.* When the measure is zero, then the random variable is said to be following the Gaussian distribution. When it is higher or lower, the random variable is said to be following the non-gaussian distribution. In practice, Kurtosis measures the pointedness ('how pointy it is') of a distribution[14].

Kurtosis can be estimated as following;

$$kurt(x) = E(x^4) - 3(Ex_2)^2 \tag{2.6}$$

Visually the Kurtosis measure for Gaussian and non-gaussian distribution can be seen as below;

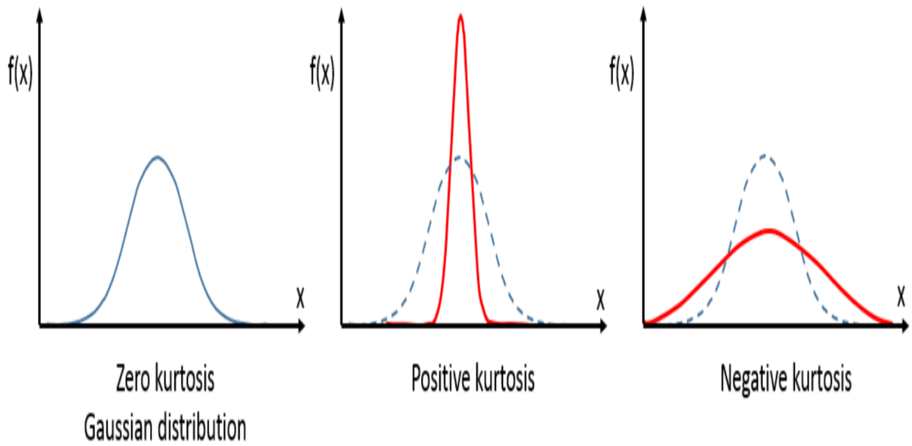


Figure 2.8: Kurtosis for Gaussian, Supergaussian and Sub-gaussian distribution[33]

NegEntropy

Entropy quantifies the expected amount of information held in a random variable and is defined as below;[14]

$$H(x) = -\int p(x)logp(x) \tag{2.7}$$

whereas;
$H(x)$ is Entropy of random variable x,

Negentropy measures the Gaussianity of a particular random variable. Considering that the Gaussian random variable has the largest entropy among all the random variables of equal variance, the NegEntropy can be considered as below;

$$J(x) = H(x_{gaussian}) - H(x) \tag{2.8}$$

whereas;
$x_{gaussian}$ - is a Gaussian random variable of some correlation (and covariance matrix) as x,
As $x_{gaussian}$ is largest of all random variables, the resulting entropy is always positive,

2.2.6 Mutual Information

Mutual Information between two random variables is the measure of the reduction of the randomness of a variable given knowledge of some other random variable and defined as below;

$$I(x, y) = H(x) - H(x|y) \tag{2.9}$$

whereas;
$I(x, y)$ - Mutual information between random variable x and y,
$H(x)$ is Entropy of random variable x,

$H(x|y)$ is a conditional Entropy of random variable x given y and written as.

$$H(x|y) = H(x,y) - H(y) \tag{2.10}$$

whereas;
$H(x,y)$ - Joint entropy of random variables x and y,
$H(y)$ is Entropy of random variable y

Using (2.9) and (2.10);

$$I(x,y) = H(x) + H(y) - H(x,y) \tag{2.11}$$

From (2.11), it can be observed that the objective to minimize mutual information $I(x,y)$ can be achieved by maximizing the joint entropy $H(x,y)$ [25].

Illustratively, it can be shown as below;

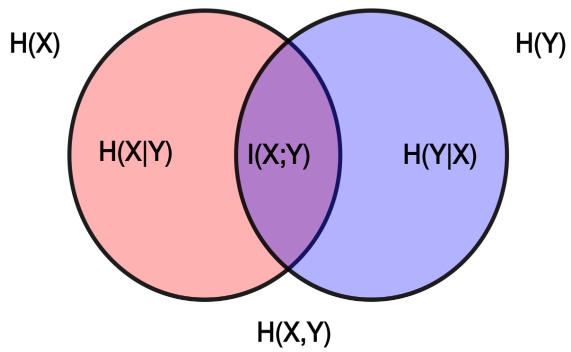


Figure 2.9: Mutual Information plot[7]

It is to be noted from the above plot that the higher the mutual information is, the higher the dependencies between random variables.

2.2.7 Assumptions made by ICA algorithms

The following assumptions need to be taken into account while applying Independent Components Analysis (ICA) to any specific applications;

- The underlying sources are statistically independent and the channel level recordings (observed random variables) are the linear combination of those original sources.
- The estimated independent components don't bring any order of importance on its own.

- To apply the ICA algorithm on any random set of random variables, out of all, more than one random variables cannot be following the Gaussian distribution due to its symmetric nature even after the rotation of the axis.
- In this work, the time delays from the true sources to the observations after mixing considered negligible.

2.3 Applying ICA to Neuronal signals processing

Several studies have been done applying Independent Component Analysis to the sets of human brain recordings which may be of non-invasive, semi-invasive or invasive in nature. Studies suggested for different sub-applications, one or more algorithm(s) perform better compared to other algorithms.
The comparison of ICA algorithms performance on the Electroencephalography (EEG) data for removing the artefacts shows that RADICAL algorithm performs better on the given data sets in which the comparison was done with the help of a measure, Mutual information based on k-neighbour statistics[18]. Recent works have tried to model the ICA for the random variables which are closely following the Gaussian distribution[15].

2.3.1 The cortical generators as Independent components

While ICA can be used to filter the artefacts, this method can as well be used to estimate the independent sources which are called cortical generators from the brain imaging data[6][32].

From the true sources until the channels observations, it is assumed to have the linear mixing which deters the original sources. So the recorded measurements at channel levels are having the linear transformation from the true sources. This also implies that the cortical source generators are considered to be statistically independent of one another while the recorded signals are the linear combinations of these sources. Statistically unmixing these combinations from the original sources is the main application of ICA.

The very same analogy is the core principle behind an ICA algorithm which makes it one of the most important algorithm to extract these statistically independent sources.

2.4 Evaluation of *Number* of Independent Components in ICA

By one of the assumptions, ICA results in the independent components without any specific order of importance, unlike PCA which orders the resulting components based upon the eigenvalue. While there are a number of efforts in optimally coming up with the number of independent components, one of them is used here for this work.

2.4.1 *icasso* package

The validation of Independent components is done with the help of *icasso* package[10]. While the estimation of the independent sources can be done with an outcome of a number of components which is as many as the number of channels measurements in the original recordings, the existence of additional optimization approach in combination with an evaluation method based on clustering techniques provide a reasonable accuracy over the selection of number of reliable estimated components/sources[11].

2.4.2 Statistical reliability

The statistical reliability of the sample makes the ICA with a biased estimation which can be corrected by bootstrapping the whole dataset. Even though the information contained in the data set remains the same, the sampling with replacement makes the estimation more reliable over different samples using different runs[11].

2.4.3 Algorithmic reliability

As ICA estimates the sources with optimization technique such as gradient descent, the initial parameters hold the key indication about the direction in which the optimization objective is met. ICA is a stochastic problem and gives different results every time it is run. Their results are different from the previous run. This stochastic algorithmic nature needs a careful approach for obtaining the reliable Independent components. With a single run, this will not be possible to find the global minima while trying to meet the objective function. One approach being, by randomizing the initial parameters over the number of runs may ensure that the ICA convergence captures many local minima. This process of smoothing across different local minima increases the reliability of the final estimates when combined with bootstrapping[11].
.

Figure 2.10 shows the possible many local minimas with respect to the different initial parameters[31];

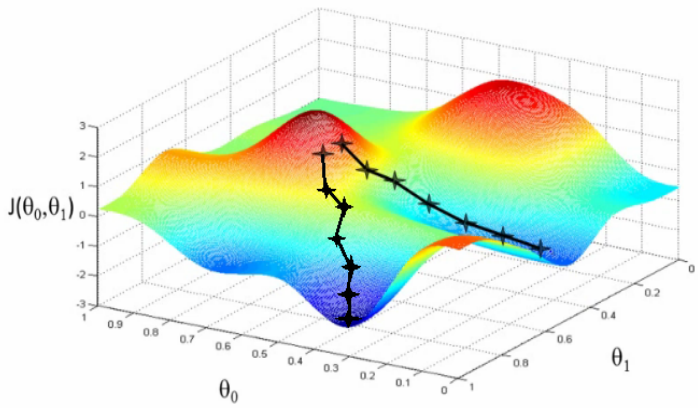


Figure 2.10: Local Minima Objective[31].

These two combined approaches remove the bias of having a single set of independent components over the whole sample but sets of bootstrapped samples[11].

2.4.4 Clustering

The clustering of all the estimated independent components from the number of runs is likely to give an indication of the reliable number of independent components and generates more reliable candidates to make an assessment

.

Clustering is an unsupervised learning technique which has been widely used in research and applications. As a result, it organizes all the input observations into a number of groups based upon each of the observations' characteristics on the multi-dimensional space. Space comes with the measure of the distance between the observations[10].

Figure 2.11 to 2.14 [29] shows the different types of clustering techniques and dendrogram for construction.

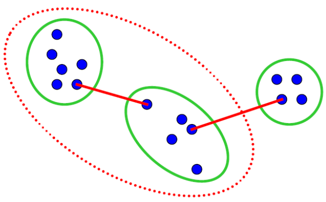


Figure 2.11: Single Linkage Clustering

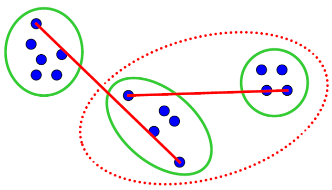


Figure 2.12: Complete Linkage Clustering

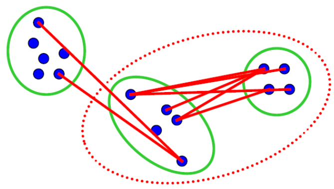


Figure 2.13: Average Linkage Clustering

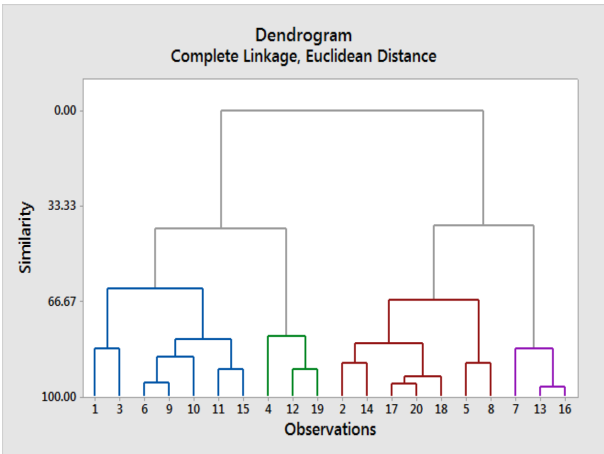


Figure 2.14: Clustering techniques

In this case, the clustering techniques in combination with the visual inspection of the clustering stability indexes for each formed cluster is likely to give an indication about the possible number of independent components. The principle being, even with randomization of the initial parameters and bootstrapping the samples over the number of runs using ICA algorithm, the estimated independent components are often predicted to be around at the high dimensionality space. The cluster formed with less number of independent components in a small area through the visual inspection in 2D space indicates that this cluster is likely to be unstable and implies that these few independent components might not be having any real significance, and thus the original number of independent components be reduced by one. The process goes on until we arrive at a reasonable number of clusters which are *compact* and *well isolated* from one another[10].

The estimated independent components at first partitioned with the help of the clustering technique. Average Linkage (AL) clustering technique has been used in this case. The dendrogram produced by this agglomerative type clustering technique shows how the cluster emerges with a single component and adding more and more to it based on the similarity measure[10].
The problem is still the evaluation of those generated clusters (groups) i.e whether each cluster is genuine, reliable and how many of them can really exist[29]. Quantifying them gives a way to know the estimate of the reliable number of (reduced) independent components. One such measure used is Clustering stability index for validating the quality of each cluster[11]
.

The clustering quality index *(Iq)* is the measure which is used to determine the compactness of each cluster. The more isolated the cluster it is with high compactness, it is likely to form an independent component itself. The quality of the clusters with the above definition can be accomplished by measuring the intra-cluster and inter-cluster distances in the cluster space. The following formula is used to calculate the Cluster Quality Index *(Iq)*, which is basically the difference between the average intra-cluster similarities and average extra-cluster similarities[10]

$$I_q(C_m) = \frac{1}{|C_m|^2} \sum_{i,j \in C_m} \sigma_{\mathrm{ij}} - \frac{1}{|C_m|C_-m|} \sum_{i \in C_m} \sum_{j \in C_-m} \sigma_{ij} \tag{2.12}$$

whereas;
C_m - set of components that belong the Mth cluster.
C_-m - set of components that do not belong to the Mth cluster
σ_{ij} - similarity of independent components matrix
σ_{ij}=$|r_{ij}|$
$|r_{ij}|$ is an absolute value of the mutual correlation coefficients with i,j=1,2,3...k.

From the above equation, it can also be noted that the Cluster Quality Index I_q with 1 is an ideal cluster which signifies an independent component. Less value of Iq implies that the cluster is less compact and homogeneous in a group.

A similarity measure which is used here is R-index which is defined as below[10][3];

$$I_R = \frac{1}{L}\sum_{m=1}^{L}\frac{\mathrm{S}_\mathrm{m}^\mathrm{in}}{\mathrm{S}_\mathrm{m}^\mathrm{ex}} \tag{2.13}$$

whereas;

$$\mathrm{S}_\mathrm{m}^\mathrm{in} = \frac{1}{|C_m|^2}\sum_{i,j\in C_m} d_{ij} \tag{2.14}$$

$$\mathrm{S}_\mathrm{m}^\mathrm{ex} = min\frac{1}{|C_m|C\_m'|}\sum_{i\in C_m}\sum_{j\in C\_m'} d_{ij} \tag{2.15}$$

whereas;
d_{ij} - similarity measure.

This $I_R$with minimum value indicates the compact and well-separated clusters. In addition, this type of clustering evaluation should be done not only using the quantitative evaluation but also using interactive visual inspection.

The Clustering validity index (I_q) is ordered from high to low before they are plotted against the independent components. The visual inspection of such a plot will give information about where there is a sudden drop in index happens. This package *icasso* is used to cluster and evaluate the number of clusters which gives information about the most reliable independent components.

2.4.5 Visualization of estimated independent components using 2-D projection

The clusters formed using the estimated independent components obtained by bootstrapping and randomization, can be visualized. However, there exists difficulty in visualizing the independent components due to its nature of the curse of dimensionality. While it is comparatively easy to visualize the parameters up to 3-dimensional space, as the number of dimensions increase, the difficulty in interpretation increases leaving the purpose more grey than precise. This needs techniques using which we can visualize the very high dimensional space into 2 dimensions. Once such technique used is called Curvilinear Components Analysis (CCA)[10].

2.4.6 Revision of number of Independent Components

Evaluation of the number of clusters provide information about the number of reliable and strong impact making independent components using which we can re-estimate the independent components[10].

Once the visual inspection is done in combination with the quality measures are analyzed, the possible number of independent components can be given as an input to the ICA algorithm[10].

As an effect of having revised the number of independent components, the small clusters merge to form big clusters as an outcome of the reduction in the input of Independent components[10].

2.5 Evaluation of *quality* of Independent components obtained from ICA algorithm

The evaluation of the quality of the independent components is done with the help of the Mutual Information, Spearman correlation coefficient, Kurtosis and Entropy measures[14][10].

2.6 Conclusion

In this chapter, the works related to the goal of this thesis work has been presented with the background information needed to read this work. It included the *icasso* package, evaluation measures used for ICA.

Chapter 3

Methods

3.1 Experimental Set up

The experiment included five patients at resting state recorded continuously for three minutes using an invasive ECoG procedure. The electrodes are placed over the left and right temporal cortex, and depth electrodes were implanted in left and/or right hippocampus. Electrodes are all positioned according to clinical standards.

The grid architecture contains the following;

Parameter	ECoG
Distance between channels, mm	10

Table 3.1: Grid parameters

3.2 State of recording

The resting state was used for all subjects.

3.3 Procedure

The detailed procedure is explained below;

1. The data is pre-processed to filter the recordings between 0.1 and 100 H_z.
2. The mutual information, Spearman correlation coefficient are estimated at the channel level recordings i.e before applying ICA algorithm.
3. The pre-processed data is then used for estimating the Independent components with the help of the package *icasso*[10]. The process includes the randomization of the initial parameters for optimization and bootstrapping the samples so

to avoid the single local minima. The procedure is run 100 times keeping the number of independent components as same as the original data for every run.

4. Similarities between the independent components are estimated and average linkage clustering is used to cluster the calculated independent components which are obtained in all the runs.

5. All the estimated independent components from all different runs are visualized in 2-dimensional space with the help of Curvilinear Component Analysis(CCA).

6. The validity of the clusters are evaluated with the help of an index called Clustering Validity Index while visually inspecting for the sudden drop in the index with respect to an increase in the number of clusters.

7. The number of clusters centred around over the various runs are considered to provide the reasonable number of independent components which in turn used to re-define the number of components in ICA algorithm as an input.

8. Once the number of independent components are validated, the quality of independent components are measured with the help of estimation of Mutual Information, Spearman correlation coefficient at the components level i.e after applying ICA [32].

9. The estimated independent components are ordered for its importance with the help of Entropy and Kurtosis using which the results are compared and discussed accounting the effect of maximizing the joint entropy in the objective functions.

3.4 ICA Algorithms

Once the data is pre-processed, the following ICA algorithms are used;

1. FastICA
2. InfoMax
3. Jade
4. OGWE
5. Pearson
6. EFICA

Each algorithm is discussed briefly in the following sections.

3.4.1 FastICA

FastICA algorithm works on maximizing non-gaussianity by forcing each of the components to be as far as from the normal distribution possible[13]. It uses neg-entropy as a measure while doing so.

$$N(x) = H(x_{gaussian}) - H(x) \tag{3.1}$$

whereas;
x - is a random vector known to be non-gaussian,
$H(x)$ is the entropy,
$H(x_{gaussian})$ is the entropy of a Gaussian random vector whose covariance matrix is equal to that of x.
Instead of using the above equation due to its associated difficulties, the approximations are used.

The algorithm works as below[27];

1. Initialize w_i (e.g random matrix)
2. $w_i{+}{=}E(\pi(w_i^T X))w_i - E(xpi(w_i^T X))$
3. $w_i{=}\frac{\mathrm{w_i^+}}{||\mathrm{w_i^+}||}$
4. For $i = 1$, go to step 7, Else, continue with step 5.
5. $w_i^+{=}w_i - \sum_{j=1}^{i-1} W_i^T w_j w_j$
6. $w_i{=}\frac{\mathrm{w_i^+}}{||\mathrm{w_i^+}||}$
7. If not converged, go to Step 2, Else go to Step 1 with $i{=}i{+}1$ until all components are extracted

whereas;
wi- is a column vector of the unmixing matrix W,
wi^+ - is a temporary variable used to calculate wi (it is the new wi before normalization)
π_i is the derivation of $\pi(.)$ and $E(.)$ is the expected value (mean).

3.4.2 InfoMax

InfoMax algorithm carries the objective function to minimize the mutual information by maximizing the joint entropy between two random variables. It uses the objective function which is logistic infomax function and the optimization algorithm which is stochastic gradient descent. InfoMax algorithm works as below[27];

1. Initialize $W(0)$ (i.e random matrix)

2. $W(t+1) = W(t) + \mu(t)(I - f(Y)Y^T)W(t)$

3. If not converged, then Step 2.

whereas;
t - represents a given approximation step,
Mu(t) - is a general function that specifies the size of the steps for the unmixing matrix updates (usually an exponential function or a constant)[27].
f(Y) - Non-linear function chosen according to the type of distribution.
I - Identity matrix of dimensions mxm.

3.4.3 JADE

Joint Approximate Diagonalization of Eigen matrices algorithm(JADE) is a blind source separation method based on diagonalization of fourth-order cumulant tensor[14]

The algorithm works as below[5];

1. Form the covariance matrix R, and compute a whitening matrix W

2. Form the sample forth order cumulants Z_z of the whitened process $z(t)$=$Wx(t)$; compute the n most significant eigenpairs $gamma, M_r|1 <= r <= n$.

3. Jointly diagonalise the set N^e=$gamma, M_r|1 <= r <= n$ by a unitary matrix U.

4. An estimate of A is A=WU.

3.4.4 OGWE

OGWE is an optimized Sinusoidal ICA/BSS algorithm and proven to be well performing in achieving statistical independence while estimating the independent components involving several applications[24]. It uses Marginal entropy in achieving statistical independence.

3.4.5 Pearson

The Pearson algorithm takes into account the measure 'skewness' while using maximum likelihood function as an objective function with the help of optimization by Hyvarinen's fixed point algorithm. The procedure is as below[16];

1. Calculate the third and fourth sample moments α_3 and α_4 for current data y_k=$W_x x$ and select the Pearson system or fixed *(tanh)* contrast according to figure3.1.

2. If the Pearson system was selected, estimate parameters of the distribution by the method of moments.

3. Calculate scores $\varphi(y_k)$ for the Pearson system or fixed contrast.

4. Calculate the demixing matrix W_{k+1} using algorithm for fixed point [16] or natural gradient descent algorithm [1].

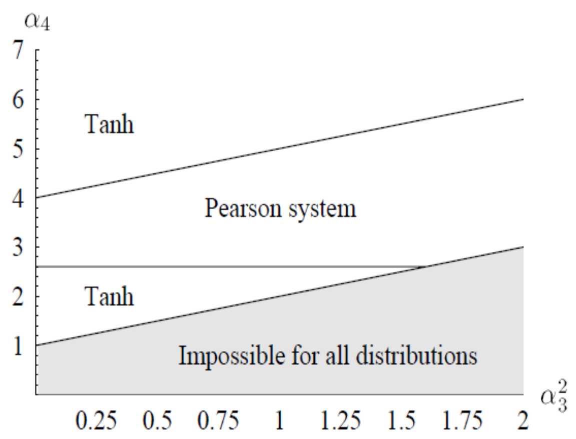


Figure 3.1: The Contrast function for Pearson ICA

The contrast functions used in the pearson ICA are presented in (α_3^2, α_4)plane. Limit for all distributions in line α_4=$\alpha_3^2 + 1$. Clearly sub-gaussian sources are defined to have kurtosis less than 2.6 and the tanh-contrast is utilized. Clearly super-gaussian sources are defined to have α_4=$\alpha_3^2 + 4$ and the *tanh* contrast is again utilized. In area between these boundaries the pearson system is used. The choice of boundaries is based on practical experience[16].

3.4.6 EFICA

The algorithm[17]consists of the following steps;

1. Run FastICA until convergence

2. Adaptive choice of different non-linearities for estimating the score functions of the sources from Step.1

3. Refining for each of the estimated component by one-unit FastICA using the non-linearities found in Step.2

4. Fine tuning additional parameters.

3.5 Evaluation Measures

The evaluation measures are those ones which are used for statistical and linear independence which are as below;

1. Mutual Information - Statistical Independence

2. Spearman correlation co-efficient - Linear Independence

3. Kurtosis - Statistical Independence
4. Entropy,Joint entropy - Statistical Independence

Chapter 4

Results and Discussion

4.1 Datasets

The datasets consists of human brain recordings with the following dimensions for all the considered five subjects;

Description	***Size***
Subject 1	52x46081
Subject 2	48x46276
Subject 3	20x46081
Subject 4	44x185098
Subject 5	22x46275

The recordings differ in the number of channels and the time stamps.

4.2 Evaluation of *Number* of Independent Components in ICA

The package *Icasso* is used to find the independent components and evaluation of those components keeping the convergence statistically and algorithmically efficient. *Icasso* comes with the ICA implemented module FastICA whereas it was applied with the following parameters for all subjects.

Subjects	***1-5***
Number of Samples for Bootrapping	100
Random Initialization	Yes
Contrast function for choosing the non-linearity	*tanh*
Approach	Symmetric
Maximum number of fine tuning	1000
Maximum number of iterations	100

Once all the independent components are obtained from all different runs, the evaluation is done with the help of Clustering technique, called Average Linked (AL) Clustering. The calculated independent components over 100 different samples each time with randomized initial parameters are used to calculate the similarity among them which are used to arrive at the reasonable number of clusters. And then the estimated independent components are visualized in two dimensions using the Curvilinear Component Analysis (CCA) technique as below for Subject-1.

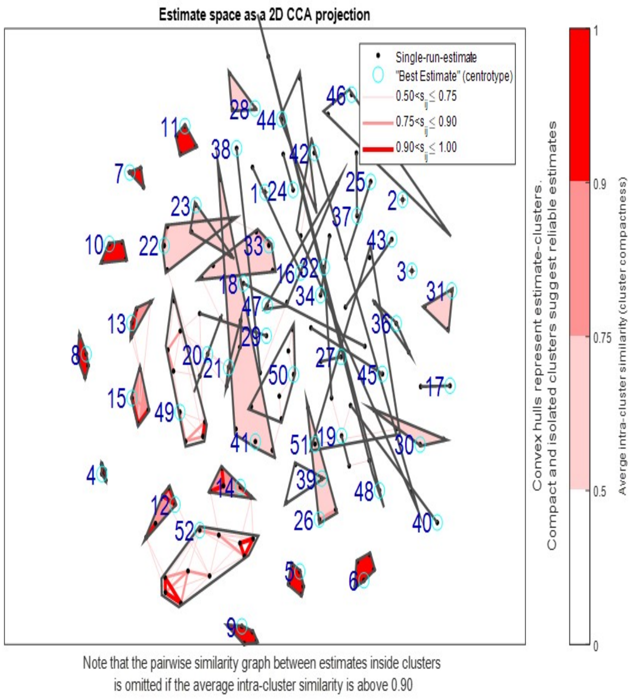


Figure 4.1: Clusters Visualization for ECoG,Subject-1

As it is seen in the above visualization of the clusters, some of them are very compact and well isolated while some are either less compact or in less proximity with each other. The compactness implies how good the estimated components on the original dimensional space are closer by and would signify the similar properties, which prompts for to be considered as a single Independent component in the reduced dimensions. The following Figure shows the Cluster Stability Index I_q which quantitatively measures the above properties.

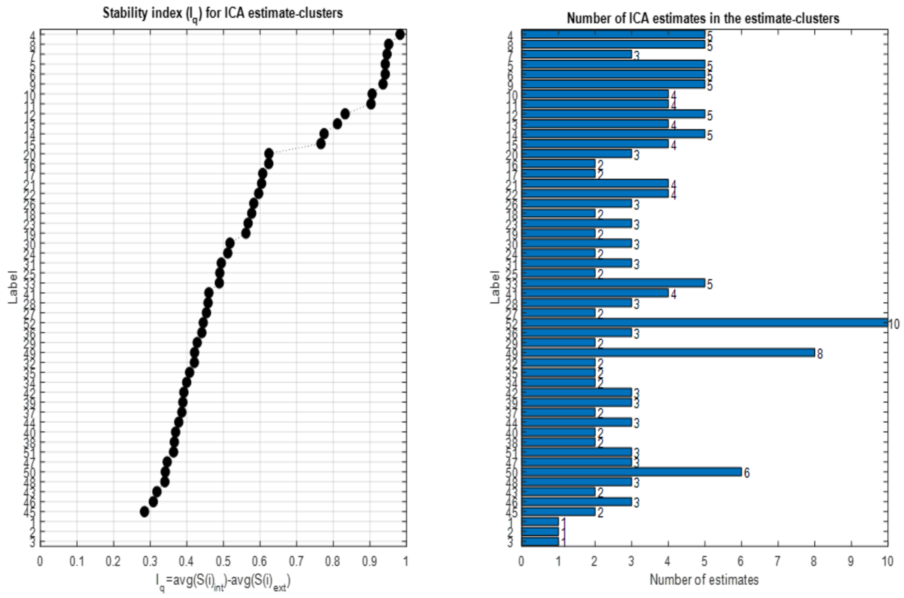


Figure 4.2: Clusters Stability Index for ECoG

The notion behind looking at the compactness and isolation of the clusters is, irrespective of the bootstrapping and random initialization over the number of trials, the independent components fall around in or around the same space.

Here it can be visually seen that the sudden (first) drop in the stability index after six number of clusters which is an indication about the number of reliable estimates for the number of independent components with stability. Based upon the above evaluation with the help of the clustering technique, the number of independent components is estimated to be of six. The summary of the original dimensions and the Independent sources are shown in the following table for all five subjects.

Subjects	***1***	***2***	***3***	***4***	***5***
Number of Channels	52	48	20	44	22
Number of Independent components	6	6	3	22	12

Table 4.1: Number of Channels and Number of Independent Components

It can be seen that the final number of total independent components for each subject is between 3 to 22 while the majority of them lie between 3 to 12. For subject four, there is not a clear indication of the reliable number of components as seen here below;

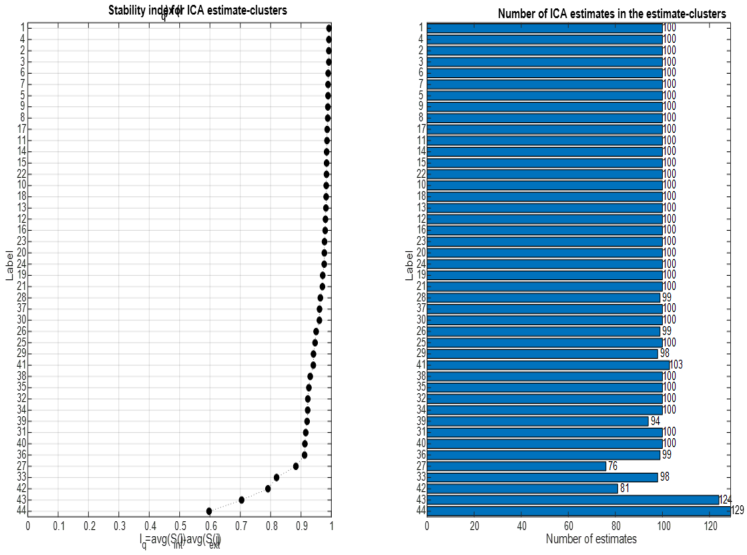


Figure 4.3: Stability Index, Subject 4

Such inconsistency may be associated with the complexity of obtaining the sources with larger number of components in combination with the high number of observations (i.e subject four has threefold of time stamps compared to other subjects). These evaluations shall be compared to the independent components projection back to the channels[32]. The column vector of the unmixing matrix is plotted to show the intensity of projection on each channel as below for each component.

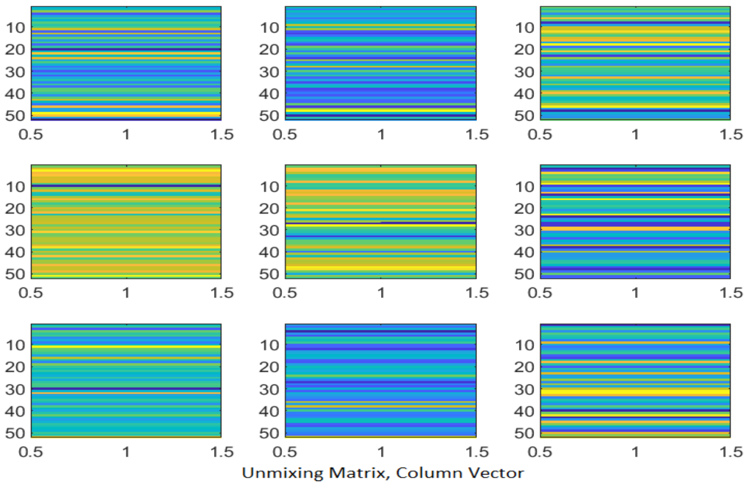
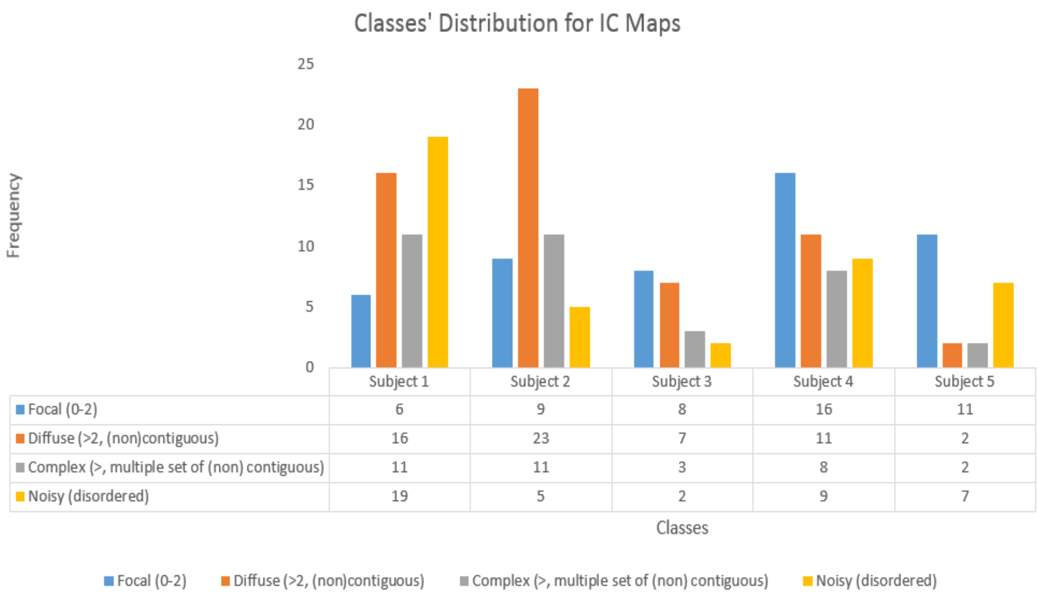


Figure 4.4: Projection of Unmixing matrix on Channels by Components 1-9, Subjects 1

Once the individual independent components are projected at the level of the channel, visual inspection is made to observe the intensity of a component to a particular channel and different classes are considered [32];

- Focal - An Independent component considered to be very active on two channels.
- Diffuse - An independent component considered to be very active in more than two channels sparingly
- Complex - An independent component considered to be very active in more than two channels multiple places across
- Noisy - Disordered without any specific interpretation

For all subjects, the number of components obtained are classified as below;



	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
Focal (0-2)	6	9	8	16	11
Diffuse (>2, (non)contiguous)	16	23	7	11	2
Complex (>, multiple set of (non) contiguous)	11	11	3	8	2
Noisy (disordered)	19	5	2	9	7

Figure 4.5: Classification of Components by Projection's intensity at Channels, Subjects 1-5

This classification shows that the distribution of a number of reliable components obtained using *icasso* package is closer to how the most probable projection looks using the above visual classification.

4.3 Evaluation of *quality* of Independent components obtained from ICA algorithm

The comparison between different ICA algorithms is done with the help of the evaluation measure such as Mutual Information for Statistical independence, Spearman correlation coefficient for linear independence, Kurtosis and Entropy measures for non-gaussianity.

4.3.1 Mutual Information

The Mutual Information (MI) between the electrodes give an indication about the shared information between channels (Sources). Once the data is pre-processed, the Mutual information shared between the recordings at channel levels is compared to the estimated Independent sources. The estimated MI at channels level are shown below;

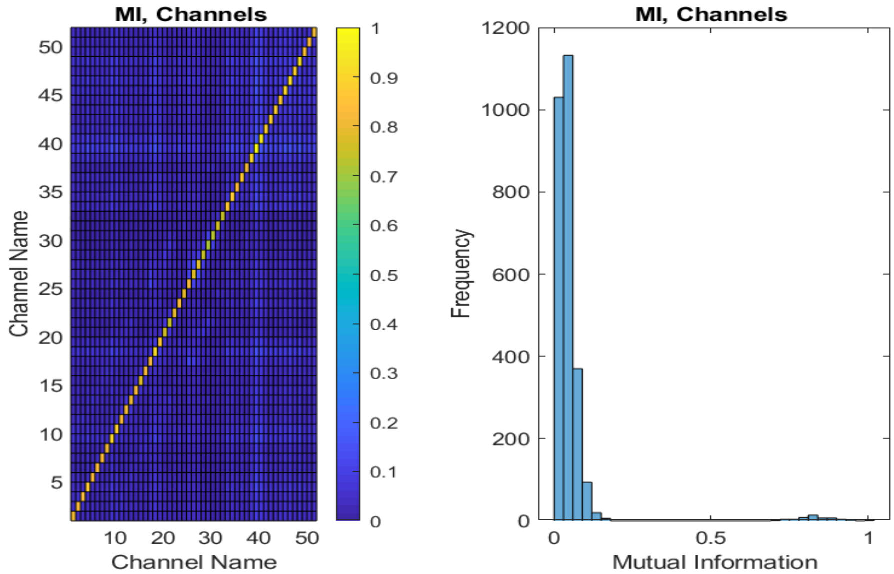


Figure 4.6: MI at Channels level

The calculated mutual information values are left-skewed as shown in the above distribution. The higher the Mutual Information, the higher the channels recordings are dependent. The objective, as defined, for an ICA algorithm is to reduce the mutual information between the channels level recordings *i.e increasing the statistical independence*.

Once the distribution of pairwise mutual information on the level of the channel recordings is observed, they are estimated at the components level. The number of independent components is kept as same as the original number of channels. This is done to ensure the objectivity on the comparison and account for the fact that some family of algorithms are considered to be underperforming when the number of independent components increases[7].

. The distribution of the mutual information at the components level are shown below;

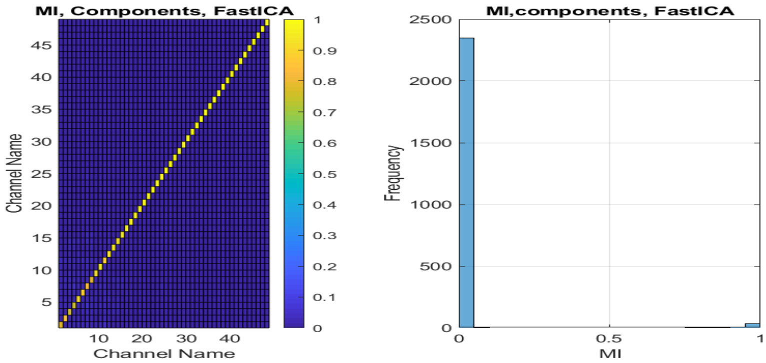


Figure 4.7: Mutual Information at Components level, FastICA

As it can be seen from the above distribution, the pairwise mutual information between the components are less than those of estimations at the channels level recordings using FastICA algorithm. Very similar observations can be found for all other ICA algorithms.

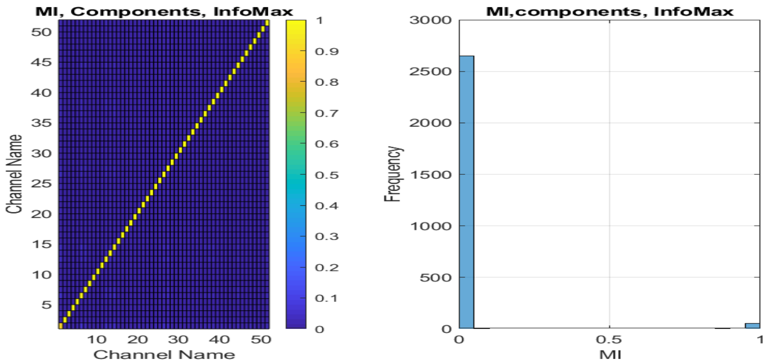


Figure 4.8: Mutual Information at Components level, InfoMax

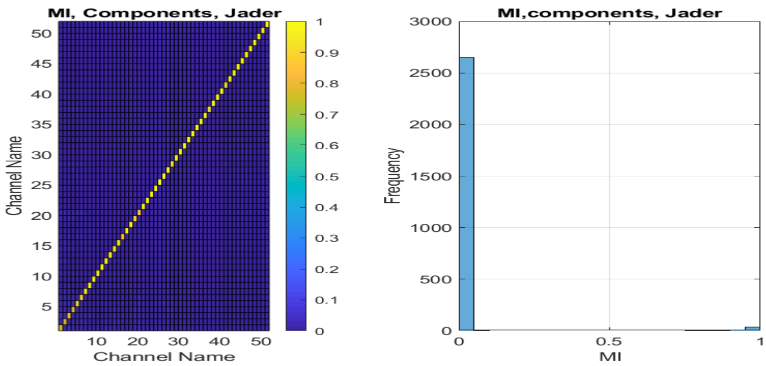


Figure 4.9: Mutual Information at Components level, Jader

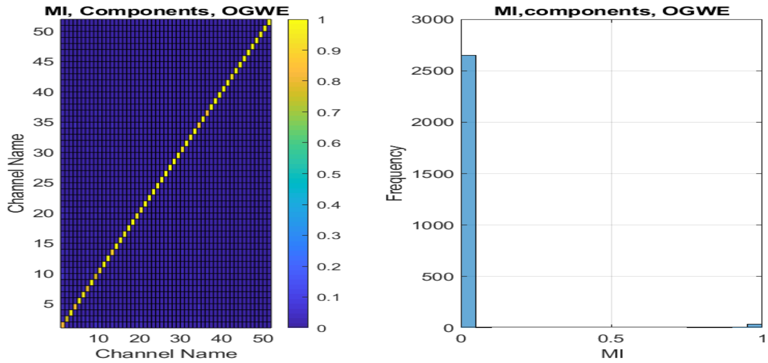


Figure 4.10: Mutual Information at Components level, OGWE

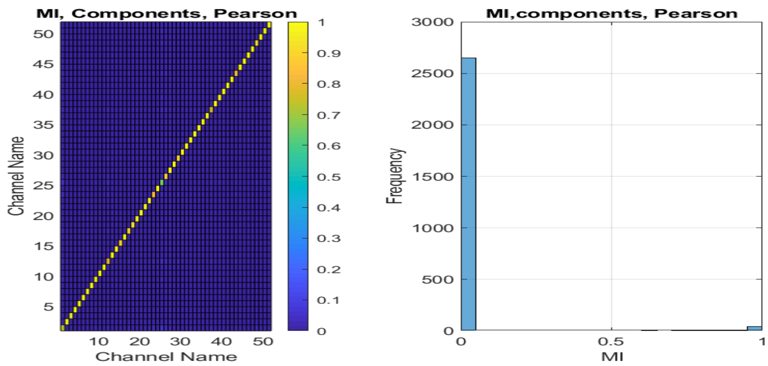


Figure 4.11: Mutual Information at Components level, Pearson

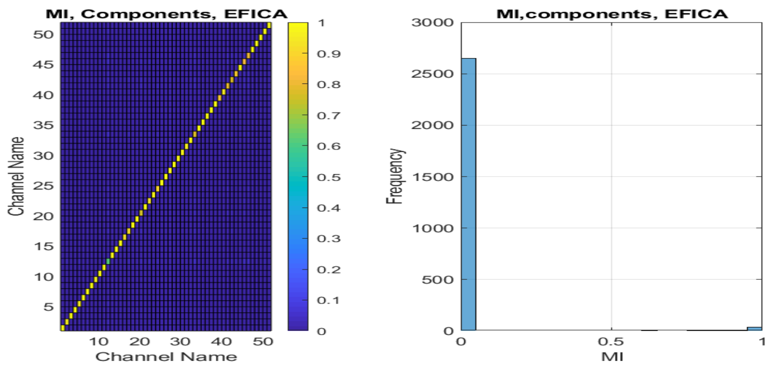


Figure 4.12: Mutual Information at Components level, EFICA

It can be observed that all the ICA algorithms meet its objective function in reducing the mutual information.

4.3.2 Kurtosis and Entropy

In this quality assessment, the estimated independent components are ordered for their importance using the measure of Kurtosis indicating the non-gaussianity. While Entropy itself does not constitute for measuring the non-gaussianity (Negentropy being a measure for non-gaussianity), it is, however, irrespective of the sign, a measure using which we can order the importance for an extent towards non-gaussianity. Individual entropies can as well be compared with the joint entropies as they both contribute for a reduction in the mutual information. These estimates are done for all the five different subjects and are shown below from 4.13 to 4.17.

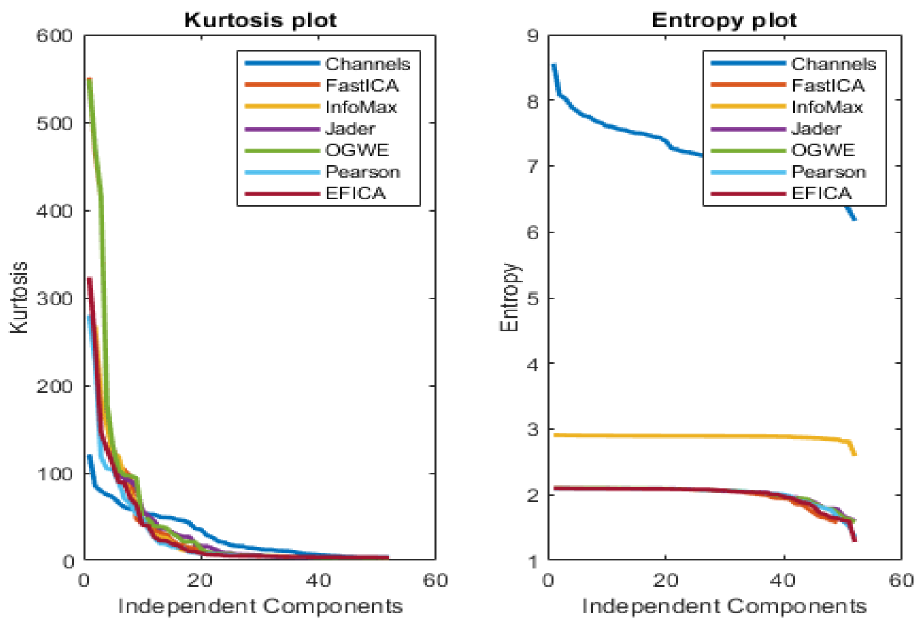


Figure 4.13: Components ordered for Kurtosis and Entropy,Subject-1.

The higher the component's Kurtosis, higher they are non-gaussian. For Channels recordings, the Kurtosis values are less compared to those of original sources. The lesser the entropy, the higher the components of nongaussian

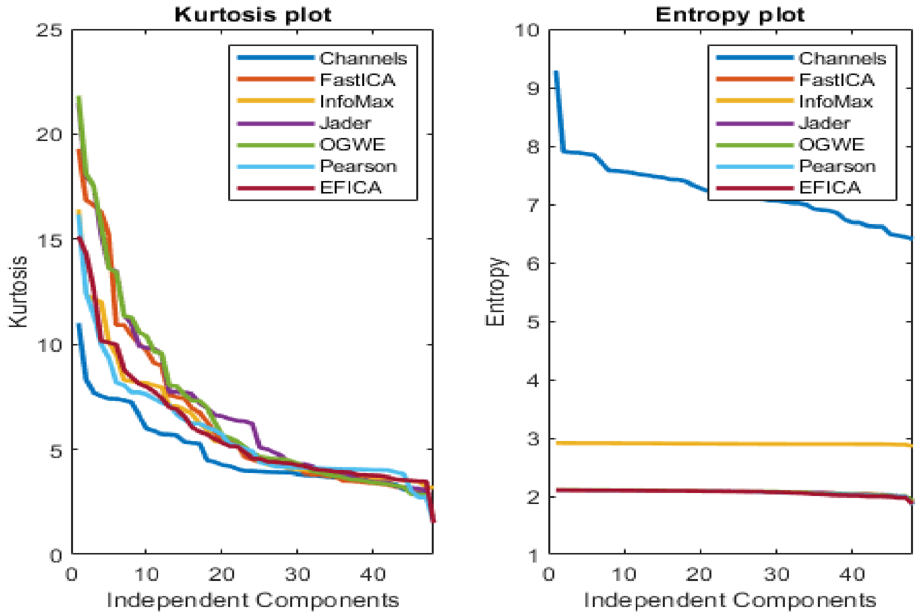


Figure 4.14: Components ordered for Kurtosis and Entropy,Subject-2

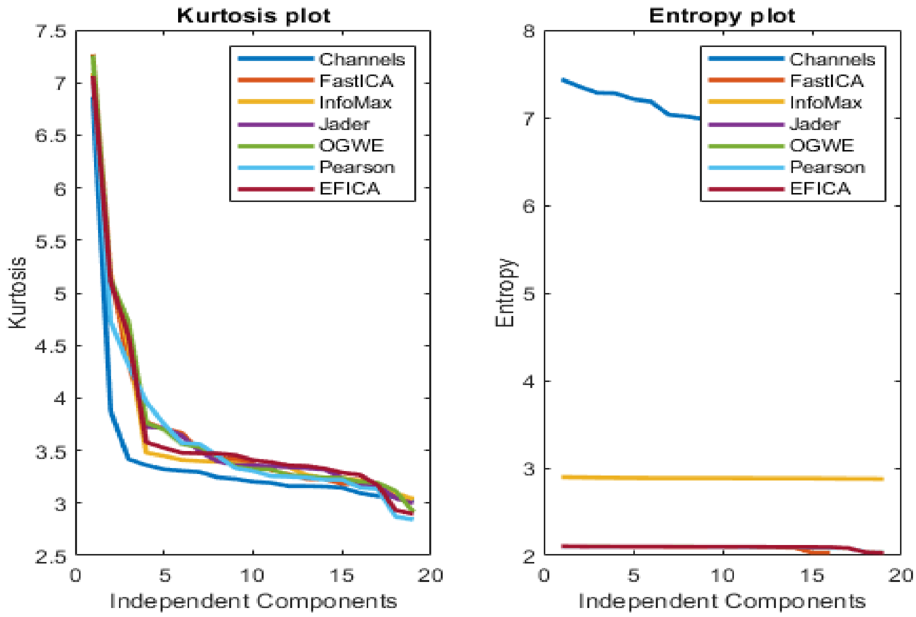


Figure 4.15: Components ordered for Kurtosis and Entropy,Subject-3

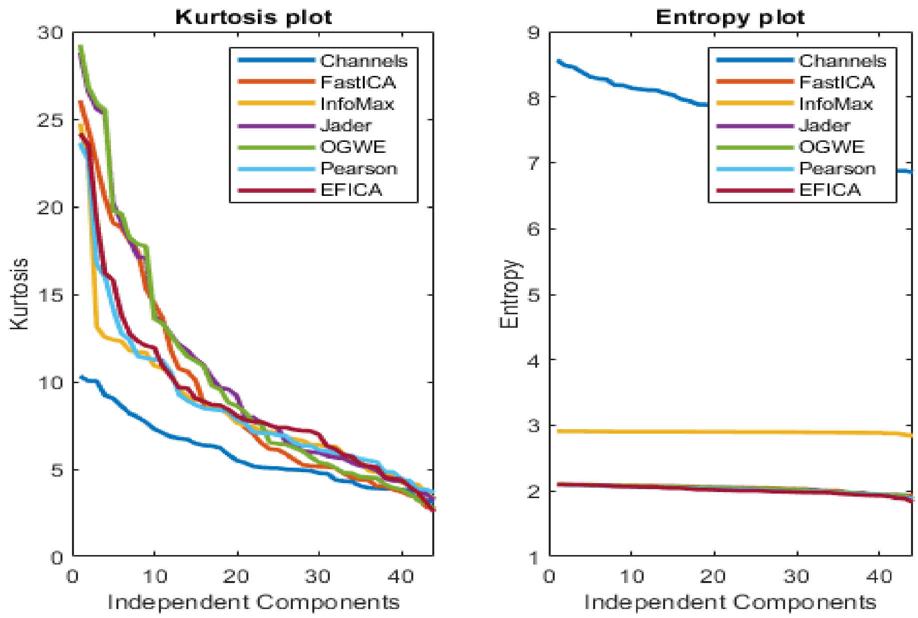


Figure 4.16: Components ordered for Kurtosis and Entropy,Subject-4

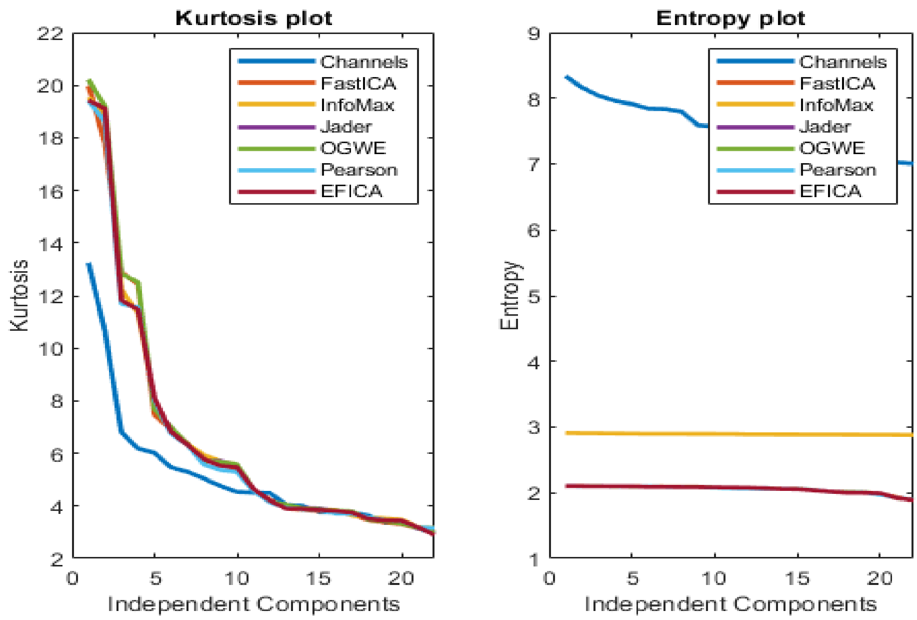


Figure 4.17: Components ordered for Kurtosis and Entropy, Subject-5

OGWE consistently produces independent components with higher Kurtosis maximizing the non-gaussianity. It is also observed that the FastICA family of algorithms (including EFICA), in all cases, closely follows in achieving the objective function of maximizing the statistical independence for estimated components. There are insignificant to moderate differences when it comes to estimating the individual kurtosis measure by different ICA algorithms except for InfoMax. As there is an issue of outliers with Kurtosis measure, applying such measure to ECoG data needs additional evaluation measures due to the nature of spikes in these data points[7].

Using the individual measure of entropies on the estimated independent components, except InfoMax, all of the other used algorithms outperforms in a decisive way when it comes to maximizing the non-gaussianity by extracting the components with smaller entropies. As the reduction of the mutual information is depending upon the individual component's entropies and the joint entropy between the components (2.11), the joint entropy results are compared with the above results and presented here below;

Description	***Channels***	***FastICA***	***InfoMax***	***Jader***	***OGWE***	***Pearson***	***EFICA***
Subject 1	0,056	3,981	5,686	3,989	3,981	3,955	3,933
Subject 2	0,067	4,092	5,724	4,088	4,089	4,072	4,069
Subject 3	0,074	4,051	5,612	4,074	4,076	4,073	4,076
Subject 4	0,104	4,016	5,690	4,003	4,002	3,968	3,953
Subject 5	0,090	4,012	5,642	4,010	4,010	4,004	4,008

Table 4.2: Mean Joint Entropy between estimated Components

It is observed that, out of all chosen algorithms, InfoMax performs better when it comes to the maximization of the joint entropy between the independent components which is its main objective function. This objective of joint entropy maximization is achieved by InfoMax, is better than other algorithms with a considerable difference. Followed by InfoMax, the other chosen algorithms come closely between them. The above results along with those from figures 4.13 to 4.17 reasons the deficiency of the InfoMax algorithm and aligned with those of this algorithm's pitfall in dealing with interdependencies i.e maximizing the joint entropy does not guarantee to be minimal[2].

4.3.3 Mean Mutual Information, Spearman Co-efficient, Kurtosis, and Entropy measures

The mean mutual information values for all the five subjects indicate that using InfoMax algorithm or FastICA algorithms, do not always give the best performing results compared to other listed algorithms i.e Jader, OGWE, Pearson,EFICA for which there are negligible differences across these. This is in contrast to the similar studies results where InfoMax outperforms most of the other algorithms[8] and likely attributed to the fact of higher individual estimated components entropies which are reported in this work. The estimated mean mutual information values as an outcome of different ICA algorithms are as below;

Description	***Channels***	***FastICA***	***InfoMax***	***Jader***	***OGWE***	***Pearson***	***EFICA***
Subject 1	0,0557	0,0226	0,0237	0,0223	0,0223	0,0222	0,0221
Subject 2	0,0668	0,0238	0,0237	0,0237	0,0237	0,0237	0,0236
Subject 3	0,0738	0,0629	0,0531	0,0530	0,0530	0,0530	0,0529
Subject 4	0,1036	0,0280	0,0300	0,0277	0,0278	0,0273	0,0280
Subject 5	0,0900	0,0462	0,0472	0,0461	0,0461	0,0462	0,0462

Table 4.3: Mean Mutual Information between estimated components

While the differences between the mutual information estimates obtained using different algorithms look smaller than those of individual entropies, the results are aligned in which the OGWE, Pearson, Jader and EFICA algorithms perform closely compared to other chosen algorithms for this work. As expected the mutual information reduction is evident for all the algorithms. The differences are noted very less in comparison.

Description	***Channels***	***FastICA***	***InfoMax***	***Jader***	***OGWE***	***Pearson***	***EFICA***
Subject 1	0,2236	0,0195	0,0192	0,0199	0,0192	0,0195	0,0192
Subject 2	0,2134	0,0209	0,0207	0,0214	0,0204	0,0213	0,0209
Subject 3	0,0166	0,0627	0,0531	0,0524	0,0523	0,0523	0,0530
Subject 4	0,1164	0,0233	0,0239	0,0233	0,0225	0,0225	0,0226
Subject 5	0,1034	0,0467	0,0464	0,0458	0,0457	0,0447	0,0453

Table 4.4: Mean Spearman Correlation Coefficient between estimated components

There is a reduction of Spearman correlation coefficient across subjects compared to channels level recordings for linear dependence except for the subject-3. The Spearman correlation coefficient shows inconsistent nature for specific recording i.e Subject-3 for repeated runs. This observation shows that there is no evidence linear dependence is not a strong and evident measure in analysing the quality of the sources.

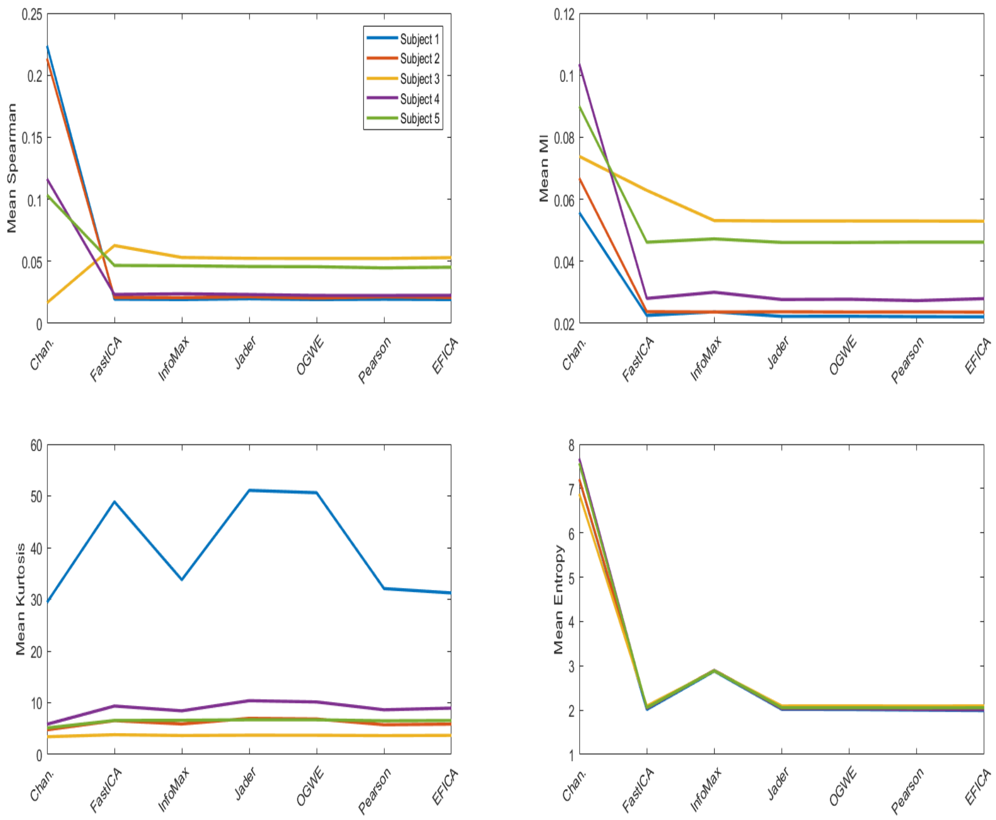


Figure 4.18: Comparison of different measures across different ICA algorithms,Subject-1-5

- The OWGE algorithm performs well overall for all subjects or closely performs to other best-performing algorithms on the Mutual information measure, closely followed by other algorithms, and lastly InfoMax.
- The reduction on the linear dependence using the Spearman correlation coefficient is efficiently achieved using again OGWE.
- However, it has been noted that the Spearman correlation coefficient does not provide any justifiable little information about statistical independence as the estimated measures at the level of the channel and components level look inconsistent at least for one of the Subject i.e Subject-3.

Chapter 5

Conclusion

This study of evaluation of independent components for the resting state Electrocorticography (ECoG) provides the total number of reliable estimates and different quality measures in comparing the different available Independent Component Analysis(ICA) algorithms. The results prove that the total number of estimated independent components range between 3 to 12 for five different subjects for up to 52 original number of observations using *icasso* package. The Projection of independent components on the channels and the visual inspection of them shows a not exactly but similar reduced active sites compared to those of the original number of channels recordings. It is also noted that with an increased number of timestamps for observed values, there exists inconsistency in obtaining the reliable estimates and this may be likely due to the restricted time windows which are more dependent.

Out of all used algorithms which are FastICA, InfoMax, Jader, OGWE, Pearson EFICA, measures for statistical independence are best achieved by the OGWE algorithm closely followed by the algorithms other than FastICA and InfoMax. Out of the chosen algorithms, it is also observed that the InfoMax algorithm does not perform as other chosen algorithms due to its interdependencies between the joint entropy and the individual components entropies while minimizing the mutual information for all the subjects. Without any solid evidence, in Electrocorticography (ECoG) recordings, the linear dependence is a far away measure in any form to discuss relating it to the statistical independence due to its inconsistency for subject-3.

Bibliography

[1] S.-I. Amari. Natural gradient works efficiently in learning. *Neural computation*, 10(2):251–276, 1998.

[2] A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6):1129–1159, 1995.

[3] J. C. Bezdek and N. R. Pal. Some new indices of cluster validity. 1998.

[4] J.-F. Cardoso. Blind separation of real signals using JADE. http://read.pudn.com/downloads158/sourcecode/math/704523/jader.m__.htm, 1997. [Online; accessed 07-May-2019].

[5] J.-F. Cardoso and A. Souloumiac. Blind beamforming for non-gaussian signals. In *IEE proceedings F (radar and signal processing)*, volume 140, pages 362–370. IET, 1993.

[6] C. Chintaluri and D. K. Wójcik. A novel method for spatial source localization using ecog and seeg recordings in human epilepsy patients. *BMC neuroscience*, 16(1):P286, 2015.

[7] M. X. Cohen. *Analyzing neural time series data: theory and practice.* MIT press, 2014.

[8] A. Delorme, J. Palmer, J. Onton, R. Oostenveld, and S. Makeig. Independent eeg sources are dipolar. *PloS one*, 7(2):e30135, 2012.

[9] R. Enatsu and N. Mikuni. Invasive evaluations for epilepsy surgery: a review of the literature. *Neurologia medico-chirurgica*, 56(5):221–227, 2016.

[10] J. Himberg and A. Hyvarinen. Icasso: software for investigating the reliability of ica estimates by clustering and visualization. In *2003 IEEE XIII Workshop on Neural Networks for Signal Processing (IEEE Cat. No. 03TH8718)*, pages 259–268. IEEE, 2003.

[11] J. Himberg, A. Hyvärinen, and F. Esposito. Validating the independent components of neuroimaging time series via clustering and visualization. *Neuroimage*, 22(3):1214–1222, 2004.

[12] R. Hindriks, C. Micheli, C. Bosman, R. Oostenveld, C. M. Lewis, D. Mantini, P. Fries, and G. Deco. Reconstruction of resting-state networks from macaque electrocorticographic data. *bioRxiv*, page 221051, 2017.

[13] A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks*, 10(3):626–634, 1999.

[14] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent component analysis*, volume 46. John Wiley & Sons, 2004.

[15] A. Hyvärinen, P. Ramkumar, L. Parkkonen, and R. Hari. Independent component analysis of short-time fourier transforms for spontaneous eeg/meg analysis. *NeuroImage*, 49(1):257–271, 2010.

[16] J. Karvanen, J. Eriksson, and V. Koivunen. Pearson system based method for blind separation. In *Proceedings of Second International Workshop on Independent Component Analysis and Blind Signal Separation (ICA2000), Helsinki, Finland*, pages 585–590, 2000.

[17] Z. Koldovsky, P. Tichavsky, and E. Oja. Efficient variant of algorithm fastica for independent component analysis attaining the cramér-rao lower bound. *IEEE Transactions on neural networks*, 17(5):1265–1277, 2006.

[18] V. Krishnaveni, S. Jayaraman, P. M. Kumar, K. Shivakumar, and K. Ramadoss. Comparison of independent component analysis algorithms for removal of ocular artifacts from electroencephalogram. *Measurement Science Review*, 5(2):67–78, 2005.

[19] J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela. Measuring phase synchrony in brain signals. *Human brain mapping*, 8(4):194–208, 1999.

[20] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski. Independent component analysis of electroencephalographic data. In *Advances in neural information processing systems*, pages 145–151, 1996.

[21] S. Makeig, T. Bell, T. Lee, T. Jung, S. Enghoff, et al. Eeglab: Ica toolbox for psychophysiological research. *WWW Site, Swartz Center for Computational Neuroscience, Institute of Neural Computation, University of San Diego California*, 2000.

[22] V. Matic, W. Deburchgraeve, and S. Van Huffel. Comparison of ica algorithms for ecg artifact removal from eeg signals. In *Proc. of the 4th Annual symposium of the IEEE-EMBS Benelux Chapter.(IEEE-EMBS)*, pages 137–140, 2009.

[23] L. Muller, L. S. Hamilton, E. Edwards, K. E. Bouchard, and E. F. Chang. Spatial resolution dependence on spectral frequency in human speech cortex electrocorticography. *Journal of neural engineering*, 13(5):056013, 2016.

[24] J. J. Murillo-Fuentes and F. J. González-Serrano. A sinusoidal contrast function for the blind separation of statistically independent sources. *IEEE transactions on signal processing*, 52(12):3459–3463, 2004.

[25] G. R. Naik and D. K. Kumar. An overview of independent component analysis and its applications. *Informatica*, 35(1), 2011.

[26] J. M. Palva, S. H. Wang, S. Palva, A. Zhigalov, S. Monto, M. J. Brookes, J.-M. Schoffelen, and K. Jerbi. Ghost interactions in meg/eeg source space: A note of caution on inter-areal coupling measures. *Neuroimage*, 173:632–643, 2018.

[27] G. SAHONERO-ALVAREZ, B.-g. La Paz, H. CALDERON, and B. La Paz. A comparison of sobi, fastica, jade and infomax algorithms.

[28] H. G. J. H. J. Sarela and A. Hyvarinen. Fast Independent Component Analysis. https://research.ics.aalto.fi/ica/fastica/, 2005. [Online; accessed 08-May-2019].

[29] A. Saxena, M. Prasad, A. Gupta, N. Bharill, O. P. Patel, A. Tiwari, M. J. Er, W. Ding, and C.-T. Lin. A review of clustering techniques and developments. *Neurocomputing*, 267:664–681, 2017.

[30] P. Tichavskỳ, Z. Koldovskỳ, and E. Oja. Speed and accuracy enhancement of linear ica techniques using rational nonlinear functions. In *International Conference on Independent Component Analysis and Signal Separation*, pages 285–292. Springer, 2007.

[31] V. Vryniotis. Gradient Descent. http://blog.datumbox.com/wp-content/uploads/2013/10/gradient-descent.png/, 2008. [Online; accessed 12-Aug-2018].

[32] D. Whitmer, G. Worrell, M. Stead, I. K. Lee, and S. Makeig. Utility of independent component analysis for interpretation of intracranial eeg. *Frontiers in human neuroscience*, 4:184, 2010.

[33] L. Wiskott. Lecture notes on Independent Component Analysis. https://www.ini.rub.de/PEOPLE/wiskott/Teaching/Material/index.html, 2016. [Online; accessed 07-May-2019].

[34] V. Zarzoso, J. J. Murillo-Fuentes, R. Boloix-Tortosa, and A. K. Nandi. Optimal pairwise fourth-order independent component analysis. *IEEE Transactions on Signal Processing*, 54(8):3049–3063, 2006.

Master's thesis filing card

Student: Marimuthu Ananthavelu

Title: An evaluation of independent component analyses with an application to resting-state Electrocorticography

Dutch title: Een evaluatie van onafhankelijke component-analyses met een applicatie voor rusttoestand Electrocorticography

UDC: 681.3*I20

Abstract:
In this thesis work, a different number of Independent Components Analysis(ICA) algorithms are applied to resting state Electrocorticography (ECoG) data on five different human subjects. At first, the total number of reliable independent components or the so called cortical sources are validated with the help of an already established method in which the randomization of initial parameters in combination with bootstrapping of the samples followed by clustering was used. Once the total number of cortical sources are evaluated, the comparison has been made between FastICA, InfoMax, Jade, OGWE, Pearson, and EFICA algorithms for the quality of the estimated components using the measures of independence which are, the reduction in the mutual information for statistical independence and Spearman correlation coefficient for linear independence between the components. The estimated independent components from each algorithm, are ordered for the importance of non-gaussianity with the help of measures namely, Entropy and Kurtosis. The results are discussed while taking into account the individual functioning of each ICA algorithm and its specific goal of optimization.

Thesis submitted for the degree of Master of Science in Artificial Intelligence, option Engineering and Computer Science

Thesis supervisor: Prof.dr. Ir. Marc Van Hulle

Assessors: Ir. Mansoureh Fahimi Hnazaee
dr. Ir. Ghumare Eshwar

Mentor: Ir. Mansoureh Fahimi Hnazaee