International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 3 (2015) pp. 7147-7164

http://www.ripublication.com

Adaptive Estimation of Blind Audio Correlation Signal Characteristics using Neural Networks and Complex Pursuit Gradient Method

[1]Chandika Mohan Babu and [2]AzemeeraSharath Naik

IEEE Member
Assistant Professor, Wireless Communication Domain (D3)
School of Electronic and Communication Engineering,
Lovely Professional University,Punjab, India.
chandika.14736@lpu.co.in
Assistant Professor, Wireless Communication Domain (D3)
School of Electronic and Communication Engineering,
Lovely Professional University,Punjab, India.
Sharath.17394@lpu.c

Abstract

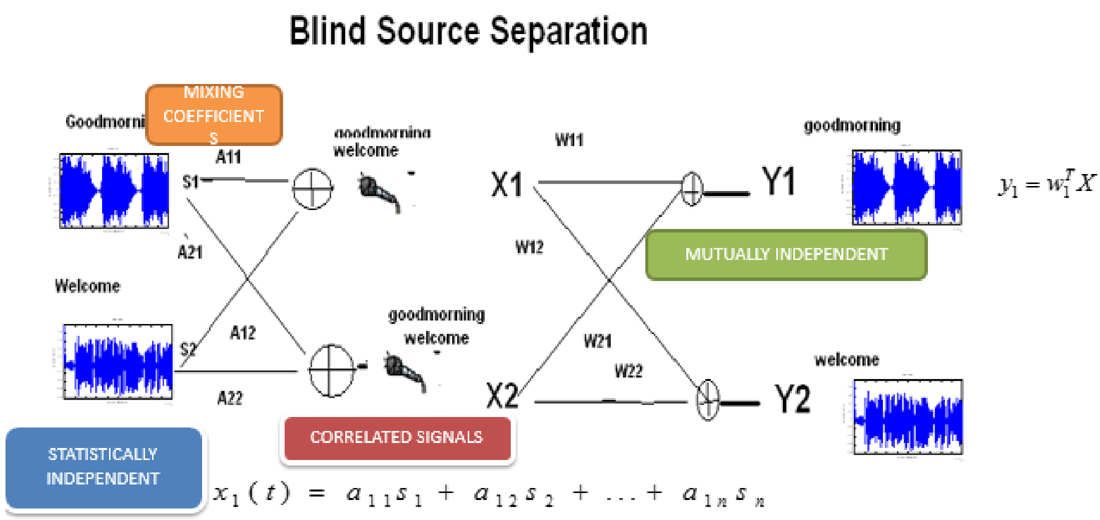
In this research paper we simulate the simulations of the real time application in the field of speech processing which overcome the effect of interference in blind source separation (BSS) using blind equalization principles with neural networks. we have done the separation of three speech signals which are taken as wav file and the work is explored for different signals by two different methods called projection pursuit method and complex pursuit gradient ascent method considering the non gaussianity and correlation in time series. We estimated the original signal by un-mixing the coefficients (W) in such way that the un-mixing coefficients extract the original signal from the mixture on the signal and the simulations are focused on correlation characteristics of signals and results were performed in Matlab simulation tool.Complex pursuit is measured interms of predictability (F), kurtosis (K) and gradient ascent which gives better result in maximizing and minimizing the complexity of the signal mixtures using the fourth order kurtosis function. The complexity of the signal mixture reduces in such a way that the un-mixing coefficients extract a non-Gaussian signal from the signal mixture and the variation can be seen by finding correlation of original signal and extracted signal which is almost similar and the process is carried in adaptive manner using neural network.Simulation results shows the correlation between the original signal and extracted signal is almost equal to 98-99 variation, as the iteration process is increased the correlation value of the corresponding signal is increasing to almost 99.89% to

100%. This shows the complexity of the signal mixture is reduced by each iteration and estimating the original signal back from a mixture signal is done.

Keywords— BSS,Non-Gaussian,ComplexPursuit Gradient Ascent,Fourth Order Kurtosis Function.

I. INTRODUCTION

Blind source separation (BSS) is the process of getting back the original signal from a mixture of different signals from a given microphone. As this signals are any kind of format say speech , image, electrical signal , or any source format (S), if this signals are unknown ,which are linearly mixed with the mixing coefficients (A) then the received output is said to be an a blind signal as we don't know what kind of signal we are receiving from the microphone. So the real challenging problems comes here is extracting the original signal back from the microphone. [1] [6]. We are estimating the original signal by un-mixing the coefficients (w) in such way that the un-mixing coefficients extract the original signal from the mixture on the signal properties [1] [3]. As the original source signal and mixing coefficients are not known, it is difficult to find the task.



$$x_1(t) = a_{11}s_1 + a_{12}s_2 + \ldots + a_{1n}s_n$$
$$x_2(t) = a_{21}s_1 + a_{22}s_2 + \ldots + a_{2n}s_n$$
$$\vdots$$
$$x_m(t) = a_{m1}s_1 + a_{m2}s_2 + \ldots + a_{mn}s_n$$

Fig 1.Blind Source Separation

Consider a situation where there are n people speaking simultaneously in a conference in m number of microphones respectively, then the each microphone is a mixture of n speakers which are linearly mixed with the distance from the microphone. The output from the each microphone is combination of different speaker with varying distance parameters and it is in the form of $x_1(t)x_2(t).....x_m(t)$.

$$
\begin{aligned}
x_1(t) &= a_{11}s_1 + a_{12}s_2 + ... + a_{1n}s_n \\
x_2(t) &= a_{21}s_1 + a_{22}s_2 + ... + a_{2n}s_n \\
&: \\
x_m(t) &= a_{m1}s_1 + a_{m2}s_2 + ... + a_{mn}s_n
\end{aligned}
\quad \textbf{(1)}
$$

Where a_{ij} and original speech signals $s_1(t), s_2(t).....s_m(t)$, are unknown and given the recorded signals $x_1(t), x_2(t).....x_m(t)$. This problem can be solved if some useful information of the signal is known to estimate the a_{ii}, and this information can be signal structure, statistically independent of each other in time series. Several papers has presented [6] [7], ICA is a classic example of blind source separation which depends on signals are statistically independent of each other at any time interval.Independent component analysis separates mixture signal into useful information components where the nature of source signal is unknown. i.e. if we consider voice signal at particular time it is always different compare to other voice signal at that particular instant of time because they are uncorrelated[6] [7] [9] , so ICA separates a set of signal mixtures into a corresponding set of statistically independent components as shown in Fig 2.

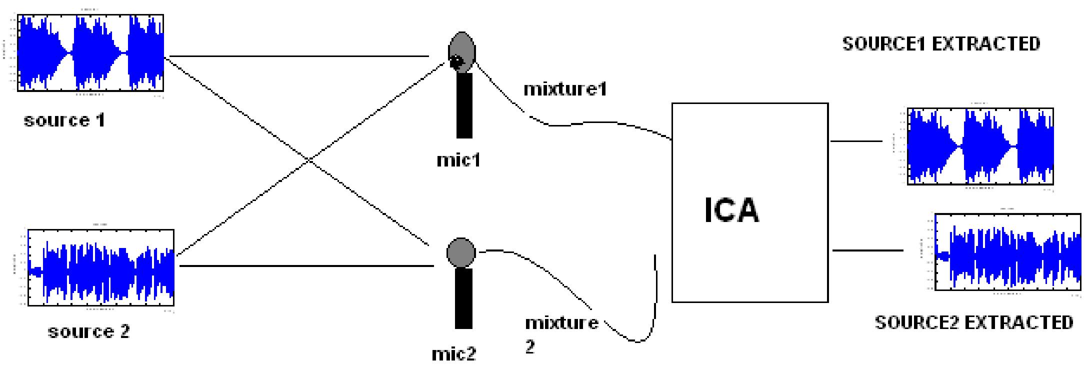


Fig 2. ICA Separation

II. COMPLEXITY PURSUIT GRADIENT ASCENT

1) ***Extracting a Single Signal***

Consider a scalar signal mixtures y_i formed by the applicationof a weight vector w_i to a set of M signals $X = (x_1, x_2,,,,,,,,,, x_M)^T$. Given that $y_i = w_i^T X$ and can be written as

$$F = \ln \frac{w_i \bar{c} w_i^T}{w_i \hat{c} w_i^T} \quad \textbf{(2)}$$

Where $\bar{C}$ is an $M \times M$ matrix of long-term covariance between signal mixtures, and $\hat{C}$ is a corresponding matrix of short-term covariance. The long term covariance $\bar{C}_{ij}$ and the short-term covariance $\hat{C}_{ij}$ between the ith and jth mixtures are defined as

$$\hat{C}_{ij} = \sum_{\tau}^{n} (x_{i\tau} - \hat{x}_{i\tau})(x_{j\tau} - \hat{x}_{j\tau})$$
$$\bar{C}_{ij} = \sum_{\tau}^{n} (x_{i\tau} - \bar{x}_{i\tau})(x_{j\tau} - \bar{x}_{j\tau}) \quad \textbf{(3)}$$

Note that $\hat{C}$ and $\bar{C}$ need only be computed once for a given set of signal mixtures, and the terms $(x_{i\tau} - \bar{x}_{i\tau})$ and $(x_{i\tau} - \hat{x}_{i\tau})$, can be precomputed using fast filtering operations.

Gradient ascent on F with respect to w_i could be used to maximize F, thereby maximizing the predictability of y_i. The derivative of F with respect to w_i is

$$\nabla w_i F = \frac{2w_i}{V_i}\bar{C} - \frac{2w_i}{U_i}\hat{C} \quad \textbf{(4)}$$

The function F can be maximized using gradient ascent to iteratively update w_i until a maximum of F is located

$$w_i = w_i + \eta \nabla w_i F \quad \textbf{(5)}$$

Where η is a small constant (typically 0.001) and this is how we extract single source.

2) Simultaneous Extraction of Signals

The gradient of F is zero at a solution,

$$w_i \bar{C} = \frac{V_i}{U_i} w_i \hat{C} \quad \textbf{(6)}$$

F corresponds to values of w_i that satisfy equation, which has the form of a generalized eigen problem. Solutions for w_i can therefore be obtained as eigenvector of the matrix $(\hat{C}^{-1}\bar{C})$, with corresponding eigen values $\gamma_i = \frac{V_i}{U_i}$. The first such eigenvectors defines a maximum in F, and each of the remaining eigenvectors defines saddle points in F.

The matrix $W = (w_1, w_2,,,,,, w_M)^T$ can be obtained using a generalized eigen values routine. Eigen value function $W^T = eig(\bar{C}, \hat{C})$ can be used from Matlab. All M signals can be recovered,

$$y = WX \quad \textbf{(7)}$$

Where each row of y is one extracted signal y_i.

If the number M of mixtures is greater than the number of source signals then a standard procedure for reducing M consists of using principal component analysis

(PCA). PCA is used to reduce the dimensionality of the signal mixtures by discarding eigenvectors of X which have eigen values close to zero.

III. PROPOSED BLOCK DIAGRAM

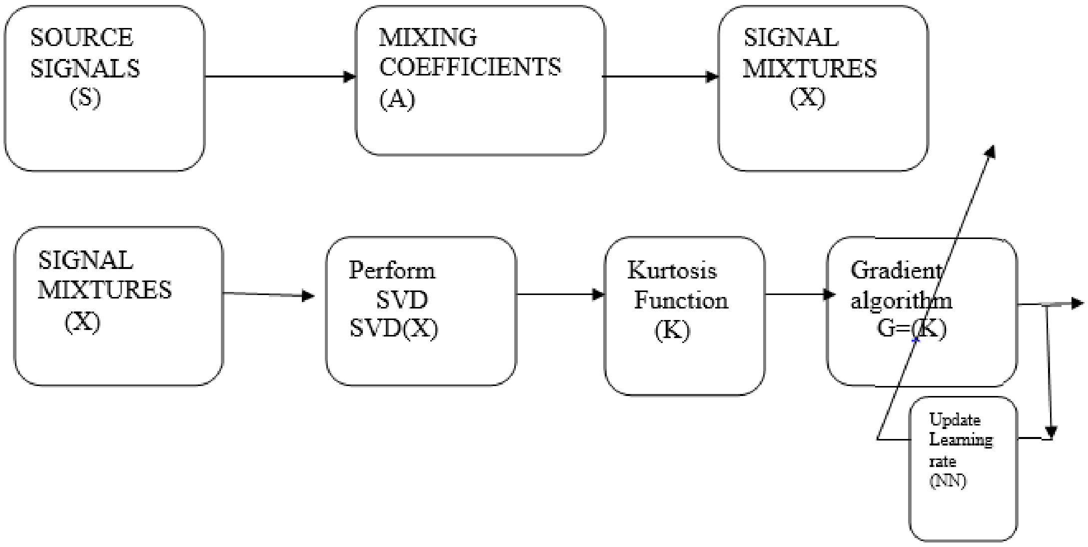
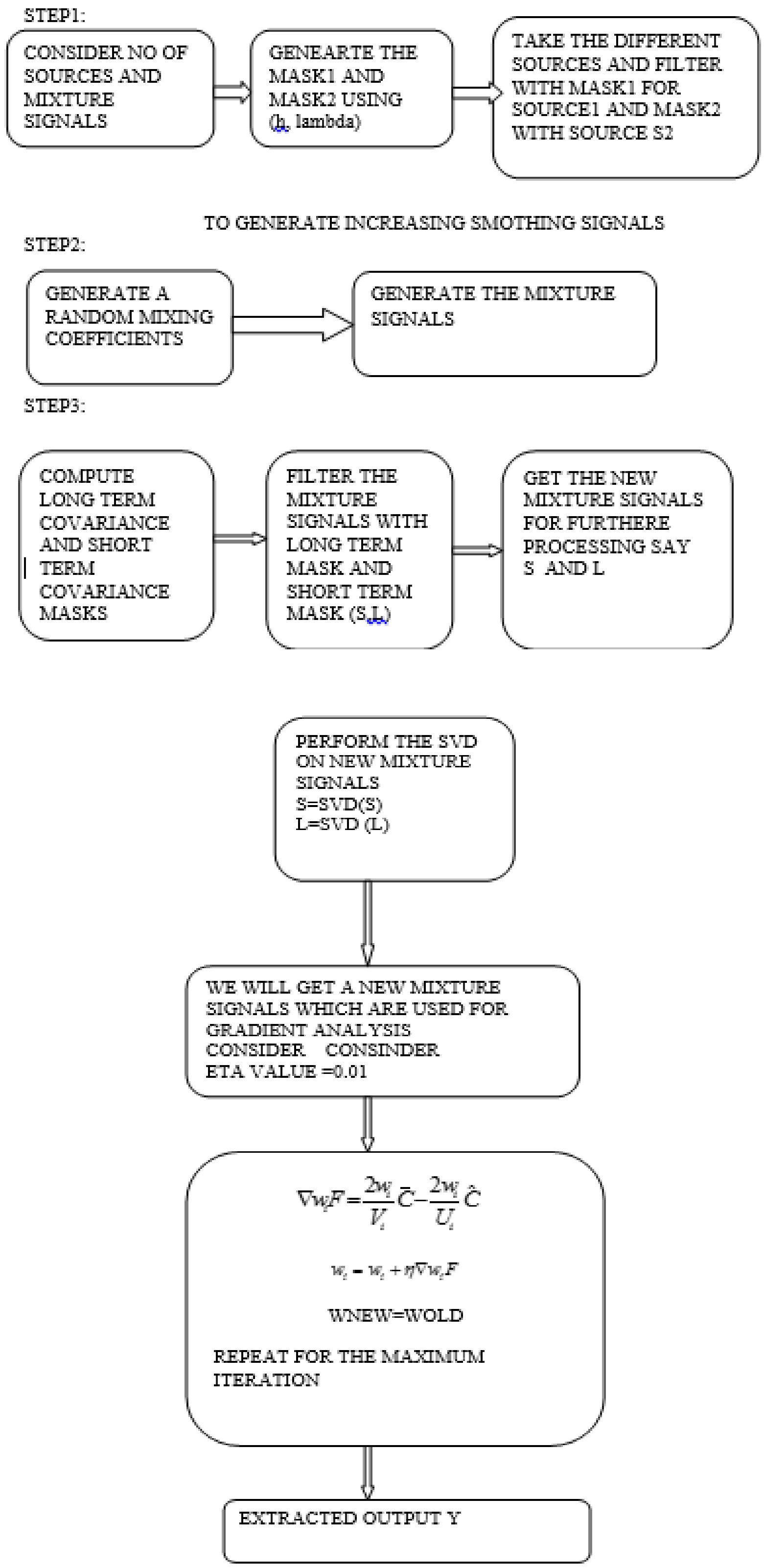


Fig. 3 Proposed Method block diagram

IV. ALGORTITHM IMPLEMENTATION

The simulations are performed using the Matlab Tool considering the source signal as wav files like audio file of police car, reading a IEEE paper e.t.c which are recorded and used for the problem definition of BSS. The algorithm implemented in 4 steps using mathematical computations.


STEP1:
CONSIDER NO OF SOURCES AND MIXTURE SIGNALS
GENEARTE THE MASK1 AND MASK2 USING (h, lambda)
TAKE THE DIFFERENT SOURCES AND FILTER WITH MASK1 FOR SOURCE1 AND MASK2 WITH SOURCE S2
TO GENERATE INCREASING SMOTHING SIGNALS
STEP2:
GENERATE A RANDOM MIXING COEFFICIENTS
GENERATE THE MIXTURE SIGNALS
STEP3:
COMPUTE LONG TERM COVARIANCE AND SHORT TERM COVARIANCE MASKS
FILTER THE MIXTURE SIGNALS WITH LONG TERM MASK AND SHORT TERM MASK (S,L)
GET THE NEW MIXTURE SIGNALS FOR FURTHERE PROCESSING SAY S AND L
PERFORM THE SVD ON NEW MIXTURE SIGNALS S=SVD(S) L=SVD (L)
WE WILL GET A NEW MIXTURE SIGNALS WHICH ARE USED FOR GRADIENT ANALYSIS CONSIDER CONSINDER ETA VALUE =0.01
$\nabla w_i F = \frac{2w_i}{V_i}\bar{C} - \frac{2w_i}{U_i}\hat{C}$
$w_i = w_i + \eta \nabla w_i F$
WNEW=WOLD
REPEAT FOR THE MAXIMUM ITERATION
EXTRACTED OUTPUT Y

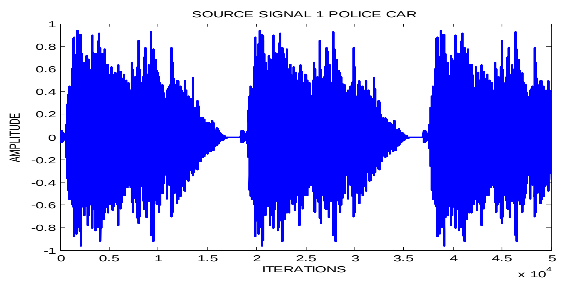
V. SIMULATIONS AND RESULTS

1) Complexity pursuit for different sources

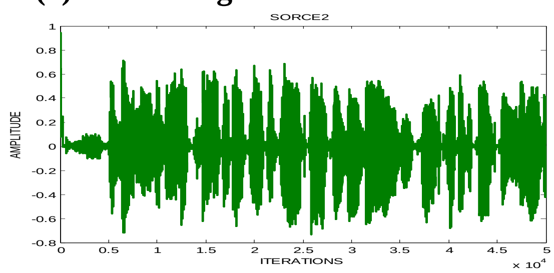
Speech source signals are statistically independent, Signal mixtures are mixture of two statistically independent signals which makes more complex which are plotted as a histogram whichyields a more bell-shaped structure[4] [10] [11], These bell-shaped histograms are referred to as normal or Gaussian. σ is the standard deviation, which is a measure of the variability of x and $\bar{x}$ is the mean of x .the first term $(1/\sqrt{2\pi\sigma^2})$ is a normalization constant which ensures that the area under the normal pdf sums to unity, derivative $dg(x)/dx$ of $g(x)$ with respect to x. [7] [3]

$$p_x(x) = \frac{dg(x)}{dx} \tag{8}$$

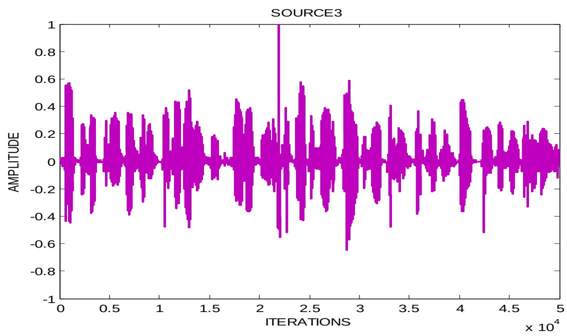
$$p_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right) \tag{9}$$



(a) Source Signal 1:Police Car Sound



(b) Source Signal 2: Reading the Stock Market



(c) Source Signal 3:Reading IEEE Paper

Fig 4. Three different sources

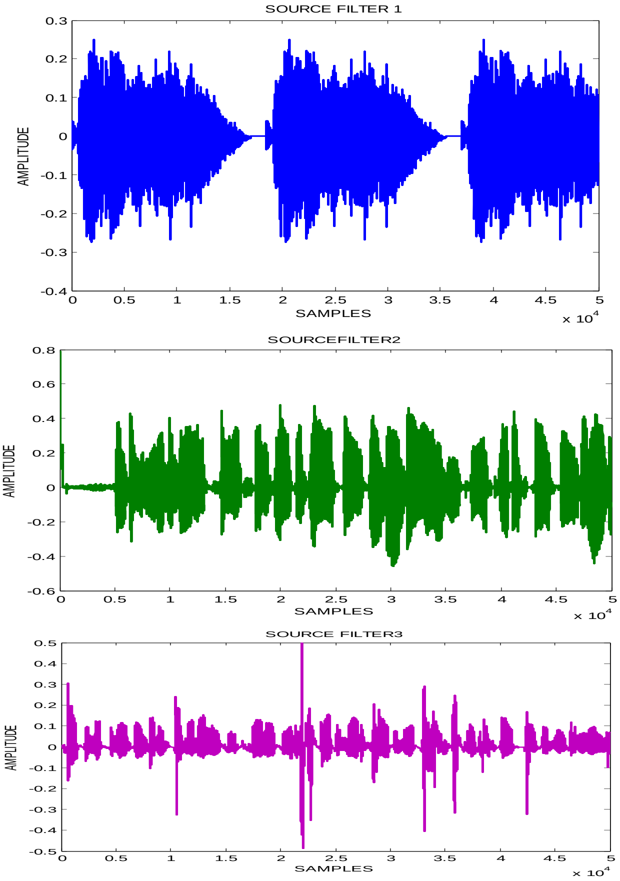


Fig 5. Filter Output after Masking(Step 1)

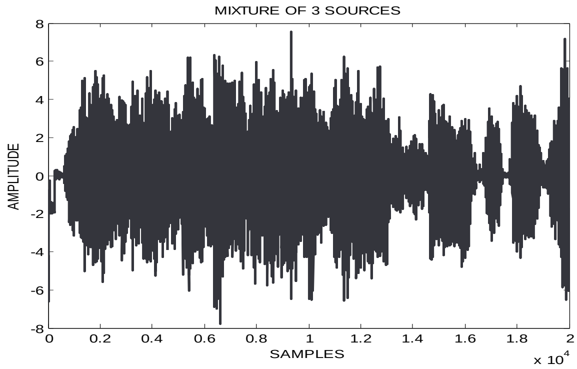
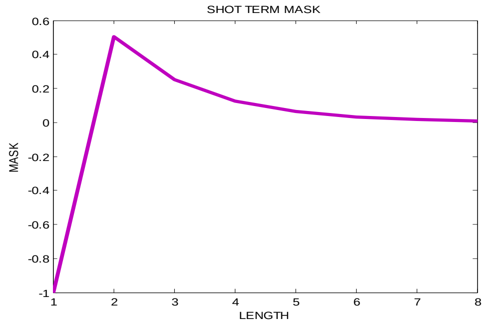
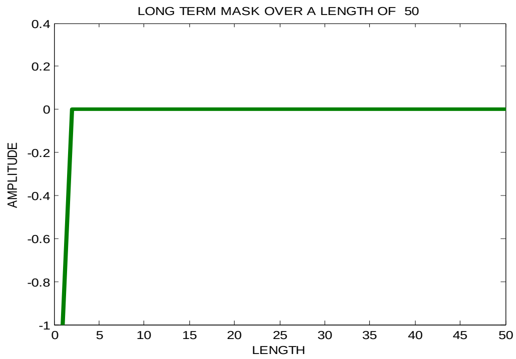


Fig 6. Mixture of Three Source Signals from Microphone (Step 2)

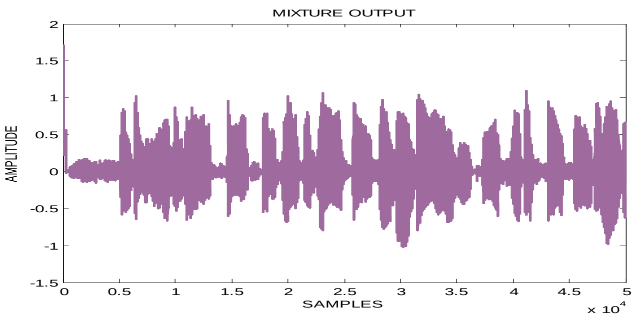


(a) Short Term Mask

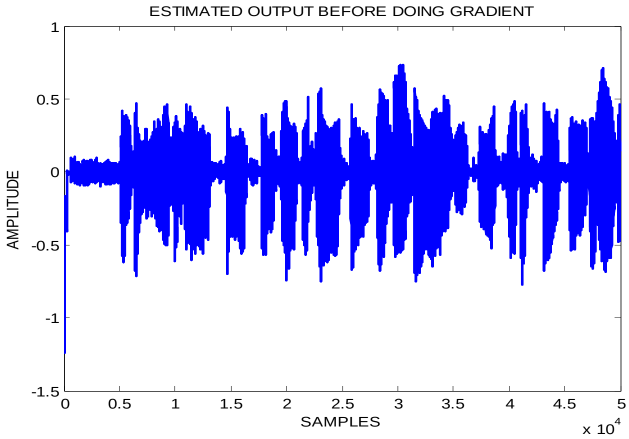


(b)Long Term Mask

Fig 7. Long Term and Shot Term Covariance for Length 50

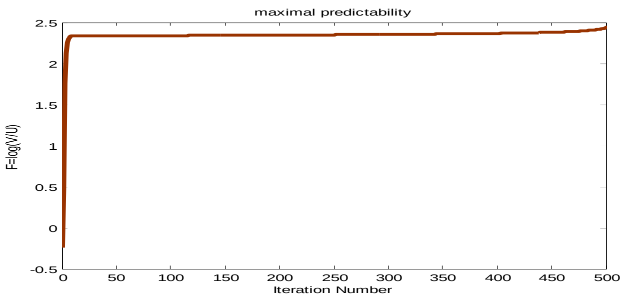


(a) Mixture Output

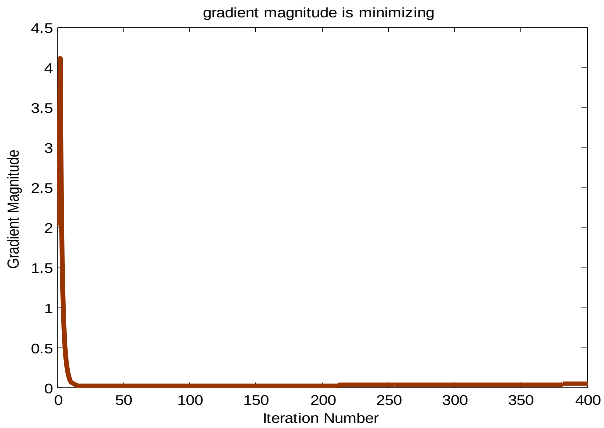


(b) Estimated Output

Fig 8. (a)Mixture output (b) estimated output



(a) Maximum Predictability using F=log(V/U)



(b) Gradient value is minimized

Fig 9. F is maximizing and Gradient values is minimizing with respect to W

2) Performing Complex Pursuit Gradient Ascent

Function F is maximized and the value is in log which is equal to 2.4295, Gradient G is -0.18218 -0.060625 0.047073 and New Mixing Matrix is W -0.073662 6.3688 8.0002

Using Grad ascent:
Correlation of signal with sources extracted by initial w
ans = 0.0029 0.0001 1.0000

Using EIG: Correlations between sources and all recovered signals

1.000	0.0033	0.0029
0.0007	1.000	0.0001
0.0003	0.0001	1.0000

This Implies we are Extracting the 3rd Signal which is in pink colour, i.e. Reading IEEE Paper

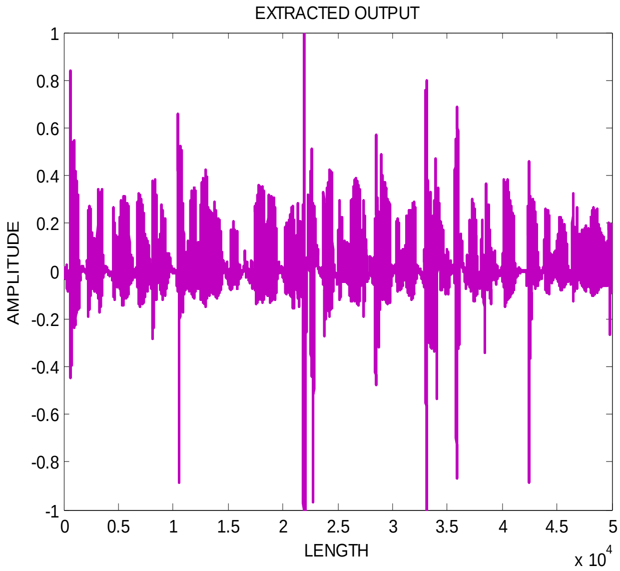
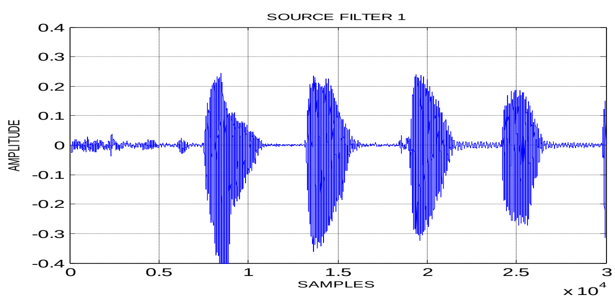


Fig 10. Extracted output

3) Complexity Pursuit for Two Different Sources with 5000/500 Iterations



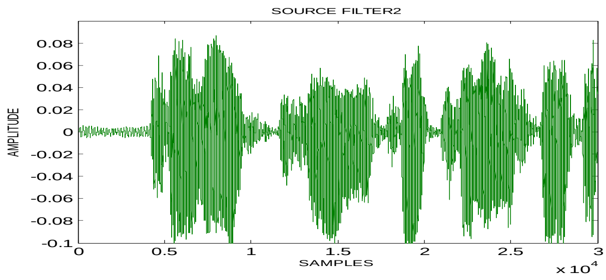


Fig 11. Filter outputs of sources

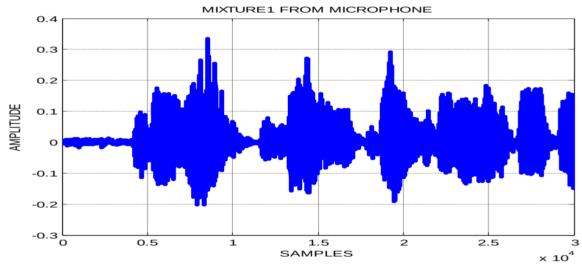
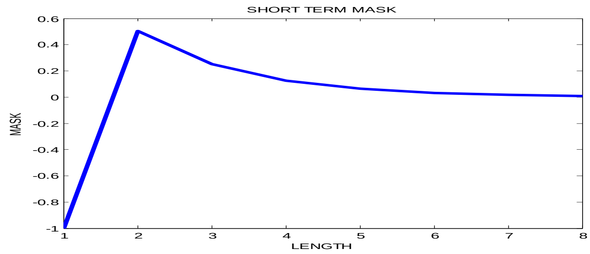
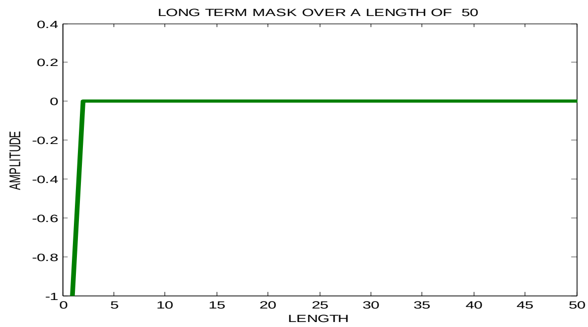


Fig 12. Mixture output from the microphone



(a) Short Term Mask



(b) Long Term Mask

Fig 13. Long Term and Short Term Covariance

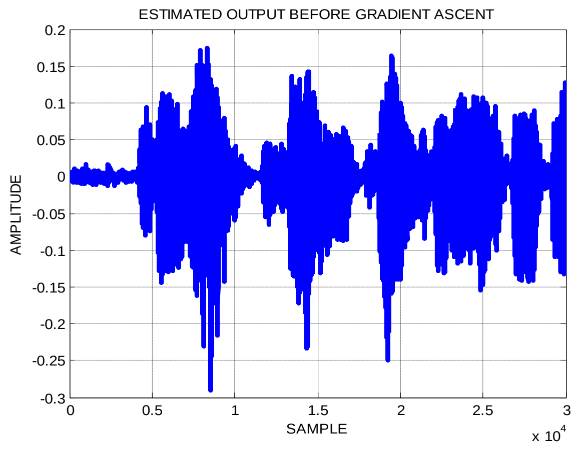


Fig 14. Expected output before Gradient Ascent

4). Implementing Complex Pursuit Gradient Ascent for 5000 Iteration

Function F is maximized and the value is in log which is equal to 1.6115,
Gradient G is 0.00036523 -0.00010216
And new Mixing matrix is W -0.34246 -1.2245

Using Grad ascent: (5000 ITERATIONS)

Correlation of signal with sources extracted by initial w
ans = 0.0021 0.0585 **0.9985**
Implies 99.85 % of source signal is extracted

Using Grad ascent (500 ITERATIONS)

Correlation of signal with sources extracted by initial w
ans = 0.0022 0.1269 0.9967
Implies 99.67 % of source signal is extracted

Extracting one source signal from a mixture of three original signals, in which third signal is recovered by estimating the gradient neural network process

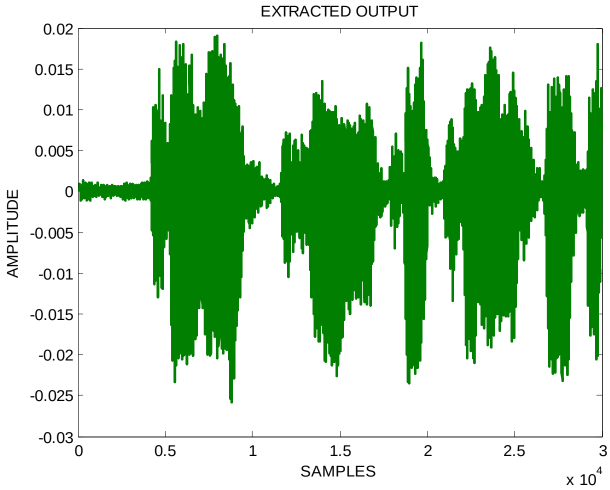
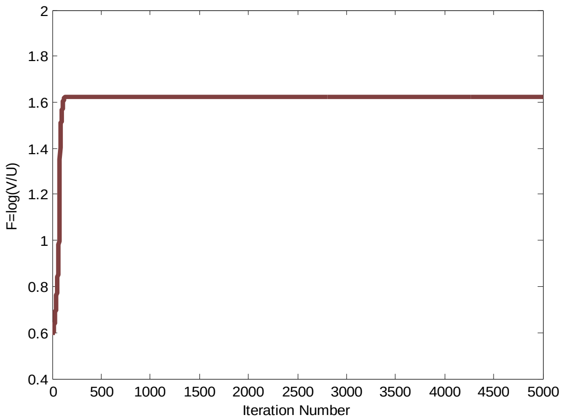
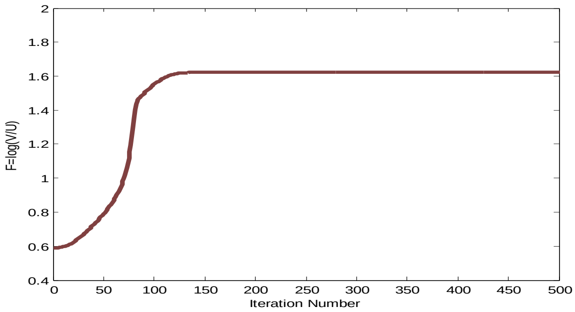


Fig 15. Estimated output and extracted output

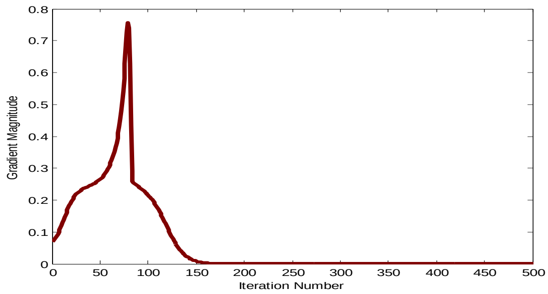


(a)500 Iterations

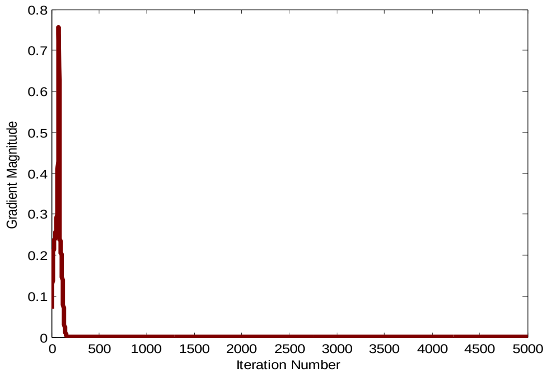


(b) 5000 Iterations

Fig 16. Gradient Function F For (a) 500 (b) 5000 Iterations



(a) 500 Iterations



(b) 5000 Iterations

Fig 17. Gradient Function For (a) 500 (b) 5000 Iterations

5) Proposed Analysis of Complex PursuitGradient Method

As in this method we are doing the long term (V) and short term (U) covariance for maximizing the temporal function which is in the form of F=ln(V/U) , this will vary with respect to weight vector w. This function is mainly depending on the value of lambda which inturn depend on the h function, so there is no relation with respect to eta values this can be explained by considering number of iterations equal to 1000 as shown in table 1.When the lambda value is 1 the corresponding correlation output is maximum at 99.85% and function F which is called as the predictability in terms of log.

So our assumptions consider the number of iterations for 5000 and keeping the eta value which is used for convergence as constant value say 0.1, so if we see the variation of parameters like H, lambda, correlation coefficient and function F.

TABLE 1.Complex Pursuit Gradient Method(ETA, F, Corr)

ETA	F VALUE	H	lambda	CORRELATION
0.1	1.8115	9e+005	1	99.85%
0.2	1.8115	9e+005	1	99.85%
0.3	1.8115	9e+005	1	99.85%
0.4	1.8115	9e+005	1	99.85%
0.5	1.8115	9e+005	1	99.85%
0.6	1.8115	9e+005	1	99.85%
0.7	1.8115	9e+005	1	99.85%
0.01	1.8115	9e+005	1	99.85%
0.02	1.8115	9e+005	1	99.85%
0.03	1.8115	9e+005	1	99.85%
0.04	1.8115	9e+005	1	99.85%
0.05	1.8115	9e+005	1	99.85%
0.06	1.8115	9e+005	1	99.85%
0.07	1.8115	9e+005	1	99.85%
0.001	1.8115	9e+005	1	99.85%
0.002	1.8115	9e+005	1	99.85%
0.003	1.8115	9e+005	1	99.85%

Similarly the table 2 signifies that the lambda value is making an important role in extracting the source signal , colour with bold text signifies , they are extracted output , whereas the other rows specify they are the mixture of two signals. And the function F= zero corresponds that the weight vector is not orthogonal to the mixture signal and when the lambda value is tending to.99 which is less than 1 gives the information about the extracted output.

So we conclude that to extract non gaussianity and correlation signal with effective predictability the proposed complex pursuit gradient method depends on the three parameters (H, lambda, F). The results provide the effective separation for the data base signal and this approach will give same results with other signals in real time applications using Matlab tool.

TABLE 2.Complex Pursuit Gradient Method (H, lambda, F)

ETA	H	Lambda	F	CORRELATION
0.1	0.5	0.25	0	70.68 % EXTRACTED OUTPUT IS MIXTURE SIGNAL
0.1	2	0.70711	0	71.85% EXTRACTED OUTPUT IS MIXTURE SIGNAL
0.1	4	0.8409	0	89.47%
0.1	10	0.93303	0	96.59%
0.1	20	0.96594	0	97.00%
0.1	**100**	**0.99309**	**0.00096261**	**99.65%**
0.1	**200**	**0.99654**	**0.00096259**	**99.85%**
0.1	**300**	**0.99769**	**0.0014717**	**99.85%**
0.1	**400**	**0.99827**	**0.0017896**	**100%**
0.1	**500**	**0.99875**	**0.0020045**	**100%**
0.1	**600**	**0.99885**	**0.0021685**	**100%**
0.1	**700**	**0.99901**	**0.0022929**	**100%**
0.1	**800**	**0.99913**	**0.002463**	**100%**
0.1	**900**	**0.99923**	**0.002463**	**100%**
0.1	**1000**	**0.99931**	**0.0025277**	**100%**
0.1	**1100**	**0.99954**	**0.002581**	**100%**

V1. CONCLUSION

We estimated the original signal by un-mixing the coefficients (W) in such way that the un-mixing coefficients extract the original signal from the mixture on the signal and the simulations are focused on correlation characteristics of signals and results are performed in Matlab simulation tool.Complex pursuit is measured interms of predictability (F), kurtosis (K) and gradient ascent, which gives better result in maximizing and minimizing the complexity of the signal mixtures using the fourth order kurtosis function the complexity of the signal mixture that reduced in such a way that the un-mixing coefficients extract a non-Gaussian signal from the signal mixture and the process is carried in adaptive manner using neural network. Simulation results shows the Correlation between the original signal and extracted signal is almost equal to 98-99 variation, as the iteration process is increased the correlation value of the corresponding signal is increasing to almost 99.89% to 100%. This shows the complexity of the signal mixture is reducing by each iteration and estimating the original signal back from a mixture signal is done with parameters. As the eta value is varied and keeping the other factors unchanged the observation is F is reached a maximum point 1.8115 and H and lambda which are dependent on each other remain same and the correlation between the original signal and the extracted signal is 99.85% , Lambda=2^(-1/H). So there is no effect of eta value in complex pursuit gradient method. In the second case when h value is changing, i.e. it is varied from 0.5 to 1100, the observation is, lambda value is depending on H value. so in order to achieve better correlation the value of H should be above 400 and as the lambda value tending to 1 , we are getting 100 % correlation of the extracted with respect to original signal.

II. REFERENCES

[1] D.Yellin and E. W Einstein. Criteria for Multichannel Signal Separation IEEE Transaction on Signal Processing, SP-42(8):2158-2168, August 1994.

[2] E. Bacharakis, A. K. Nandi, and V. Zarzoso. Foetal ECG Extraction using Blind Source Separation and Blind Deconvolution. Neural Computation, 7(6):1129-1159,1995

[3] L. D. Lathauwer, B. D. Moor, and J. Vandwalle. Blind Source Separation by Higher-Order Singular Decomposition. In Proceedings IEEE/ATHOS Signal Processing Conference on Higher-Order Statistics, pages 134-138, Spain, june 1995.

[4] A. K. Nandi and V. Zarzoso. Fourth-Order Cumulant Based Blind Source Separation. IEEE Signal Processing Letters, 3(12):312-314,December 1996.

[5] Hyvarinen, A, Complexity Pursuit: Separating interesting Components from time series. Neural Computation, 13, 883-898. 2001

[6] Brown, GD, Yamada, S, & Sejnowski, TJ. Independent components analysis (ica) at the neural cocktail party. *Trends in neuroscience*, **24**(1), 54–63. 2001

[7] Porrill, J, & Stone, JV. *Independent components analysis for signal separation and dimension reduction*. 118-124.1997 (August). Psychology Department, Sheffield University, Sheffield, UK, available from http://www.shef.ac.uk/˜pc1jvs

[8] *Nguyen T. K. Cuong. "Removing Noise and Artifacts from EEG Using Adaptive Noise Cancelator and Blind Source Separation", IFMBE Proceedings, 2010*

[9] *Hyvarinen, A.. "Independent component analysis: recent advances", Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences ,2012..*

[10] *Y. Yang. " Complexity Pursuit For Unifying Model", Neural Process Letter, 2010.*

[11] *Georgescu, Vasile, and Dumitru Ciobanu. "separating independent components underlying parallel foreign exchange rates: reconstructing and forecasting strategies", World Scientific Proceedings Series on Computer Engineering and Information Science, 2013.*