# EKENS: A Learning on Nonlinear Blindly Mixed Signals

W.Y.Leong, J.Homer
School of Information Technology and Electrical Engineering
The University of Queensland
Australia
leong@itee.uq.edu.au, homerj@itee.uq.edu.au

**Abstract-** *We present experimental results of the blind separation of independent sources from their nonlinear mixtures. The proposed EKENS algorithm is a generalization of natural gradient algorithm and Gram-Charlier series, which is extended in two ways: (1) to deal with nonlinear mapping, and (2) to be able to adapt to the actual statistical distributions of the sources by estimating the kernel density distribution at the output signals. In this paper, the observations are modelled based on nonlinear generative multilayer perceptrons analysis. The theory of the EKENS learning algorithm is discussed. Simulations show that the EKENS algorithm is able to find the underlying sources from the observation, even though the data generating mapping is nonlinear and unknown.*

## I .Introduction

Linear Independent Component Analysis (ICA)[1] and linear Blind Source Separation (BSS) from linear mixtures are relatively well-established approaches with many techniques[2, 3]. These unsupervised blind learning methods are often based on a generative approach, where the goal is to explain how the observations were generated from the independent components sources. It is assumed that there exist certain source signals which have generated the observed data through an unknown mapping. The goal of the proposed generative learning is to identify both the source signals and the unknown generative mapping.

However, it is evident that nonlinear mixing is far more difficult than the linear case. This nonlinear mapping problem has attracted attention recently [4, 5]. In the non-linear mixture model, the linear ICA theory and the equivariant property might not be able to deduce the nonlinear mapping. Therefore, the blind separation algorithms for the linear mixture model generally fail to extract the independent sources from the non-linear mixtures.

In this paper we consider ICA as the problem of transforming a set of patterns $\boldsymbol{x}$ (vectors of size $n$, often called observations), whose components are not statistically independent from one another, into patterns $\mathbf{y} = \boldsymbol{W}(\boldsymbol{g}(\boldsymbol{x}))$ whose components are as independent from one another as possible. In the nonlinear mapping case, $\boldsymbol{g}$ is a nonlinear multilayer network and $\boldsymbol{W}$ is the unmixing matrix. In the blind source separation application, one further assumes that the observations are the result of a mixture of statistically independent sources, $s_i$, i.e. $\boldsymbol{x} = \mathbf{M}(\mathbf{As})$, $s_i$ being the $i$th component of $\mathbf{s}$, $\mathbf{A}$ is a $n$x$n$ unknown full-rank and non singular mixing matrix, and $\mathbf{M}$ is the set of invertible nonlinear transfer functions. The purpose of BSS is to recover the sources from the observations.

Generally, the nonlinear mapping is rather unconstrained and difficult, and normally demands a good dependence measure. Many contributions to the nonlinear problem already exist. Taleb and Jutten [4] proved that the source independence assumption is not strong enough in the general nonlinear case. They proposed a direct estimation of the score functions [4] by minimizing the mean square error of the parameter vector in post-nonlinear mixtures. The results show that the least mean square estimation of the score functions performs well with hard nonlinearities. Also, Taleb [6] investigated a structured nonlinear model framework, and proposed a stochastic algorithm designed to deal with the parametric nonlinear mixtures. Valpola, Giannakopoulos, Honkela and Karhunen [5] proposed an alternate approach using Bayesian ensemble learning in nonlinear ICA. The ensemble learning provides the necessary regularisation for nonlinear ICA by choosing the model and sources that have most probably generated the observed data.

Our goal is to develop a new method, which we term EKENS, for inferring the original sources from the nonlinear observations alone. The nonlinear mapping from the unknown sources to the observations is modelled with the multi-layer perceptron (MLP) network. The main objective of this paper is to investigate a class of nonlinear transformations in mixing systems for which an iterative and equivariant treatment is presented. The paper is organized as follows: In section II, the nonlinear model is presented and a set of learning rules is derived in section III based on Gram-Charlier criterion. The learning rules are verified via simulation in section IV. The concluding remarks are discussed in section V.

## II. Equivariant Kernel Nonlinear Separation (EKENS)

Genetic algorithms (GA) are currently one of the most popular class of stochastic optimisation techniques. In this work, we propose a new GA--Equivariant Kernel Nonlinear Separation (EKENS) algorithm. This consists of obtaining an estimate of the pdf $\boldsymbol{F}(\boldsymbol{x})$ and, subsequently, application of the nonlinear function $\boldsymbol{f}_y(\boldsymbol{y}) = \frac{d}{dy}\log\boldsymbol{F}(\boldsymbol{y})$. We employ Gram-Charlier and Edgeworth series expansion [7, 8] to approximate the probability distribution $\boldsymbol{F}(\boldsymbol{x})$. The key idea of these expansions is to write the characteristic distribution function of the probability density function of $\boldsymbol{F}$. These distribution functions are approximated to characterise the distribution properties. The summary of the Gram-Charlier and Edgeworth series is as below.

## Gram-Charlier Kernel Density Estimation, *F*

*Preliminary*: The global separation system for the nonlinear mixtures consists of both nonlinear and a linear stage[9]. The nonlinear stage consists of $n$ parametric nonlinear functions to cancer the post-distortion. The linear stage consists of a regular separating matrix $\boldsymbol{W}$ devoted to the separation of the linear mixture. The nonlinear stage can be performed by constructing a nonlinear transform $\boldsymbol{g}$ to isolate each component of the observation vector, $\boldsymbol{x}$,

$$\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x} - \eta_0 * \boldsymbol{F}(\boldsymbol{x}) \tag{2.1}$$

where $\eta_0$ is a positive adaptation step size and $\boldsymbol{F}(\boldsymbol{x})$ is the probability density function of the nonlinear observation $\boldsymbol{x}$.

When the true probability distribution function (pdf) of a random variable $\boldsymbol{x}$ is unknown, yet believed to be similar to a normal one, it is quite natural to approximate it with a pdf of the normal form. The Gram-Charlier and Edgeworth series [7, 8] optimizes the determination of the probability density function (pdf) by data moments. The pdf $\boldsymbol{F}(\boldsymbol{x})$ can be investigated via Hermite's polynomials ($\boldsymbol{H}_n(\boldsymbol{x})$) truncated expansion:

$$\begin{aligned}\boldsymbol{F}(\boldsymbol{x}) &= \frac{1}{\sigma\sqrt{2\pi}}\exp\left[\frac{-(\boldsymbol{x}-\boldsymbol{m})^2}{2\sigma^2}\right]\sum_{n=1}^{k}\boldsymbol{C}_n \bullet \boldsymbol{H}_n(\boldsymbol{x}) \\ &= \alpha_{\boldsymbol{m}\bullet\sigma}(\boldsymbol{x})\sum_{n=1}^{k}\boldsymbol{C}_n\boldsymbol{H}_n(\boldsymbol{x})\end{aligned} \tag{2.2}$$

where $\alpha_{\boldsymbol{m}\bullet\sigma}(\boldsymbol{x})$ is the Gaussian distribution having $\boldsymbol{m}$ as mean and $\sigma$ as standard deviation.

**Hermite's Polynomials Determination, $\boldsymbol{H}_n$**

The Hermite's polynomials are evaluated by the following iteration rules:

$$\boldsymbol{H}_n(\boldsymbol{x}) = \frac{(-1)^n}{\alpha_\sigma(\boldsymbol{x})}\boldsymbol{D}_x^n\alpha_\sigma(\boldsymbol{x}) \tag{2.3}$$

where $\alpha_{\sigma,\boldsymbol{m}}$ is the standard Gaussian function defined as

$$\alpha_{\sigma,\boldsymbol{m}}(\boldsymbol{x}) = \frac{1}{\sigma\sqrt{2\pi}}\exp[\boldsymbol{x}-\boldsymbol{m}]^2/2\sigma^2 \tag{2.4}$$

where $\boldsymbol{m}$ and $\sigma$ are the mean and standard deviation of the Gaussian which are given by

$$\begin{cases}\boldsymbol{m} = \boldsymbol{u}^1, \\ \sigma = \sqrt{\boldsymbol{u}^2 - (\boldsymbol{u}^1)^2}\end{cases} \tag{2.5}$$

and the terms $\boldsymbol{u}^k$ are $k$-order input moments defined as

$$\boldsymbol{u}^k = \int_{-\infty}^{+\infty}\boldsymbol{x}^k\boldsymbol{F}(\boldsymbol{x})\boldsymbol{dx} \tag{2.6}$$

The operator is defined as

$$\boldsymbol{D}_x^n = \frac{\boldsymbol{d}^n}{\boldsymbol{dx}^n} \tag{2.7}$$

It can be shown that $\boldsymbol{H}_n(\boldsymbol{x})$ can be calculated from $\boldsymbol{H}_{n-1}(\boldsymbol{x})$ and $\boldsymbol{H}_{n-2}(\boldsymbol{x})$. The Hermite's polynomials terms [8] are obtained as follow:

$$\begin{cases}\boldsymbol{H}_0(\boldsymbol{x}) = 1, \\ \boldsymbol{H}_1(\boldsymbol{x}) = \dfrac{(\boldsymbol{x}-\boldsymbol{m})}{\sigma^2}, \\ \boldsymbol{H}_2(\boldsymbol{x}) = \dfrac{(\boldsymbol{x}-\boldsymbol{m})^2}{\sigma^4} - \dfrac{1}{\sigma^2}, \\ \boldsymbol{H}_3(\boldsymbol{x}) = \dfrac{(\boldsymbol{x}-\boldsymbol{m})^3}{\sigma^8} - 3\dfrac{(\boldsymbol{x}-\boldsymbol{m})}{\sigma^4}, \\ \boldsymbol{H}_4(\boldsymbol{x}) = \dfrac{(\boldsymbol{x}-\boldsymbol{m})^4}{\sigma^8} - 6\dfrac{(\boldsymbol{x}-\boldsymbol{m})^2}{\sigma^6} + \dfrac{3}{\sigma^4} \\ \boldsymbol{H}_5(\boldsymbol{x}) = \dfrac{(\boldsymbol{x}-\boldsymbol{m})}{\sigma^2}\left(\dfrac{(\boldsymbol{x}-\boldsymbol{m})^4}{\sigma^8} - 10\dfrac{(\boldsymbol{x}-\boldsymbol{m})^2}{\sigma^6} + \dfrac{15}{\sigma^4}\right)\end{cases} \tag{2.8}$$

**Coefficient $\boldsymbol{C}_n$ Determination**

In Gram-Charlier's expansion, the input data are used to determine the moments up to order $k$ and the expansion gives the pdf for the continuous random variable $\boldsymbol{x}$. The coefficient $\boldsymbol{C}_n$ can be calculated via Hermite's polynomials and moments by the following orthogonalization process:

$$\begin{aligned}\boldsymbol{C}_n &= \frac{\int\boldsymbol{F}(\boldsymbol{x})\boldsymbol{H}_x(\boldsymbol{x})dx}{\int\boldsymbol{H}_x(\boldsymbol{x})\sigma(\boldsymbol{x})dx} \\ &= \frac{\sigma^{2n}}{\boldsymbol{n}!}\int\boldsymbol{F}(\boldsymbol{x})\boldsymbol{H}_n(\boldsymbol{x})dx\end{aligned} \tag{2.9}$$

It can be shown [8] that

$$\boldsymbol{C}_{(2n+\delta_d)} = \sum_{a=0}^{n}\frac{\sigma^{2\alpha}}{2^\alpha\alpha!}\sum_{\beta=0}^{2n+\delta_d-2\alpha}\frac{(-1)^{2n+\delta_d-\alpha-\beta}}{(2n+\delta_d-2\alpha-\beta)!\beta!}\boldsymbol{m}^{2n+\delta_d-2\alpha-\beta}\boldsymbol{u}^\beta, \tag{2.10}$$

where $n$ =0, 1, 2, 3, … and $\delta_d$ is defined as

$$\delta_d = \begin{cases}+1 & \text{if } (2n+\delta_d) \text{ is odd} \\ 0 & \text{if } (2n+\delta_d) \text{ is even}\end{cases} \tag{2.11}$$

Further, it can be shown that $\boldsymbol{C}_n$ is determined by $\boldsymbol{C}_{n-1}$ and $\boldsymbol{C}_{n-2}$. Thus, $\boldsymbol{C}_n$ can be calculated via an iterative process, which requires only the direct calculation of $\boldsymbol{C}_0$ and $\boldsymbol{C}_1$. The coefficients $\boldsymbol{C}_n$ are obtained as a function of the Gaussian parameters. The first six coefficients are

$$\begin{cases}\boldsymbol{C}_0 = \boldsymbol{u}^0, \\ \boldsymbol{C}_1 = \boldsymbol{u}^1 - \boldsymbol{m}\boldsymbol{u}^0, \\ \boldsymbol{C}_2 = \dfrac{1}{2}(\boldsymbol{u}^2 - 2\boldsymbol{u}^1\boldsymbol{m} + \boldsymbol{u}^0\boldsymbol{m}^2 - \sigma^2\boldsymbol{u}^0), \\ \boldsymbol{C}_3 = \dfrac{1}{3!}(\boldsymbol{u}^3 - 3\boldsymbol{u}^2\boldsymbol{m} + 3\boldsymbol{u}^1\boldsymbol{m}^2 - \boldsymbol{u}^0\boldsymbol{m}^3) - \dfrac{\sigma^2\boldsymbol{u}^1}{2} + \dfrac{\boldsymbol{u}^0\sigma^2\boldsymbol{m}}{2}, \\ \boldsymbol{C}_4 = \dfrac{1}{4!}[\boldsymbol{u}^4 - 4\boldsymbol{u}^3\boldsymbol{m} + 6\boldsymbol{u}^2(\boldsymbol{m}^2 - \sigma^2) - 4\boldsymbol{u}^1(\boldsymbol{m}^3 - 3\sigma^2\boldsymbol{m}) \\ \quad + \boldsymbol{u}^0(\boldsymbol{m}^4 + 3\sigma^4) - 6\sigma^2\boldsymbol{m}^2], \\ \boldsymbol{C}_5 = \dfrac{1}{5!}[\boldsymbol{u}^5 - 5\boldsymbol{u}^4\boldsymbol{m} + 10\boldsymbol{u}^3(\boldsymbol{m}^2 - \sigma^2) - 10\boldsymbol{u}^2(\boldsymbol{m}^3 - 3\sigma^2\boldsymbol{m}) \\ \quad + 5\boldsymbol{u}^1(\boldsymbol{m}^4 - 6\sigma^2\boldsymbol{m}^2 + 3\sigma^4) - \boldsymbol{u}^0(\boldsymbol{m}^5 - 10\sigma^2\boldsymbol{m}^3 + 15\sigma^4\boldsymbol{m})],\end{cases} \tag{2.12}$$

**Weight Determination**

In all simulations, the length $t$ of the source signal sequence is 50 and the total number of iterations is 3500, where one iteration involves processing all the observations. As shown by the results, the EKENS algorithm is able to recover the sources from nonlinear mixtures that involved relatively smooth nonlinearities. The experimental results of the above algorithms are presented for 2, 4 and 6 mixtures of sources. This section reports the performance by presenting some illustrative examples. The six independent and zero mean source signals are given in eq(3.1).

Following the Gram-Charlier expansion based estimation of the pdf $\boldsymbol{F}(\boldsymbol{x})$, the nonlinear density function $\boldsymbol{f}_y(\boldsymbol{y})$ is estimated, $\boldsymbol{f}_y(\boldsymbol{y})=\dfrac{\boldsymbol{d}}{\boldsymbol{dy}}\log\boldsymbol{F}(\boldsymbol{y})$. The iterative equivalent gradient algorithm for the estimation of the unmixing matrix $\boldsymbol{W}$ is then formed as follow:

$$\boldsymbol{W}(k+1) = \boldsymbol{W}(k)+ \eta_0\,[\boldsymbol{I} - (\boldsymbol{f}_y(\boldsymbol{y}_k))\boldsymbol{y}_k^{T}]\boldsymbol{W}(k) \qquad (2.13)$$

where $\eta_0$ is a positive adaptation step size, $\boldsymbol{I}$ is the identity matrix. At each iteration $k$, a new estimated density $\boldsymbol{f}_y(\boldsymbol{y}_k)$ is calculated using Gram-Charlier procedure, with the separation output signal $\boldsymbol{y}_k$ being the new observed signal $\boldsymbol{x}$.

## III. Experiments

In this section several experiments have been performed to evaluate the validity and performance of the EKENS algorithm compared with the conventional ICA based EASI [10] and Infomax [3, 11] algorithms. These experiments are used to assess the EKENS algorithm's ability to perform blind source separation in several nonlinear mixtures. The learning scheme for all the experiments is the same. First, normalization is used to reduce the dimension of the mixtures to $n$. It is also used to find sensible initial values for the posterior means of the mixtures. The initial weights $\boldsymbol{W}_0$ of the network have random values. The pre-processed mixtures will make the separation easier. The mixtures are then adjusted via the iterated weight matrix $\boldsymbol{W}$.

$$\begin{aligned}
s_1(t) &= \cos((1{:}t)*(4*\mathrm{pi})/t)\\
s_2(t) &= \sin((1{:}t)\bullet(10*\mathrm{pi})/t)\\
s_3(t) &= \cos((1{:}t)\bullet(2*\mathrm{pi})/t)\\
s_4(t) &= \mathrm{sawtooth}(t)\\
s_5(t) &= \mathrm{sinc}(1{:}t)\\
s_6(t) &= \mathrm{square}(1{:}t)
\end{aligned} \qquad (3.1)$$

We consider a two-channel, a four-channel and a six-channel nonlinear mixture with *tanh* nonlinearities as given in eq(3.3). The corresponding mixing matrices as given in eq(3.2).

$$\boldsymbol{A}_{2\text{ mixes}}=\begin{bmatrix}1 & 0.3\\ 0.75 & 1\end{bmatrix};\ \boldsymbol{A}_{4\text{ mixes}}=\begin{bmatrix}1 & 0.75 & 0.55 & 0.55\\ 0.45 & 0.9 & 0.75 & 0.7\\ 0.65 & 0.65 & 0.85 & 0.5\\ 0.5 & 0.55 & 0.5 & 0.95\end{bmatrix};$$

$$\boldsymbol{A}_{6\text{mixes}}=\begin{bmatrix}0.9 & 0.65 & 0.5 & 0.9 & 0.5 & 0.55\\ 0.65 & 0.95 & 0.5 & 1 & 0.55 & 0.55\\ 0.55 & 0.55 & 0.9 & 0.55 & 0.55 & 0.55\\ 0.55 & 0.55 & 0.55 & 1 & 0.55 & 0.55\\ 0.55 & 0.55 & 0.55 & 0.55 & 1 & 0.55\\ 1 & 0.65 & 0.65 & 0.65 & 0.65 & 1\end{bmatrix}; \qquad (3.2)$$

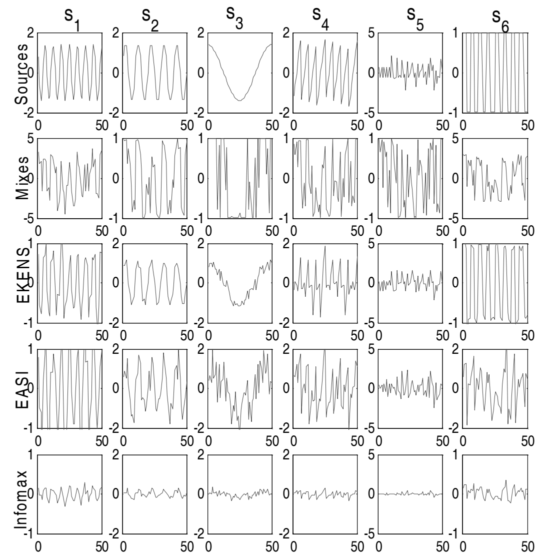


Figure 1: Top row: The original sources; Second row: The mixtures; Third row: The recovered sources using EKENS algorithm; Fourth row: The recovered sources using EASI algorithm; Fifth row: The recovered sources using Infomax algorithm.

$$\begin{aligned}
e_1&= (\boldsymbol{A}s_1) +\tanh(5* \boldsymbol{A}s_1)\\
e_2&= \tanh(\boldsymbol{A}s_2)\\
e_3&= \tanh(5*\boldsymbol{A}s_3)\\
e_4&= \tanh(\boldsymbol{A}s_4)\\
e_5&=\tanh(\boldsymbol{A}s_5)\\
e_6&=3*\tanh(\boldsymbol{A}s_6);
\end{aligned} \qquad (3.3)$$

The EASI, EKENS and Infomax have the advantage of learning the output nonlinearities during sampling. They are therefore adaptive to the actual statistical distributions of the sources. The sources (eq(3.1)) are independent because the values of one source does not convey any information about the other source. Our tests of nonlinear ICA were mainly aimed at showing this adaptability of the method to different nonlinear source distributions. The learning rate is fixed at $\eta_0 = 0.0001$.

In general, nonlinear mapping is quite complicated and difficult. It is interesting to evaluate the performance of the conventional linear ICA method and new proposed algorithm to deal with nonlinear mapping problems. For the 6 source mixing case, the separation results are depicted in Fig1, the original unknown sources (eq.3.1) are shown on row 1. The output displays for EKENS, EASI and Infomax are in row 3, 4 and 5 respectively. As

| Sources | Original *Kurtosis* | Mixes *Kurtosis* | *Kurtosis* (*6 sources*) | | | *Kurtosis* (*4 sources*) | | |
|---|---|---|---|---|---|---|---|---|
| | | | Recovered | | | Recovered | | |
| | | | EKENS | EASI | INFOMAX | EKENS | EASI | INFOMAX |
| 1 | -1.5554 | -1.3833 | -1.4139 | -0.5839 | -0.7961 | -1.4207 | -0.8405 | -1.2543 |
| 2 | -1.5594 | -1.6906 | -1.5806 | -1.0055 | -0.7462 | -1.6065 | -1.0055 | -0.3416 |
| 3 | -1.5594 | -1.9243 | -1.4895 | -0.7159 | -0.2504 | -0.6985 | -0.2504 | -0.1307 |
| 4 | -1.2665 | -1.5047 | -0.6941 | -0.8427 | -0.4347 | -0.1071 | -1.1176 | -0.7370 |
| 5 | -0.4634 | -1.6482 | -0.3111 | -0.1152 | 1.1721 | | | |
| 6 | -2.0396 | -1.5284 | -2.0053 | -0.8358 | -1.0126 | | | |

Table 1: The zero-mean white sources with sub-gaussian distribution tested with 4 and 6 nonlinear mixing.

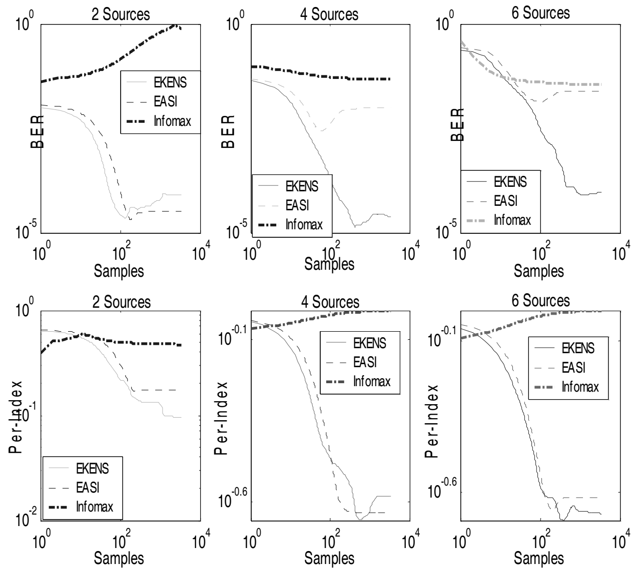


Figure 2: Top : BER; Bottom: Performance Index using a) ____ EKENS, b) ----- EASI and c) -.-.-.Infomax algorithm

shown in Fig.1, the EKENS is able to recover the sources from the nonlinear distorted mixtures. The reconstructed signals are quite accurate. Also, the EASI algorithm because of its equivariant characteristic estimates relatively well, but the output is relatively noisy. From our observation, the linear Infomax algorithm is not able to estimate the nonlinear mixture accurately. This experiment shows that this linear method fails to deduce the number of sources and the outputs.

A measure of separation performance is given by the similarity in *kurtosis* of the source and the corresponding unmixed signals. Table 1 shows the *kurtosis* results for 4 and 6-mixing. It can be seen that the EKENS algorithm provides, in general, better kurtosis matching of source and output signals. Note, the source signals are sub-gaussian and hence have negative kurtosis.

Another measure of separation performance is given by the performance matrix $\mathbf{P}=\boldsymbol{W}\bullet\boldsymbol{A}$. The $\bullet$ denotes multiplication. Perfect separation corresponds to $\mathbf{P}$ = identity. Table 2 shows that the performance matrix **P** for the EKENS and EASI algorithms is close to the identity matrix. This shows a clear separation of all sources from their nonlinear mixtures.

Figure 2 compares the bit-error-rate (BER) and performance index of the separated output over 3500 iterations. The simulation results suggest that both EASI and EKENS algorithms are sufficient to separate the true sources. The displayed BER and performance index are very low for EKENS methods between 2 and 6 mixtures. EKENS shows BER of $10^{-5}$, $10^{-5}$ and $10^{-4}$ for 2, 4 and 6 mixing respectively. However, EASI algorithm appears to suffer performance degradation with increased number of mixtures. EASI method shows BER of $10^{-1.5}$ for 6-mixing. Note both EKENS and EASI display fairly good performance values. The low performance index values for the EASI method however might be due to a lot of gaussian noise in the separated output signals. Also, we notice that the conventional linear Infomax fails to extract the nonlinear distorted mixtures.

## IV. Conclusions

In this paper, we have proposed EKENS to the problem of source separation in nonlinear mixtures, which consists of Gram-Charlier kernel density estimation. Also, we compared and investigated the EKENS, EASI and Infomax algorithms - methods for performing ICA by minimizing the mutual information of the estimated components. The experimental results show that EKENS algorithm has outperformed EASI and Infomax algorithms in a two layer nonlinear separation with lower BER and better independence. Also, the simulation results suggest that the EKENS algorithm is able to well-separate 2, 4 and 6 mixture signals; the EASI algorithm is able to well-separate 2 mixed signals and to a less extent 4 mixed signals; the linear Infomax fails to separate any of the nonlinear mixtures. In future, the proposed EKENS will be enhanced to deal with multi-layer nonlinear network.

## References:

[1] A.Hyvarinen, J.Karhunen, E.Oja, *Independent Component Analysis*. Canada: John Wiley & Sons, Inc, 2001.

[2] P.Comon, "Independent Component Analysis, A New Concept?," *Special Issue on Higher-Order Statistics*, vol. 36, pp. 287-314, April 1994.

[3] A.Hyvarinen and E.Oja, "Independent Component Analysis: A Tutorial," Helsinki University of Technology, Finland April 1999.

[4] A.Taleb, C.Jutten, "Source Separation in Post-Nonlinear Mixtures," *IEEE Transaction on Signal Processing*, vol. 47, pp. 2807-2820, 1999.

[5] H.Valpola, X.Giannakopoulos,A.Honkela,J.Karhunen, "Nonlinear Independent Component Analysis Using Ensemble Learning: Experiments and Discussion," Helsinki University of Technology, Neural Networks Research Centre, Finland 2000.

[6] A.Taleb, "A generic framework for blind source separation in structured nonlinear models," *IEEE Trans. on Signal Processing*, vol. 50, pp. 1819 - 1830, 2001.

[7] A.Pelliccioni, "The Gram-Charlier Method to Evaluate the Probability Density Function in Monodimensional Case," *Il Nuovo Cimento*, vol. 20, 1997.

[8] M.O'brien, "Using the Gram-Charlier expansion to produce Vibronic Band Shapes in Strong Coupling," *Journal Physics: Condensed Matter*, vol. 4, pp. 2347-2359, 1992.

[9] A.Taleb, "Source Separation in Structured Nonlinear Models," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 6, pp. 3613-3516, 2001.

[10] J.F.Cardoso and B.H.Laheld, "Equivariant Adaptive Source Separation," *IEEE Transactions on Signal Processing*, vol. 44, pp. 3017-3024, Dec 1996.

[11] Bell and Sejnowksi, "An Information Maximization Approach to Blind Separation and Blind Deconvolution," *Neural Computation*, vol. 7, pp. 1129-1159, 1995.

| Performance Matrix, **P** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *EKENS* | | | | *EASI* | | | | *Infomax* | | | |
| 0.9620 | 0.1545 | 0.0554 | -0.0877 | 1.3051 | 0.2750 | 0.0536 | -0.1015 | 0.1424 | 0.1156 | 0.0791 | 0.0994 |
| 0.0138 | 0.9770 | -0.0761 | -0.0405 | -0.3092 | 1.0223 | 0.0726 | 0.0998 | 0.0514 | 0.1417 | 0.0676 | 0.0850 |
| -0.0006 | 0.1117 | 0.8654 | 0.3408 | 0.2080 | 0.2668 | 0.8768 | 0.2551 | 0.0796 | 0.1021 | 0.1121 | 0.0993 |
| 0.4626 | -0.1840 | -0.4033 | 0.9369 | 0.1310 | -0.0682 | 0.1339 | 1.1218 | 0.0870 | 0.0842 | 0.0754 | 0.1339 |

Table 2: The performance matrix **P** for 4 mixed sources after separation. **P** is close to the identity matrix after rescaling and reordering.