International Journal of Modern Communication Technologies & Research (IJMCTR)
ISSN: 2321-0850, Volume-3, Issue-1, January 2015

Application of Signal Separation Algorithms for Artifact Removal from EEG Signals

Prof. A. Janardhan, Prof. K. Kishan Rao

Abstract— **Contamination of Electroencephalographic (EEG) activity by eye movements, blinks, heart and line noise is serious problem for EEG interpretation and analysis. Often Blind Signal Separation (BSS) algorithm is applied to EEG recordings to remove wide variety of artifacts. Our goal is to quantify which Signal Separation algorithm would be most effective. More specifically, we evaluate the influence of noise and artifacts on the performance of FpICA, SOBI ,JADE and AMUSE algorithms. These algorithms were applied to simulated EEG signal with artificial artifacts.**

Index Terms— **EEG,BSS , artifact, RMSE, Correlation coefficient**

I. Introduction

The Electroencephalogram (EEG)[1, 2] is a medical examination based on brain's electric activity. The signal is recorded using electrodes placed on the scalp of the patient. The signal contains among the useful information, which allow scientists to view the cerebral activity, redundant or noise information, artifacts (extra-cerebral signals). In order to conclude that something is wrong or that the patients have a disease further processing is necessary.

EEG obtained from scalp electrodes is a sum of the large number of neurons potentials of the order of few micro volts . The interest is in studying the potentials in the sources inside the brain and not only the potentials on the scalp, which globally describe the brain activity. Direct measurements from the different centers in the brain require placing electrodes inside the head, which means surgery. This is not acceptable because of the risk for the subject. Another possibility is to calculate the signals of interest from the EEG obtained on the scalp. These signals are weighed sums of the neurons activity, the weights depending on the signal path from the brain cell to the electrodes. Because the same potential is recorded from more than one electrode, the signals from the electrodes are supposed to be highly correlated. If the weights were known, the potentials in the sources could be computed from a sufficient number of electrode signals. Independent component analysis (ICA), sometimes referred to as blind signal separation or blind source separation, is a mathematical tool that can help solving the problem.

In this paper we used Independent Component Analysis (ICA) to remove artifacts from EEG signals. Our goal is to quantify which ICA algorithm would be the most effective and perform well. More specifically, we were interested to compare the performance of FpICA, SOBI, JADE, and AMUSE algorithms. These algorithms were applied to simulated EEG data with added artificial artifacts. The paper is organized as follows: firstly we introduce Methods and Artifact modeling , ICA concept and different algorithms used in this paper. Then in section 3 we explain Comparison approach used , while in section 4 we apply the comparison approach with subsequent section for Conclusions.

II. METHODS

The main approach in extracting an ECG artifact is to apply an ICA algorithm to the EEG signal. Outputs of all algorithms will be mutually independent components and some of them will be estimates of the ECG artifact. After we identify them, the EEG can be reconstructed without these corrupting components. In order to validate effectiveness of the ICA algorithms we need to compare the ICA components with the ECG artifact.

Unfortunately, we do not have much a prior knowledge about the artifact. In this work, we overcome this problem by creating an artificial ECG artifact that was added to an uncorrupted neonatal EEG signal.

A. *Artifact modeling*

After studying properties of ECG artifact signal[12] it was decided to model it as a combination of two signals. Those signals should approximate the main characteristics of ECG artifacts that occur on the EEG channels. The first simulated artifact is a spike train signal simulating corrupting QRS complexes of the ECG. The occurrence of these spikes is not strictly periodical but correlated with heart rate. The other simulated artifact is a sine wave with a frequency of 2 Hz. This sine wave corresponds to the pulsation artifact of an electrode close to a blood vessel. The frequency value is chosen as a frequency that is close to the heart rate of neonates. According to ICA standards both of these signals are created to be zero–mean and standardized to a unit variance. In this way ECG artifact is modeled as a combination of a highly dynamic source (spike train) and a slowly varying time–correlated source (sine).

Manuscript received December 30, 2014.
Prof.A.JANARDHAN, ECE, Jayamukhi Institute of Technological Sciences,Warangal,Telangana State, India, Mobile No 09989814991.
Prof.K.Kishan Rao, ECE, Vagdevi College of Engineering, Warangal, Telangana State, India, Mobile No 09440775866.

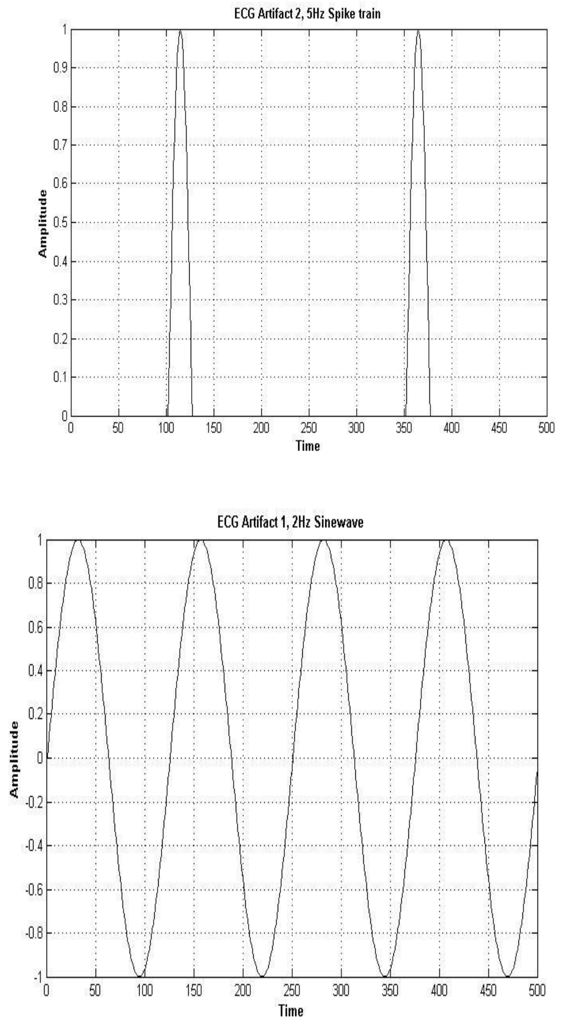

Figure 1. Artificially created ECG artifact SPIKE TRAIN "S1" and SINE "S2" sources

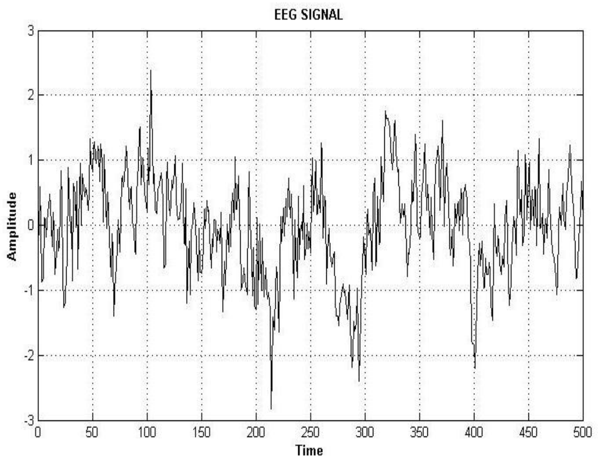

Figure 2. Simulated EEG Signal "S3" without artifacts.

In our simulations we used Figure 2 a clean simulated single trial of 0.8 seconds, EEG signal with sampling frequency of 250 Hz. The spike train artifact source , Sine wave artifact source and EEG signal are synthetically mixed by randomly generated nonsingular mixing matrix as shown in Figure 3.

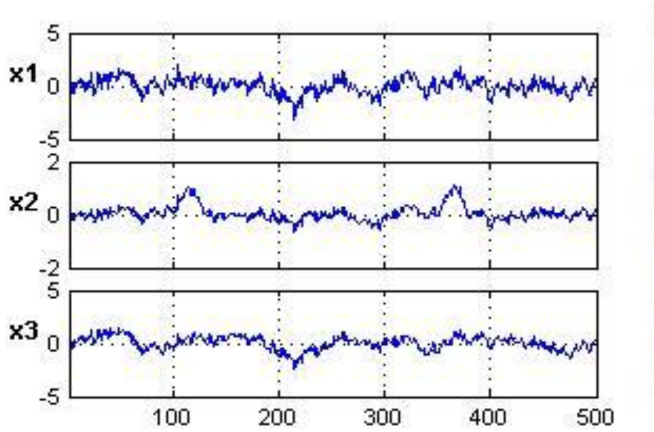

Figure 3.The EEG mixed with artifacts of Sine wave and Spike train sources.

The two independent sources Sine wave and Spike train and EEG were mixed using random mixing matrix H as shown below.

$$\begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} 0.149 & 0.561 & 0.68 \\ 0.324 & 0.002 & 0.14 \\ 0.029 & 0.017 & 1.94 \end{bmatrix} \begin{bmatrix} s1 \\ s2 \\ s3 \end{bmatrix} \qquad \text{.............(1)}$$

B. *Independent Component Analysis (ICA)*

The separation effectiveness of ICA algorithms [3] are related to statistical properties of the mixing signals on which they are applied. Algorithms based on the time structure of the data set have advantages in separating sine shaped signals. In our work we used SOBI and AMUSE as time structure based algorithms. On the other hand, for separation of the spike train signal, it is more suitable to use algorithms which are based on maximizing non–Gaussianity. Therefore, we selected FpICA and JADE. It is good to remark that although the main characteristic of a spike train signal is its "spiky" nature and super–Gaussian distribution, this signal is also time correlated to some extent.

a) ICA Concept.

We consider the linear ICA model with instantaneous mixing. Assume that we observe 'm' linear mixtures $x_1,..., x_m$ of 'n' independent components (sources) $s_1,.....,s_n$. Then we can define the ICA model

$$x(t)=Hs(t) \qquad \text{..(2)}$$

Where the sources $s=[s_1,s_2,....s_n]$ are mutually independent random variables and H is an 'm x n' unknown invertible mixing matrix. The goal is to find only from observations 'x', a matrix W such that the output

$$y(t)= Wx(t) \qquad \text{..(3)}$$

Signal 'y' represents independent components that are actually estimates of sources 's'. There is one limitation in the ICA method in the sense that an estimated signal 'y' cannot determine the variance of a source 's' due to scaling ambiguity. That is, there exists an infinite number of factors 'α'.

$\alpha_n = s_n(t)/y_n(t) = s_n/HWs_n$ ....(4)

Fortunately, we can always choose 'α' in that way that we create a unit variance signal, but this still leaves the ambiguity of the sign.

b) ICA ALGORITHMS

In order to estimate independent components in 'y' the de–mixing matrix W, numerous ICA algorithms have been developed with various approaches. We have used four of them in our comparison and they are overviewed:

(i) Second Order Blind Identification (SOBI)

When the sources are individually correlated in time, but mutually uncorrelated, an ICA algorithm based on second order statistics can be derived [4]. Mathematically, this means that for all time lags 't' the source correlation matrices are diagonal:

$$R_x(\tau)=E\{x(t)x(t+\tau)^{\tau}\} \quad \text{............................(5)}$$
$$= AR_s(\tau)A^T$$

where Rs represents the correlation matrix of the source signals. Considering that this equation holds for all values of 't', the mixing matrix A is the one that jointly diagonalizes all the correlation matrices.

(ii) JADE

Another signal source separation technique is the Joint Approximation Diagonalisation of Eigen matrices (JADE) algorithm [8]. This approach exploits the fourth order moments in order to separate the source signals from mixed signals. At the beginning, the whitening matrix P and the signal z = Px are estimated. After that, the cumulants of the whitened mixtures Q_i^z are computed. An estimate of the unitary matrix R is obtained by maximizing the criteria λiVi by means of the joint digitalization. If λiVi cannot be exactly jointly digitalized, the maximization of the criteria defines a joint approximate digitalization. An orthogonal contrast is optimized by finding the rotation matrix R such that the cumulant matrices are as diagonal as possible:

$$R = \arg\min_R \sum_i \mathrm{Off}\,(R^T Q_i^z R) \quad \text{........................(6)}$$

The mixing matrix A is calculated as $\hat{A} = RP^{-1}$ and the independent components are estimated as $y = \hat{A}^{-1}x = \hat{W} x$.

(iii) FpICA

FpICA algorithm [5] is a fixed-point iteration scheme for finding a maximum of the non-Gaussanity. It uses kurtosis and computations can be performed either in batch mode or in a semi-adaptive manner. It uses deflation approach to update the columns of separating matrix W and to find the independent components one at a time. More recent versions are using hyperbolic tangent, exponential or cubic functions as contrast function.

$$w^*(k)=C^{-1}E\{x\, g(w(k-1)^T x)\}-E\{g'(w(k-1))^T x)w(k-1) \quad \text{...(7)}$$

$$w(k) = w^*(k)w^*(k)^T C\, w^*(k) \quad \text{.............(8)}$$

where g can be any suitable non-quadratic contrast function, with derivative g’; and C is the covariance matrix of the mixtures, x. $w^*(k)^T x(t)$; t = 1, 2,…equals one of the sources.

(iv) AMUSE

Algorithm for Multiple Unknown Source Extraction (AMUSE) [6] is based on EVD of a single time-delayed covariance matrix for pre-whitened data. This algorithm uses the sub-optimal time delay, however it could be set by a value. It is relatively fast but very sensitive to additive sensor noise. This algorithm belongs to Second Order Statistics (SOS) since it uses time delayed covariance matrix.

III. COMPARISON APPROACH

In order to compare different ICA algorithms [8,9] it is necessary to create a function that will match independent components with previously created ECG artifact sources. To accomplish this we had to estimate the correlation of the artifact sources and the independent components. Components with the highest correlation coefficients are selected for comparison as matching components. Independent components are created to be zero–mean and they are standardized to unit variance.

In this experiment we have varied the impact of the ECG artifact by changing non–zero values in mixing matrix "H". In that way, our EEG signal was corrupted by ECG artifact ranging from random sparse to random uniform mixing. The quality of source separation was estimated with three parameters "r", "IS" , "ET" . We have calculated these parameters for all impact levels of the ECG artifact, assuming different mixing values, the performance criteria are averaged over 100 Monte Carlo simulations. Estimated source output plots are shown in figures 4-7 to compare performance of algorithms.

A. Correlation based criterion

The Spearman correlation coefficient 'r'[10] proves to be a good choice to compare the original ECG artifact source and the independent component because it is dependent on the (relative) shape of the signal. It shows normalized values between 0 and 1 regardless of the signal sign. The Spearman correlation coefficient is calculated according to the following formula:

$$r = 1-6 \sum d^2/ N(N^2-1) \quad \text{......................................(9)}$$

where d is the difference in statistical rank of the corresponding signal and N is the signal length. Correlation index "r" is calculated for every ICA algorithm and for both types of ECG artifact. Good separation quality of an ICA algorithm is indicated by a higher value of 'r'.

B. Index of Separability

This index of Separability (IS) [8]is computed from the N x N transfer matrix G=WH between the original sources and the estimated ones. In order to obtain the IS it is necessary to take the absolute value of the elements of "G" and to normalize the rows gi by dividing each element by the maximum absolute value of the row. The rows of the resulting matrix "G' " are:

$$g_i' = |g_i| / \max |g_i| \quad \text{..(10)}$$

The IS is obtained from the new G' matrix:

$$IS = \sum_{j=1}^{N}(\sum_{i=1}^{N}(G'(i,j)-1) \;\; / \; N(N-1) \quad \text{.... (11)}$$

The small value of IS indicates good separation.

C. Scatter Plot

Using the fact that TWO independent signals have a rectangular shape in their own plan is called the scatter plot of the two signals. A scatter plot [9] can suggest various kinds of correlations between variables with a certain confidence interval. Scatter plots display trends strong or weak correlation . One of the most powerful aspects of a scatter plot, however, is its ability to show nonlinear relationships between variables. Scatter plots of sources, and the estimated signals are plotted to show that the separation was done.

IV. RESULTS

In this experiment we randomly added two artifacts to the clean simulated EEG signal. In order to make a fair comparison of BSS algorithms, the same mixing matrix "H" was used for each algorithm. In SOBI algorithm, the number of time delayed Covariance Matrices was set to 100 and in FpICA the nonlinear was set to 'tanh', Fluctuation parameter set to 0.001 and Number of trials set to 10. All other algorithms like JADE and AMUSE values are set to default. The results of Spearman Correlation Coefficient are shown in Table-1 .

Artifact	FpICA	SOBI	JADE	AMUSE
Sine	0.870	0.999	0.722	0.995
Spike	0.685	0.397	0.321	0.399

Table. 1 Spearman Correlation Coefficient.

Another way to estimate the effectiveness of algorithm is to compare Index of Separation and execution time of each algorithm. The results of "IS" and "ET" are shown in Table -2 .

Parameter	FpICA	SOBI	JADE	AMUSE
Index of Separation	0.204	0.164	0.226	0.173
Execution time (ET) algorithm	0.3 Sec	0.47 Sec	0.09 Sec	0.05Sec

Table 2 Index of Separation and Execution times.

In order to compare different Algorithms the estimated signals are plotted as shown.

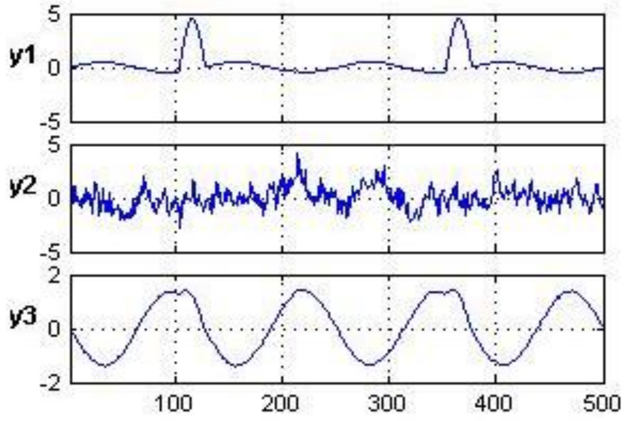

Figure 4.FpICA Estimated outputs

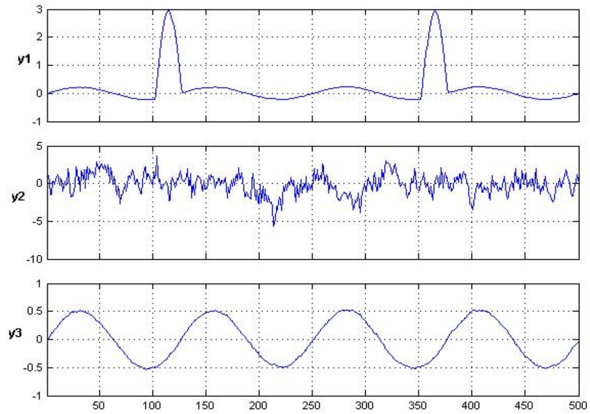

Figure 5. SOBI Estimated outputs

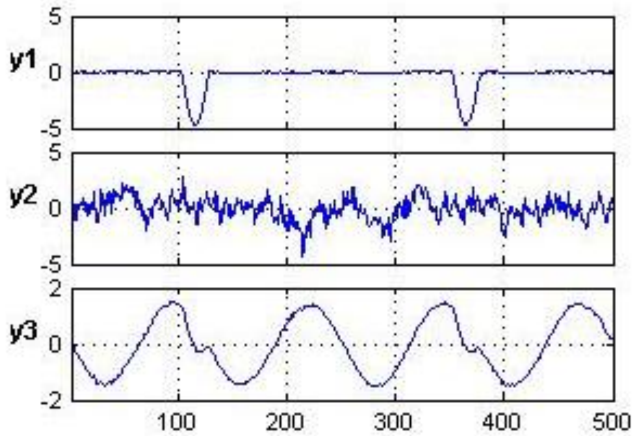

Figure 6. JADE Estimated outputs

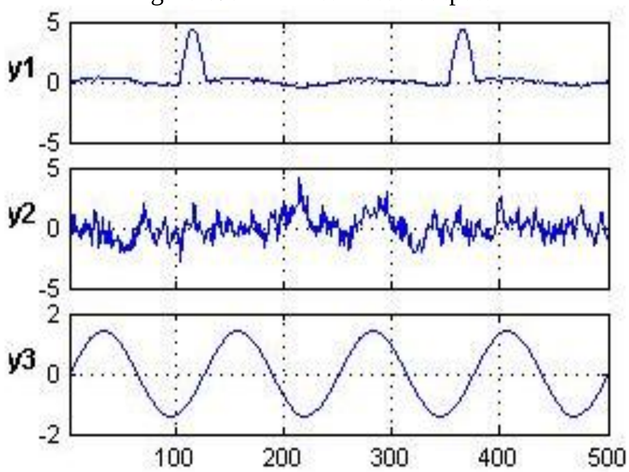

Figure 7.AMUSE Estimated outputs

To show the trends in correlation of data , scatter plots for FpICA, SOBI, JADE and AMUSE algorithms are shown below.

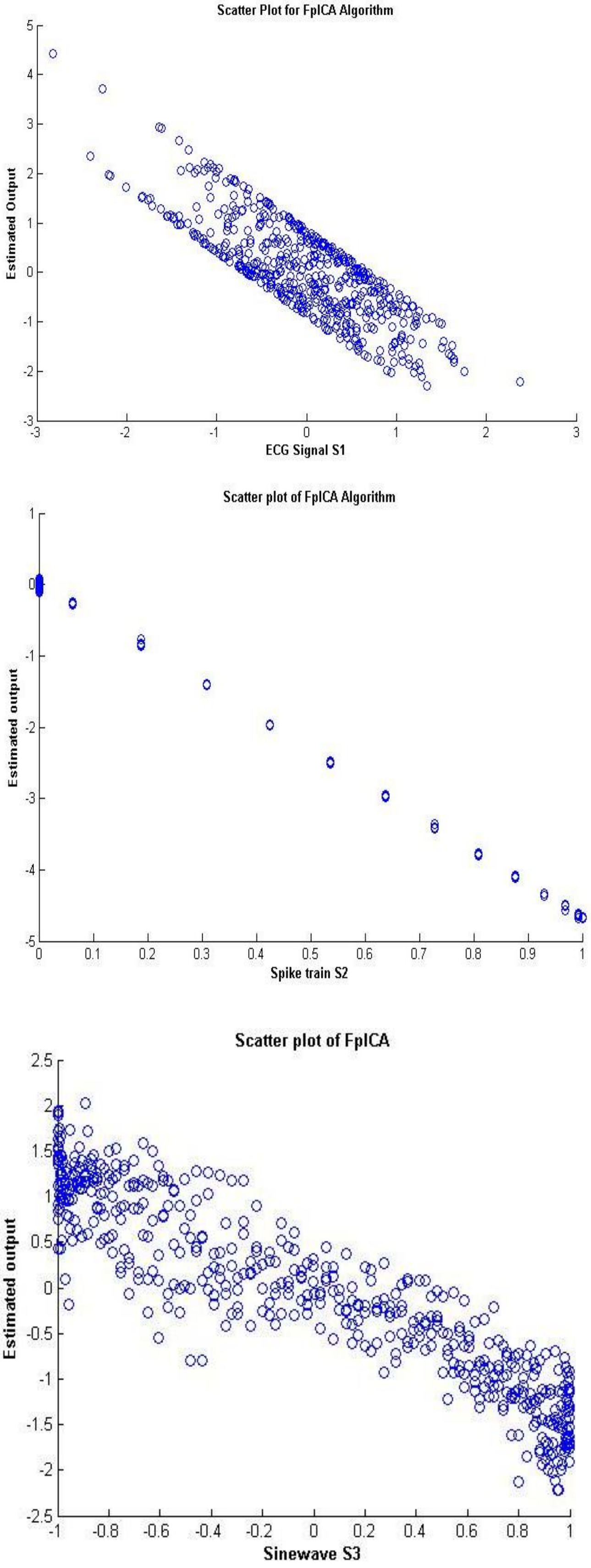

Figure 8. Scatter plot of FpICA Algorithm.

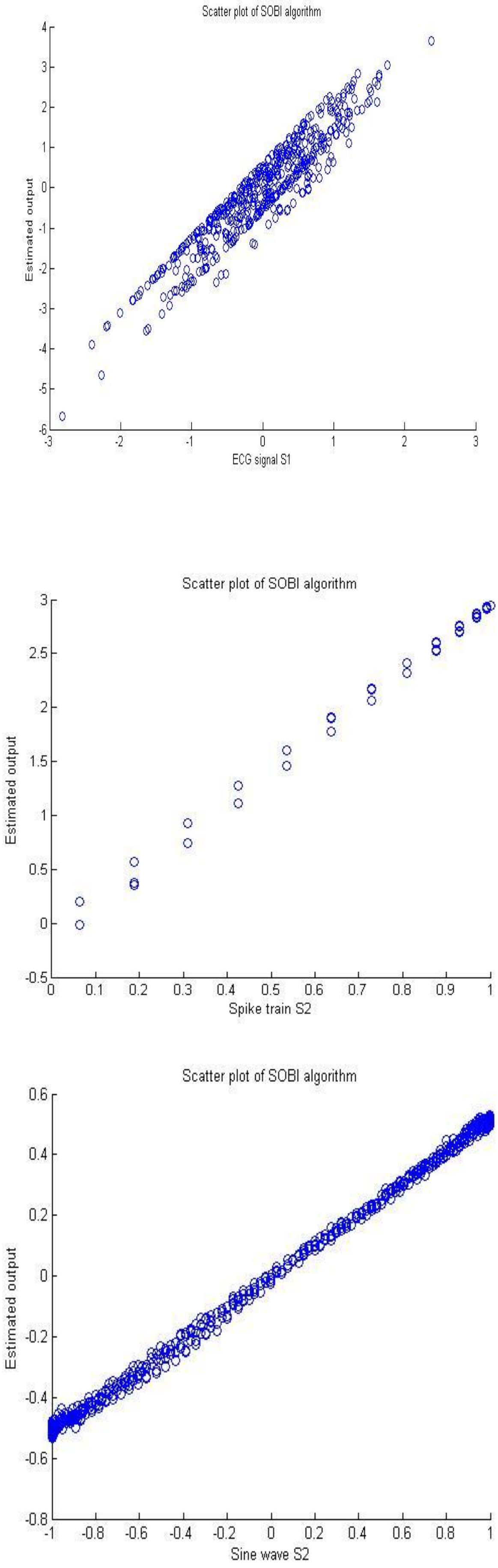

Figure 9. Scatter plot of SOBI Algorithm.

V. CONCLUSION

The experiment results confirm that the time structured based algorithms SOBI and AMUSE are efficient in extracting Sine wave. In separating the Spike train signal, the algorithms based on maximizing non-gaussianity FpICA and JADE little out performed time structure based SOBI and AMUSE. Similarly in-terms of execution time AMUSE and JADE out performed .Generally when all parameters correlation coefficient and Index of separation are taken in to account, we may conclude that FpICA for Spike train extraction and SOBI for Sine wave extraction should be applied. As both artifacts extraction is concerned which is required in here, SOBI should be the best option.

REFERENCES:

[1] Croft RJ Barry (2000) "Removal of ocular artifact from the EEG a review" Clinical Neurophysiology, 30(I), pp- 5-19.

[2] V.Krishanaveni, S.Jayaraman and K Ramadoss "Application of Mutual Information based least dependent Component Analysis (MILCA) for removal of Ocular Artifacts from Electroencephalogram.

[3] Aapo Hyvarinen and Erkki Oja" Independent Component Analysis : Algorithms and Applications" Neural network Research Center, Helsinki University of Technology, Neural Networks,13(4-5):411-430, 2000

[4] A.Beloucharani K.Abed-Meraim, J F Cordoso and E Moulines, "Second Order Blind Separation of temporally correlated Source", Proc. Int. Conf. on Digital Signal Proc. (Cyprus),pp 346-351, 1993.

[5] A.Hyvarian and E. Oja "A Fast Fixed-Point Algorithm for Independent Component Analysis", Neural Computation, 9(7):1483-1492.1997

[6] L.Tong, V.Soon, Y F Huang and R.Liu, " Indeterminacy and Identifiability of blind identification, IEEE Trans, CAS, vole , 38,pp-499-509, March 1991

[7] A Hyvarian, J Karthunen and E. Oja, "Independent Componenet Analysis". John Wiley & Sons, 2001.

[8] Ricardo Salido-Rulz,Rebeca Romo-Vazquez,Radu Ranta,Lorenzo Leija "Analysis of 5 Source Separation Algorithms on Simulated Signals", CNRS , Mexico.

[9] Ali Mansour,Mitusuru Kawamoto and Noboru Ohninshi " A Survey of The Performance Indexes of ICA algorithm"Proc. IASTED International Conf, Modelling, Identification, and Control" Feb 18-21,Innsbruk , Austria.

[10] M.Klemm, J.Haueisen and G.Ivanova "Independent Component Analysis: comparison of algorithms for the investigation of surface electrical brain activity". Med Biol Eng Comput, Vol 47, pp 413-423 , 2009

[11] A.Cichocki, S.Amari, K. Siwek , T.Tanaka et al., ICALAB Toolboxes, http://www.bso.brain.riken.jp/ICALAB.

[12] A Garcés Correa , E Laciar , H D Patiño , M E Valentinuzzi " Artifact removal from EEG signals using adaptive filters in cascade" 16th Argentine Bioengineering Congress and the 5th Conference of Clinical Engineering-2007, Journal of Physics: Conference Series 90 (2007) 012081.

AUTHORS PROFILE

PROF A. Janardhan obtained M Tech degree in Advanced electronics from Jawaharlal Nehru Technological University, Hyderabad, India in 1986 . Worked at ISRO Bangalore, DRDL Hyderabad, CMC R&D Centre and Portal Player India Pvt Ltd, Hyderabad. Started teaching career by joining Sri Venkateswara Engineering College as Professor ECE in 2009 and subsequently joined as Professor ECE at JITS Warangal, Telangana State. Currently doing PhD in electronics and communication engineering (Digital Signal Processing) at Jawaharlal Nehru Technological University ,Hyderabad , India. Research interest include Machine Learning, Source separation Algorithms ,Signal processing in Geophysics and Communication networks.

PROF K. Kishan Rao obtained B.E., & M.E., Degrees from Osmania University ,Hyderabad, India in 1965 and 1967 respectively. Obtained Ph.D from IIT Kanpur in 1973. Joined REC(NIT) Warangal on 1st September , 1972 as Lecturer in ECE Department. Promoted as Assistant Professor in May ,1974 and as Professor in March.1979. Held Positions of Head of Department, Chief Warden, Dean (Academic Affairs) and retired as PRINCIPAL on 22nd March,2002.Worked on 5 Research Projects of MHRD., DOE & DST. Conducted 8 FDP programs. Delivered more than 60 Expert Lectures. Published 51 Research papers in National Journals and 40 Research Papers in International Journals. Guided 3 candidates for their Ph.D. Degree . At present 8 candidates are working for their Ph.D. Degree in fields of Wireless Communications and Digital Signal Processing.