International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July-2013 763
ISSN 2229-5518

ICA Based Source Separation

Anamika Mishra, Shubhendu Kumar Sarangi, P. Shivani Sahoo

Abstract— Independent Component Analysis is a statistical and computational method for finding underlying factors or components form multivariate (multidimensional) statistical data. What distinguishes ICA from other method is that it looks for components that are both statistically independent and non-Gaussian. ICA defines a generative model for the observed multivariate data, which is typically given as a large database of samples. In the model, the data variables are assumed to be linear mixtures of some unknown latent variables and the mixing system is also unknown. The latent variables are assumed non-Gaussian and mutually independent and they are called the independent components of the observed data.In this paper FastICA algorithm is used to separate various types of signals.

Index Terms— Independent Component Analysis, FastICA

—————————— ◆ ——————————

1 Introduction

Independent Component Analysis is a signal processing technique whose goal is to express a set of random variables as linear combinations of statistically independent component variables. To define ICA, assume n linear mixtures $x_1 \ldots x_n$ of $\boldsymbol{n}$ independent

$$x_j = a_{j1}s_1 + a_{j2}s_2 + \ldots + a_{jn}s_n, \text{ for all } j. \quad (1)$$

Now, ignoring the time index t, in the above mixing model, each mixture x_j and each independent component s_k can be assumed as a random variable. Defining the ICA model in matrix notation,

$$x = As \quad (2)$$

Where, **x** is the random vector with elements that are the mixtures $x_1 \ldots x_n$ and **s** is the random vector with elements $s_1, \ldots s_n$, and **A** is the mixing matrix with elements a_{ij}. The statistical mode in eq.(2) is called independent component analysis, or ICA model. Sometimes we need the columns of matrix A; denoting them by a_i, the model can be written as,

$$x = \sum_{i=1}^{n} a_i s_i \quad (3)$$

The mixing matrix is assumed to be unknown. Here, all we observe is the random vector x and we aim to estimate both **A** and **s** using it. The ICA estimation is done under two important assumptions, (i) the components s_i, are statistically independent and, (ii) the independent components must have non-Gaussian distributions. For simplicity, it is also assumed that the mixing matrix is square. Then, after estimating the matrix A, we can obtain the independent components by:

$$s = Wx \quad (4)$$

Where, **W** is the demixing matrix and can be calculated by the inverse of **A**.

2 Methodology

There are several types of ICA algorithms among which the FastICA has been used in this work. Basically, it consists of two steps, the preprocessing step and the FastICA algorithm itself. The preprocessing consists further of two steps known as centering and whitening. The centering step is done by subtracting the mean of the observed data x, so as to make x a zero-mean variable. A whitening step is done to remove the correlation between the components of the observed data. One popular method for whitening is to use the eigenvalue decomposition (EVD) of the covariance matrix of the mixed signal. The final step is the FastICA algorithm which is briefly summarized as follows:

1. Choose an initial (e.g. random) weight vector **w**.
2. Let $\mathbf{w}^+ = E\{\mathbf{x}g(\mathbf{w}^T\mathbf{x})\} - E\{g'(\mathbf{w}^T\mathbf{x})\}\mathbf{w}$
3. Let $\mathbf{w} = \mathbf{w}^+ / \|\mathbf{w}^+\|$
4. If not converged, go back to 2.

The term *converge* in step 4 above refers to the condition that the value of **w** in the current iteration is same as the previous **w** value. ***E*** denotes the expectation. The function $g(.)$ can be chosen either as:

$$g(u) = \tanh(a_1.u) \text{ or } g(u) = u.\exp(-u^2/2),$$

Where, a_1 is any value that fulfils $1 \leq a \leq 2$.

3 Results and Discussions

We have taken sine, square and triangular wave of different frequencies, with different harmonics and different combinations, these signals are mixed in a random proportion to get

***Anamika Mishra** , Dept. of Electronics & Communication Engg. , ITER, SOA University, Odisha,India*
mishra.srujana089@gmail.com
***Shubhendu Kumar Sarangi**, Dept. Electronics & Instrumentation Engg. , ITER, SOA University, Odisha,India*
shubhendu1977@gmail.com
***P. Shivani Sahoo**, Dept. of Electronics & Communication Engg. , ITER, SOA University, Odisha,India*
pshivanisahoo@gmail.com

ISSN 2229-5518

the mixed signals and then are extracted back using the algorithm. Some speech signals are also taken into consideration. The order in which the outputs are recovered could not be predicted. The variances of the source signals are assumed to be equal to one.

We have taken six combinations of different signals, mixed them and then separated them using the FastICA algorithm. The joint distributions of the signals are also analysed. The type of signals taken, number of samples taken and the type of functions used are mentioned in a table below.

3.1 Simulation result1:

Signals used	sine wave, square wave
Number of samples taken	10000
Functions used	tanh(*u*), 1-tanh²(*u*)

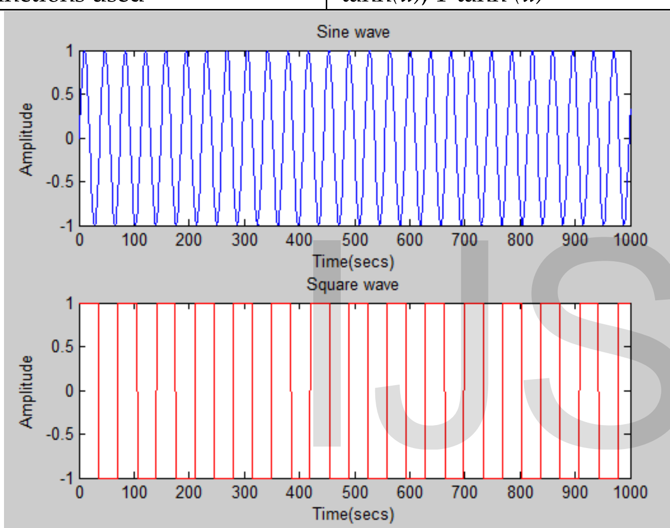


Figure 3.1(a): Original signals taken in first simulation

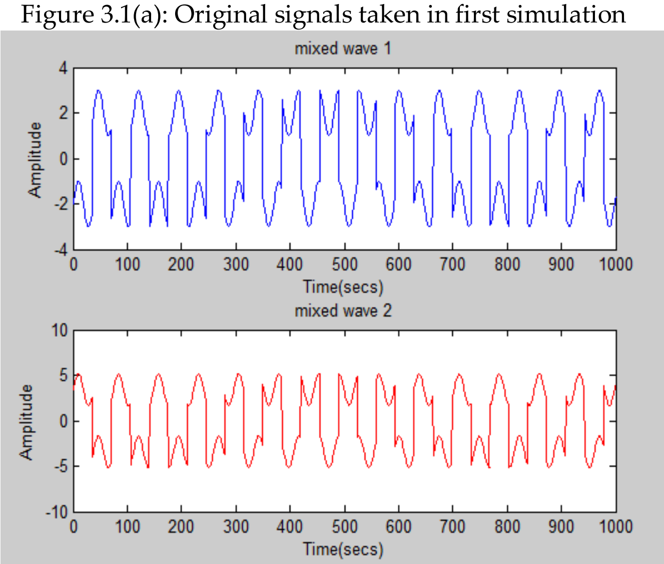


Figure 3.1(b): Mixed signals in first simulation

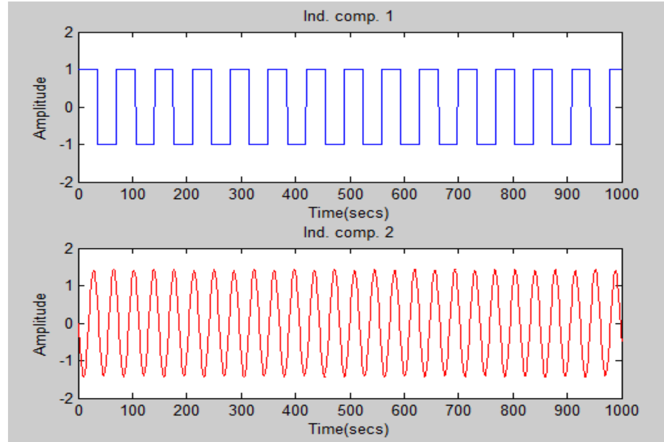


Figure 3.1(c): Recovered signals in first simulation

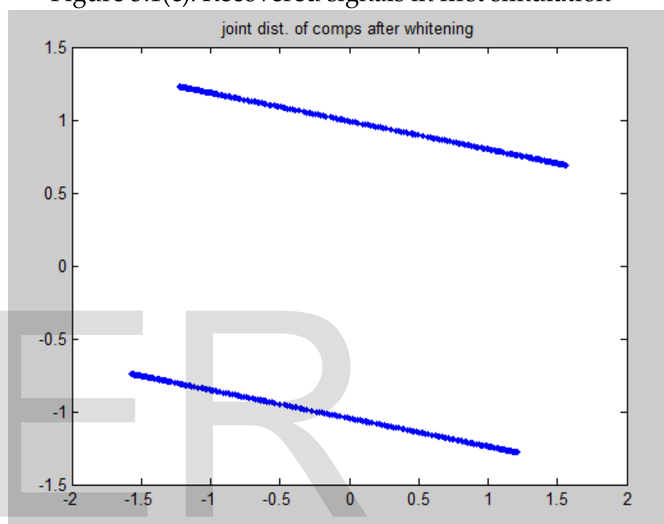


Figure 3.1(d): Joint distribution of signals in first simulation

3.2 Simulation result2:

Signals used	sine wave, square wave, 7th harmonics of sine wave
Number of samples taken	10000
Functions used	tanh(*u*), 1-tanh²(*u*)

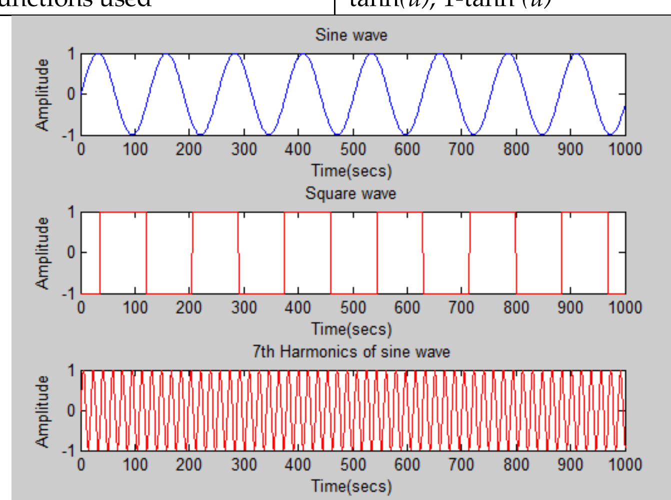


Figure 3.2(a): Original signals taken in second simulation

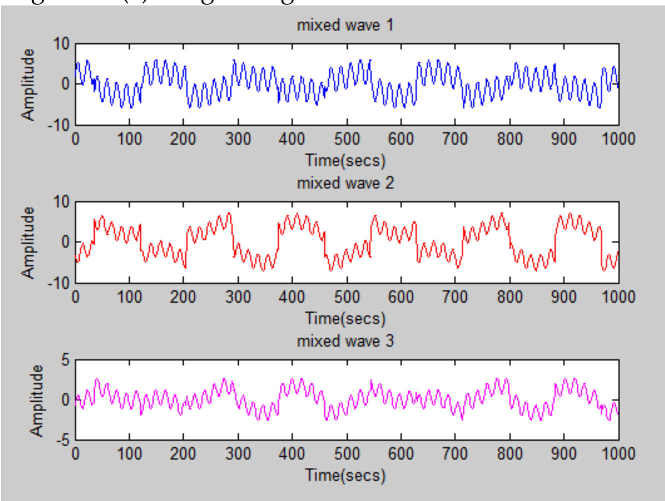


Figure 3.2(b): Mixed signals in second simulation

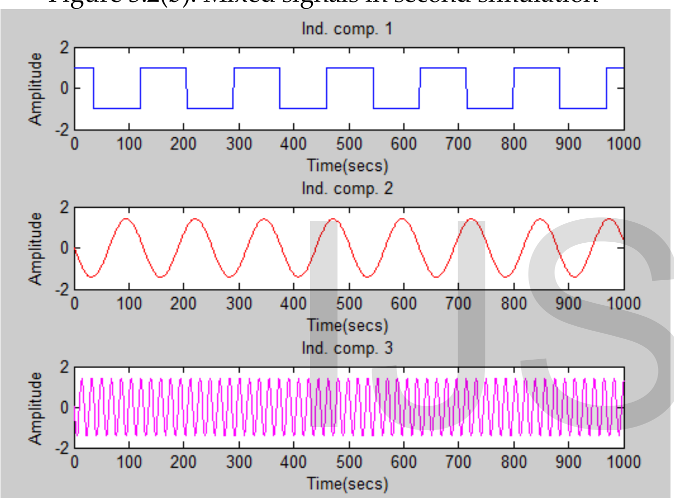


Figure 3.2(c): Recovered signals in second simulation

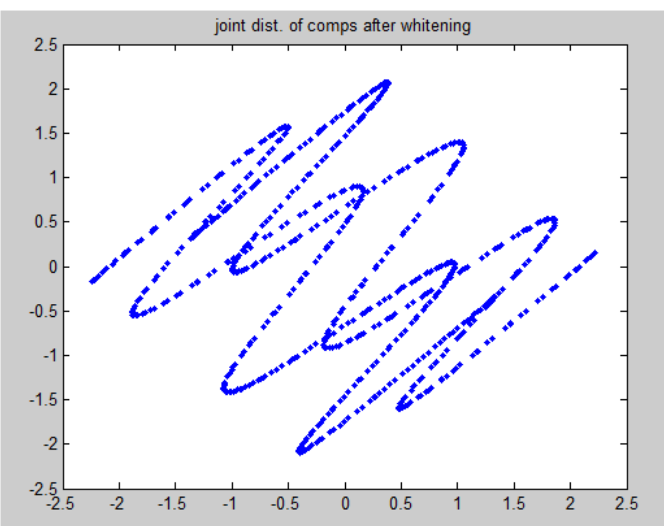


Figure 3.2(d): Joint distribution of signals in second simulation

3.3 Simulation result3:

Signals used	5^{th} harmonics of sine wave, 3^{rd} harmonics of square wave
Number of samples taken	10000
Functions used	$\tanh(u)$, $1-\tanh^2(u)$

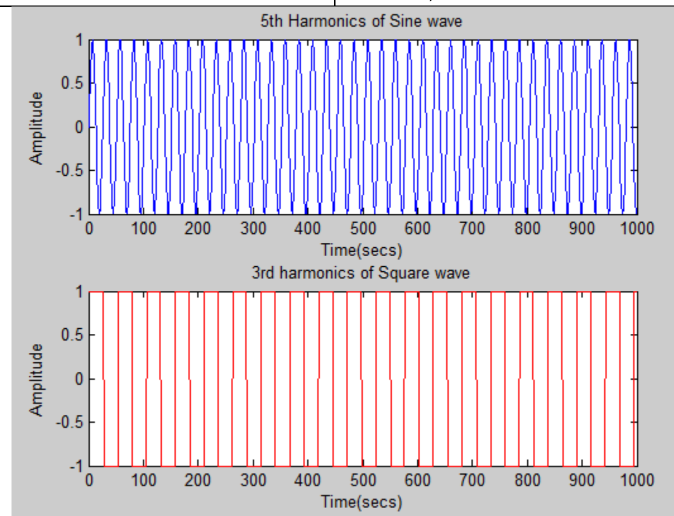


Figure 3.3(a): Original signals taken in third simulation

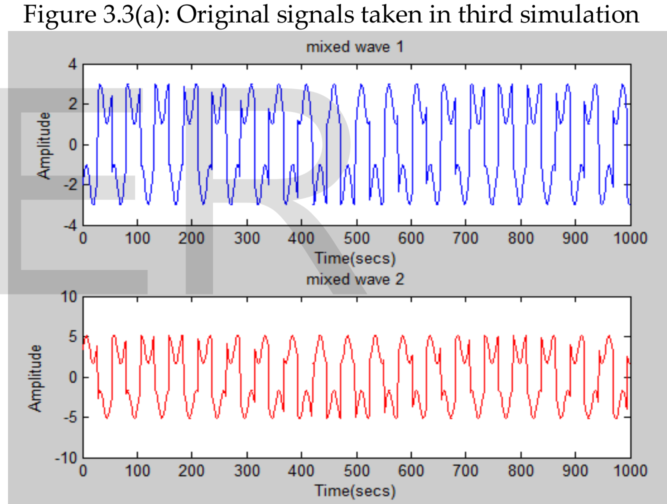
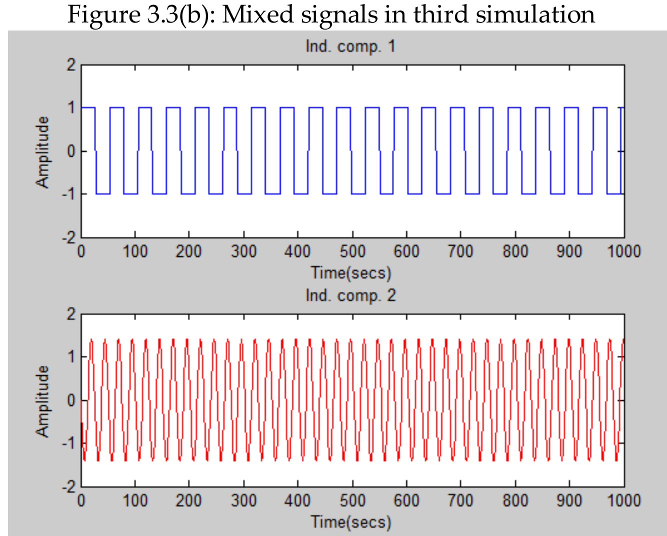


Figure 3.3(b): Mixed signals in third simulation



ISSN 2229-5518

Figure 3.3(c): Recovered signals in third simulation

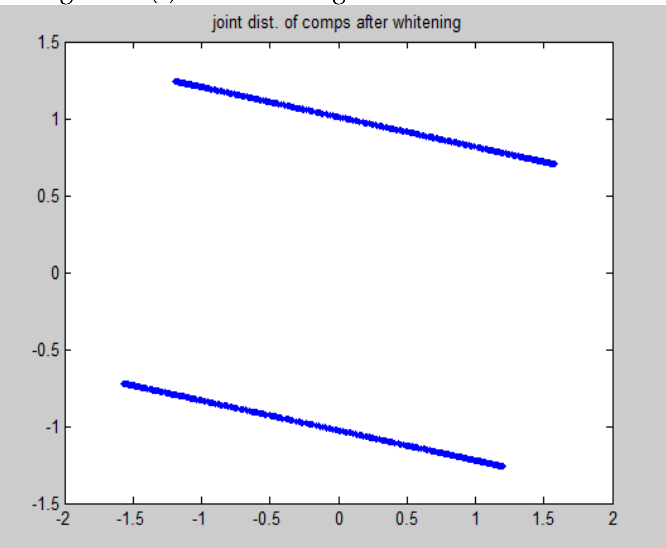


Figure 3.3(d): Joint distribution of signals in third simulation

3.4 Simulation result4:

Signals used	Two different speech signals
Number of samples taken	500000
Functions used	tanh(*u*), 1-tanh²(*u*)

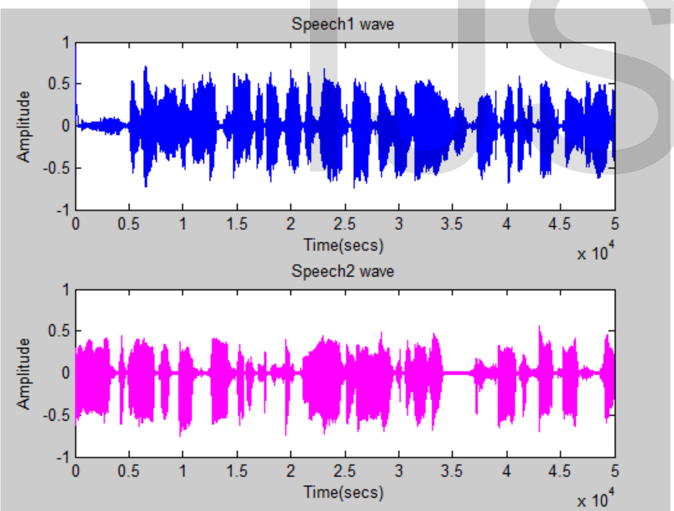


Figure 3.4(a): Original signals taken in fourth simulation

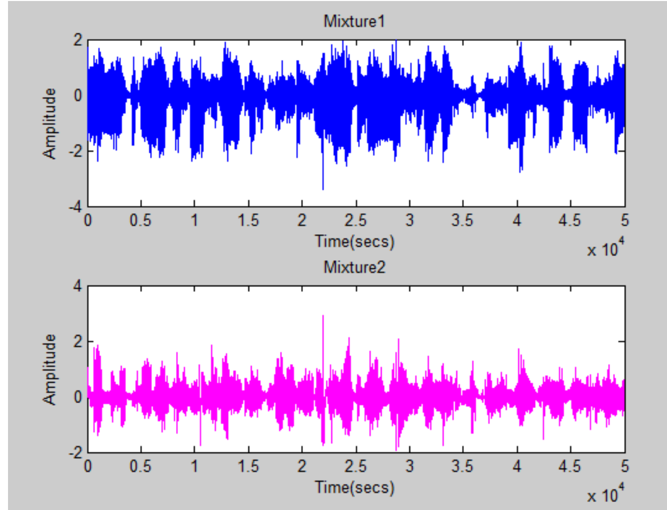


Figure 3.4(b): Mixed signals in fourth simulation

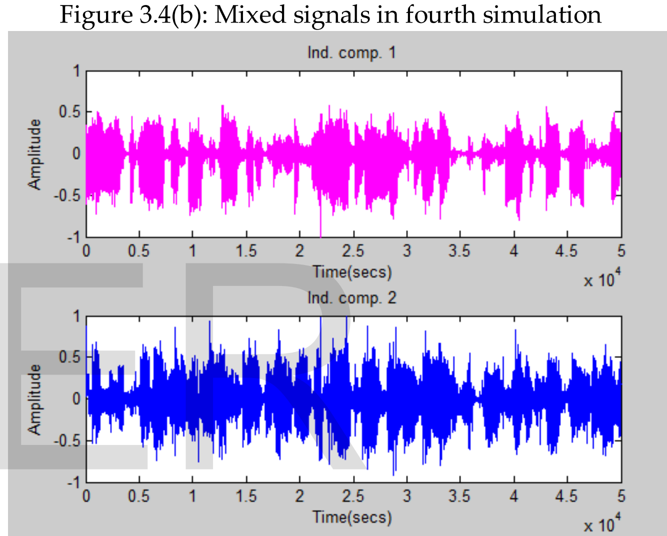


Figure 3.4(c): Recovered signals in fourth simulation

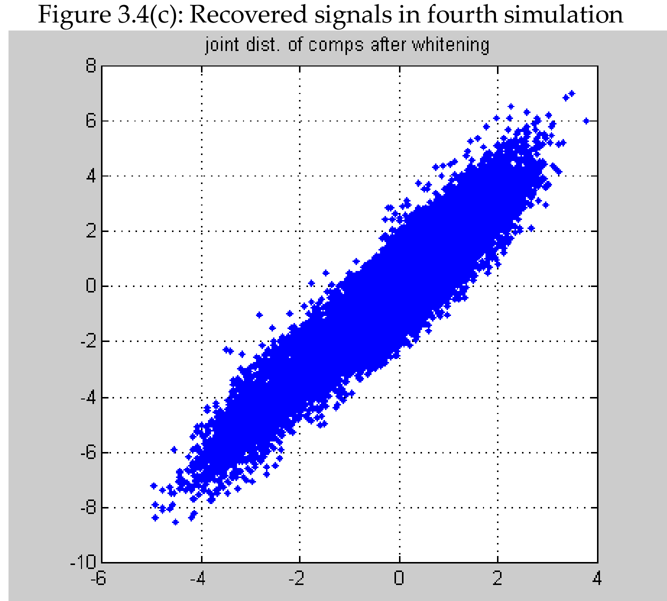


Figure 3.4(d): Joint distribution of signals in fourth simulation

3.5 Simulation result5:

Signals used	Sine, square and triangular wave
Number of samples taken	10000
Functions used	tanh(u), 1-tanh$^2(u)$

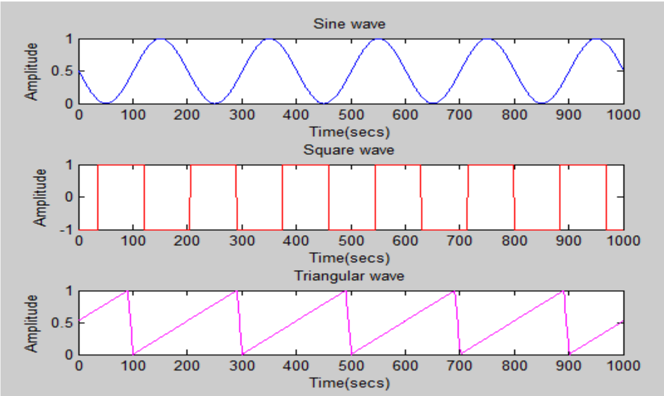


Figure 3.5(a): Original signals taken in fifth simulation

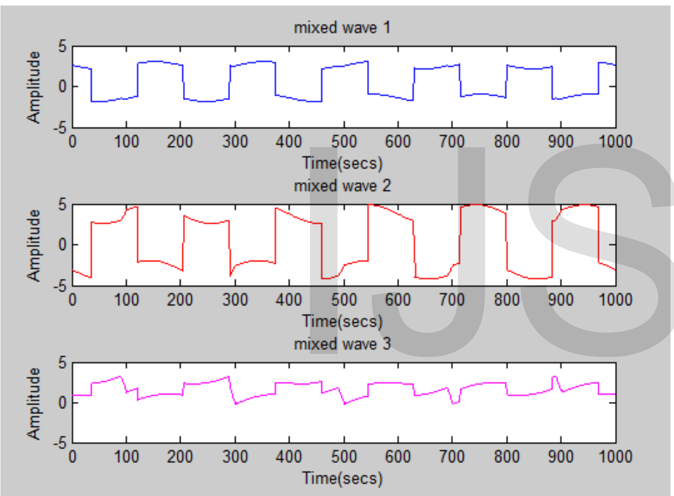


Figure 3.5(b): Mixed signals in fifth simulation

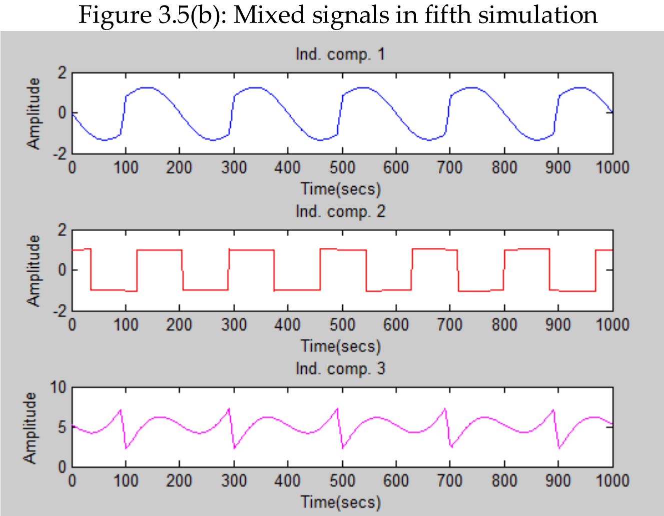


Figure 3.5(c): Recovered signals in fifth simulation

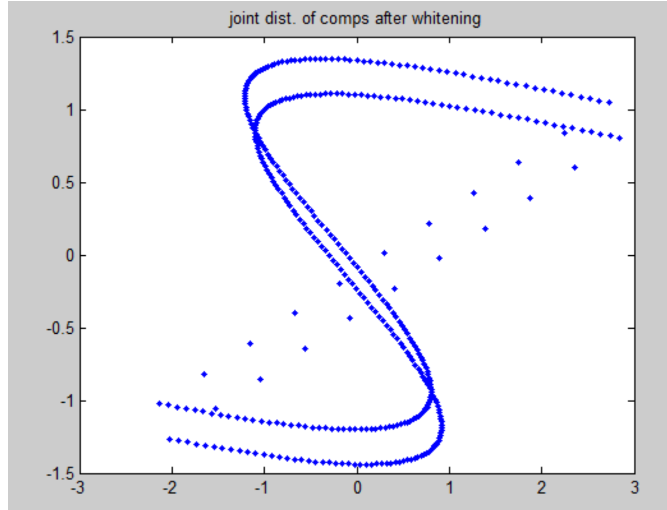


Figure 3.5(d): Joint distribution of signals in fifth simulation

4 Conclusion

In this paper, various mixed signal separation using the FastICA algorithm is shown .The result shows that the Fast ICA algorithm is an effective tool for denosing of mixed signal. The error signal between the original and recovered signal decreases by increasing the number of iterations.

Acknowledgment

The authors thank Archana Sarangi and Alka Barik of Department of ECE, ITER, SOA University for their support.

References

[1] Comon, P. (1994). Independent component analysis—a new concept? *Signal Processing*, 36:287–314.

[2] Hyvarinen, J. Karhunen, and E. Oja, — Independent component analysis: *Algorithms and applications,*‖ Neural Netw.,vol.13.,pp.411-430,May 2000.

[3] Hyvärinen, A. (1998a). Independent component analysis in the presence of gaussian noise by maximizing joint likelihood. *Neurocomputing*, 22:49–67.

[4] Hyvärinen, A. (1999a). Fast and robust fixed-point algorith ms for independent component analysis. *IEEE Trans-actions on Neural Networks*, 10(3):626–634.

[5] Hyvärinen, A. (1999e). Survey on independent component ana lysis. *Neural Computing Surveys*, 2:94–128.

[6] Simon Haykin, Communications Research Laboratory, McMaster University, Hamilton, Ontario, Canada, *"Adaptive Filter Theory"*, Fourth Edition, Pearson Education, Inc. 2002.