International Journal of Computer Engineering Research Vol. 2(1), pp. 1- 7, February 2011
Available online at http://www.academicjournals.org/JCER
ISSN 2141-6494 ©2011 Academic Journals

*Full Length Research paper*

# Blind signal separation based on generalized laplace distribution

**M.EL-Sayed Waheed, Osama Abdo Mohamed and M.E. Abd El-Aziz***

Department of Mathematics, Faculty of science, Zagzig University, Egypt.

Accepted 18 September, 2010

**Blind Signal Separation is the task of separating signals when only their mixtures are observed. Recently, Independent Component Analysis has become a favorite method of researchers for attacking this problem. We propose a new score function based on Generalized Laplace Distribution for the problem of blind signal separation for supergaussian and subgaussian. To estimate the parameters of such score function we used Nelder-Mead algorithm for optimizing the maximum likelihood function of Generalized Laplace Distribution. To blindly extract the independent source signals, we resort to FastICA approach. Simulation results show that the proposed approach is capable of separating mixture of signals.**

**Key words:** Independent component analysis (ICA), generalized laplace distribution (GLD), maximum likelihood (ML), Nelder-Mead (NM).

## INTRODUCTION

A blind source separation (BSS) algorithm aims to recover sources from a number of observed mixtures. The problem that it is solving can be formulated statistically as follows: given M-dimensional random variable vector $\mathbf{x}(t)=[x_1(t),..,x_M(t)]^T$ that arises from linear combination of the mutually independent components of N-dimensional unknown random variable $\mathbf{s}(t)=[s_1(t),..,s_N(t)]^T$ represented mathematically as

$$\mathbf{x}(t)=\mathbf{A}\mathbf{s}(t) \quad \mathbf{t}=1,2,..,M \tag{1}$$

Where $\mathbf{x}\in\mathbf{R}^M$, $\mathbf{s}\in\mathbf{R}^N$ and A is an M x N mixing matrix. Here, R denotes the field of real numbers. The class of algorithms that handle such a problem is also called independent component analysis (ICA). When the number of the mixtures is equal to that of the sources (that is, M=N), the objective can be refined to find an N x N invertible square matrix W such that

$$\mathbf{u}(t)=\mathbf{W}\mathbf{x}(t) \quad \mathbf{t}=1,2,..,N \tag{2}$$

Where the components of estimated source $\mathbf{u}(t)=[u_1(t),...,u_N(t)]^T$ are mutually independent as much as possible. This must be done as accurately as possible with the assumption that no more than one source has a Gaussian distribution. Current algorithms can meet this objective within a permutation and scaling of the original sources. In general, the majority of BSS approaches perform ICA, by essentially optimizing the negative log-likelihood (objective) function with respect to the unmixing matrix W such that

$$L(\mathbf{u},\mathbf{W})=\sum E[\log p_{u_i}(u_i)]-\log|\det(\mathbf{W})| \tag{3}$$

Where E [.] represents the expectation operator and $p_{u_i}(u_i)$ is the model for the marginal probability density function (pdf) of $u_i$, for all $i=1,2,...,n$. Normally, matrix W is regarded as the parameter of interest and the pdfs of the sources are considered to be nuisance parameters. In effect, when correctly hypothesizing upon the distribution of the sources, the maximum likelihood (ML) principle leads to estimating functions, which in fact are the score functions of the sources (Cardoso, 1998).

*Corresponding author. E-mail: abd_el_aziz_m@yahoo.com.

$$\varphi_i(u_i) = -\frac{d}{du_i}\log p_{u_i}(u_i) \quad (4)$$

In principle, the separation criterion in (3) can be optimized by any suitable ICA algorithm where contrasts are utilized (Cardoso, 1998). A popular choice of such a contrast-based algorithm is the so-called fast (cubicly) converging Newton-type (fixed-point) algorithm, normally referred to as FastICA (Hyvarinen and Oja, 1997), and based on:

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mathbf{D}(\mathrm{E}[\varphi(\mathbf{u})\mathbf{u}^T] - diag(E[\varphi_i(u_i)u_i]))\mathbf{W}_k \quad (5)$$

Where, as defined in Karvanen and Koivunen (2002)

$$\mathbf{D} = \mathrm{diag}(1/(\mathrm{E}[\varphi_i(u_i)u_i] - \mathrm{E}[\varphi_i'(u_i)])) \quad (6)$$

With $\varphi(t) = [\varphi_1(u_1), \varphi_2(u_2), ...., \varphi_n(u_n)]^T$ being valid for all $i = 1,2,...,n$. In the ICA framework, accurately estimating the statistical model of the sources at hand is still an open and challenging problem (Cardoso, 1998). Practical BSS scenarios employ difficult source distributions and even situations where many sources with very different pdfs are mixed together. Since these densities are often unknown, unrealistic assumptions about the score functions employed that can seriously compromise the performance and convergence properties of the algorithms in question can be made. This calls for a FastICA method that introduces source adaptively through a well-matched parametric (adaptive) score function (Kokkinakis and Nandi, 2006).

## GENERALIZED LAPLACE DISTRIBUTION (GLD)

Subbotin (1923) proposed a generalization of the Laplace distribution with pdf:

$$f_i(x \mid \mu_i, \sigma_{p_i}, p_i) = \frac{1}{2p_i^{1/p_i}\sigma_{p_i}\Gamma(1+1/p_i)}\exp\left(-\frac{|x-\mu_i|^{p_i}}{p_i\sigma_{p_i}^{p_i}}\right) \quad (7)$$

Where $-\infty < \mathbf{x} < \infty$, $\mu_i$ is the location parameter, $\sigma_{p_i}$ is the scale parameter, $p_i > 0$ is the shape parameter and $\Gamma(\alpha)$ is the Gamma function, defined by

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt$$

The incomplete gamma function defined by

$$\gamma(\alpha, x) = \int_0^x z^{\alpha-1}e^{-z}dz$$

The complementary incomplete gamma function defined by

$$\Gamma(\alpha, x) = \int_x^\infty z^{\alpha-1}e^{-z}dz$$

The generalized Laplace is sometimes referred to as the exponential power function distribution. This distribution is widely used in Bayesian inference (Box and Tiao, 1962; Tiao and Lund, 1970). Estimation issues related to Equation (7) are discussed in [Agr 'o, 1995; Zeckhauser and Thompson, 1970). Using the definition of the incomplete gamma functions, one can write the cdf corresponding to (7) as

$$F_i(u) = \left(1/(2\Gamma(1/p_i))\right)\begin{cases}\Gamma\left(\frac{1}{p_i}\right)+\gamma\left(\frac{1}{p_i}, \frac{(\mu_i-u)^{p_i}}{p_i\sigma_{p_i}^{p_i}}\right) & u>\mu_i \\ \Gamma\left(\frac{1}{p_i}, \frac{(\mu_i-u)^{p_i}}{p_i\sigma_{p_i}^{p_i}}\right) & u\le\mu_i\end{cases} \quad (8)$$

### Example for GLD

Consider random numbers generated form GLD with parameters $p = [2, 4, 3]$, $\mu = [= -2, 0, 4, ]$ and $\sigma = [0.5, 0.1, 0.9]$ in which its probability density function (pdf) (g, h, f) respectively as shown in Figure 1. In this example we see that the GLD contain (Laplace and Gaussian as special case).

## THE OBJECTIVE FUNCTION

Based on Equation (4) we can obtain family of parametric or score functions by twice differentiable GLD of Equation (7). By substituting Equation (7) into (4) for the source estimates $u_i$, it quickly becomes obvious that our proposed objective function

$$\varphi_i(u_i \mid \mu_i, \sigma_{p_i}, p_i) = sign(u_i - \mu_i)\frac{|u_i - \mu|^{p_i-1}}{\sigma_{p_i}^{p_i}} \quad (9)$$

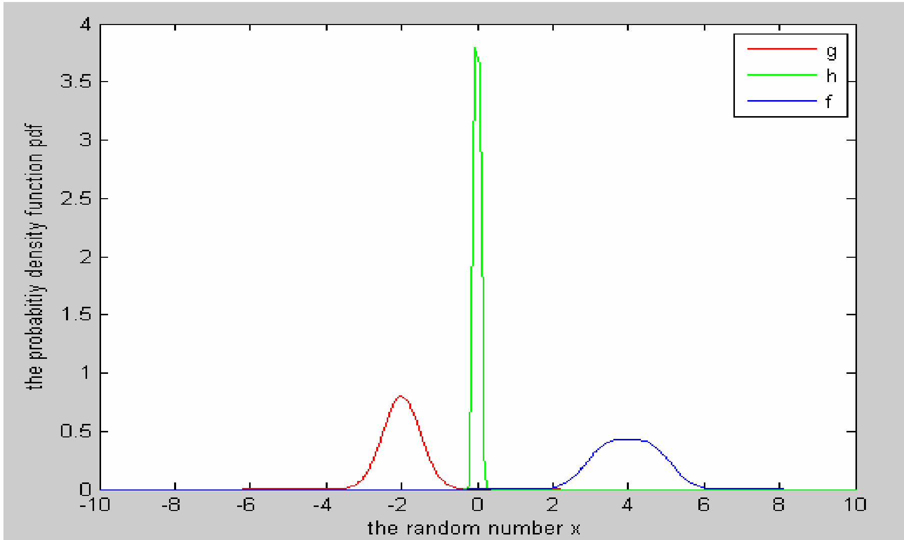


**Figure 1.** Probability density function for GLD with different parameters.

This objective function can be used to modeling large amount of signals such as speech and types of challenging heavy- and light-tailed distributions. We can obtain special case of Equation (9) at $\mu_i = 0$, $p_i = 1$ and $\sigma_i = 1$

$$\varphi_i(u_i) = sign\ (u_i) \tag{10}$$

In which this special case is the standard threshold activation function which is suitable for speech signals or (Laplacian pdf).

## ESTIMATION OF THE GLD PARAMETERS

To refine those further, we can resort to ML. For a sequence of mutually independent Data $\mathbf{u} = [u_1, u_2, ..., u_N]$ of sample size N with density as defined in Equation (7) $g_i(u_i \mid \mu, \sigma_p, p)$ the ML estimates are uniquely defined by their log-likelihood function as

$$L(u \mid \mu, \sigma_p, p) = -\log \prod_{i=1}^{N} g(u_i \mid \mu, \sigma_p, p)$$

$$= -\sum_{i=1}^{N} \log(g(u_i \mid \mu, \sigma_p, p)) \tag{11}$$

Usually, ML parameter estimates are obtained by first differentiating the log-likelihood function in Equation (11) with respect to the GLD parameters and then by equating those derivatives to zero (Shin et al., 2005).

### Estimation of the location and scale parameters

By deriving the log-likelihood function with respect to $\mu$ and $\sigma_p$ and by equalizing the obtained expressions to zero, we have the following equations:

$$\frac{\partial L}{\partial \mu} = -\sum_{i=1}^{N} | u_i - \mu |^{p-1} sign(u_i - \mu) = 0 \tag{12}$$

$$\frac{\partial L}{\partial \sigma} = -N + \frac{1}{\sigma_p^p} \sum_{i=1}^{N} | u_i - \mu |^p = 0 \tag{13}$$

The Equation (12) does not have, in general, an explicit solution and is solved by means of numerical methods, while from Equation (13) we get the maximum likelihood estimator of $,\sigma$ as follow:

$$\hat{\sigma}_p = \left( \frac{\sum_{i=1}^{N} |u_i - \mu|^p}{N} \right)^{1/p} \tag{15}$$

### Estimation of the shape parameter p

The methods presented in literature are based on the likelihood function and on indices of kurtosis.

### *Estimation of p by means of the maximum likelihood method*

If we want to determine the maximum likelihood estimator of the shape parameter p, the equation that we obtain by deriving the log-likelihood function (11) is:

$$\frac{\partial L}{\partial p} = -\frac{N}{p^2}[\log(p) + \Psi(1+1/p) - 1] + \frac{1}{p^2\sigma_p^2}\sum_{i=1}^{N} | u_i - \mu |^p -$$
$$\frac{1}{p\sigma_p^p}[\log(\sigma_p)\sum_{i=1}^{N} | u_i - \mu |^p - \sum_{i=1}^{N} | u_i - \mu |^p \log | u_i - \mu \quad (14)$$

Where $\Psi(.)$ is the digamma function, which is the first derivative of the logarithm of the gamma function. The equation (14) can be solved by using numerical methods. Moreover, Agr`o (1995) uses this method showing that it does not work well for small samples, even though it provides good results for samples of size greater than 50 - 100.

### *Estimation of p by means of indices of kurtosis*

These estimation procedures take into account the relationship between the shape parameter $p$ and the kurtosis. The usually used indices of kurtosis are:

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\Gamma(1/p)\Gamma(5/p)}{[\Gamma(3/p)]^2} \quad (15)$$

$$VI = \frac{\sqrt{\mu_2}}{\mu_1} = \frac{\sqrt{\Gamma(1/p)\Gamma(3/p)}}{\Gamma(2/p)} \quad (16)$$

$$I = \frac{1}{VI} \quad (17)$$

$$\beta_p = \frac{\mu_{2p}}{\mu_p^2} = p+1 \quad (18)$$

Where

$$\mu_r = \sigma_p^p \, p^{r/p} \frac{\Gamma[(r+1)/p]}{\Gamma(1/p)} \quad (19)$$

Is the absolute moment of grade $r$. The index $\beta_p$, called generalized index of kurtosis, the estimators of the indices of kurtosis above described are given by:

$$\hat{\beta}_2 = \frac{n\sum_{i=1}^{n}(u_i - M)^4}{[\sum_{i=1}^{n}(u_i - M)^2]^2} \quad (20)$$

$$V\hat{I} = \frac{\sqrt{n\sum_{i=1}^{n}(u_i - M)^2}}{\sum_{i=1}^{n}| u_i - M |} \quad (21)$$

$$\hat{I} = \frac{1}{V\hat{I}} \quad (22)$$

$$\hat{\beta}_p = \frac{n\sum_{i=1}^{n}| u_i - M |^{2\hat{p}}}{(\sum_{i=1}^{n}| u_i - M |^{\hat{p}})^2} = \hat{p} + 1 \quad (23)$$

Where $M$ is the arithmetic mean.

Alternative method used to maximize the ML equation in (11) to ensure the estimated parameters, this done by resorting to the Nelder-Mead (NM) method direct search method. The appeal of the NM optimization technique lies in the fact that it can minimize the negative of the log-likelihood objective function given in Equation (11), essentially without relying on any derivative information. Despite the danger of unreliable performance (especially in high dimensions), numerical experiments have shown that the NM method can converge to an acceptably accurate solution with substantially fewer function evaluations. Good numerical performance and a significant improvement in computational complexity for our estimation method. Therefore, optimizations with the NM technique produce a good estimation for three parameters in GLD.

## SIMULATIONS

Here, simulation results are shown to verify the performance of the proposed algorithm using Generalized Laplace Distribution as objective (cost) function in which we used NM to estimate the parameters before using FastICA.

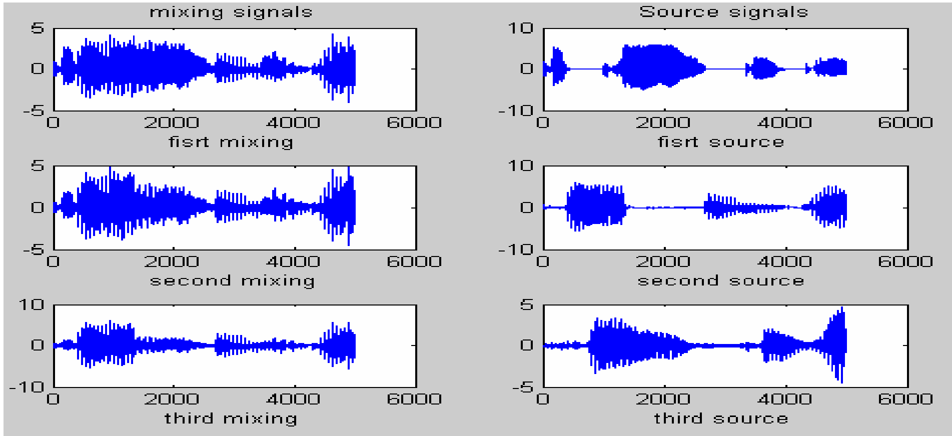


**Figure 2.** The mixing signals in left and original signals in right.

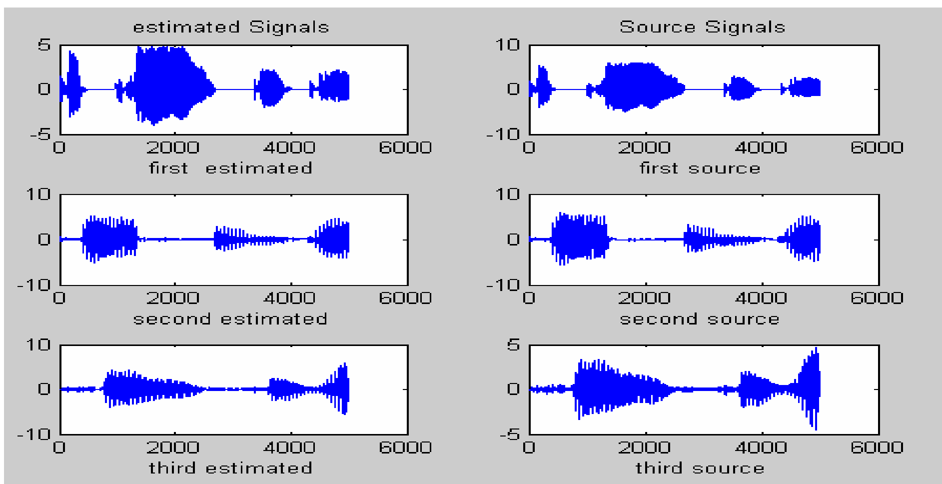


**Figure 3.** The estimated signals in left with scales, permutation and original signals in right.

**Example 1**

Consider three speech signals as sources, mixing matrix A and demixing matrix W are given as follow

$$\mathbf{A}=\begin{pmatrix} .56 & .79 & .37 \\ .75 & .65 & .56 \\ .17 & .32 & .48 \end{pmatrix} \text{ And } \mathbf{W}=\begin{pmatrix} .0109 & .0340 & .260 \\ .0024 & .0467 & .0415 \\ .0339 & .0192 & .0017 \end{pmatrix}$$

By using the equation $\mathbf{x}=\mathbf{A}\,\mathbf{s}$ we obtain mixed signals as shown in Figure (2) where mixing signals in left and original signals in right. We recover the source by using FastICA and we show the estimated signals in left with scales, permutation and original signals in right in Figure (3).

**Example 2**

Consider a three sources in which they are random number from GLD but with parameters $p=[2,1,6]$, $\mu=[=\text{-}2,0,4,]$ and $\sigma=[0.5,0.1,0.9]$

$$\mathbf{s}_1 = \mathrm{GLDrnd}(\mu(1), \sigma(1), p(1), n),$$
$$\mathbf{s}_2 = \mathrm{GLDrnd}(\mu(2), \sigma(2), p(2), n),$$
$$\mathbf{s}_3 = \mathrm{GLDrnd}(\mu(3), \sigma(3), p(3), n),$$

At n=600

$$\mathbf{A}=\begin{pmatrix} .50 & .6 & .37 \\ .35 & .65 & .60 \\ .25 & .92 & .5 \end{pmatrix} \text{ And } \mathbf{W}=\begin{pmatrix} 2.55 & \text{-}1.40 & .\text{-}0.84 \\ 2.03 & \text{-}9.91 & 9.80 \\ \text{-}.4333 & 0.84 & \text{-}0.312 \end{pmatrix}$$

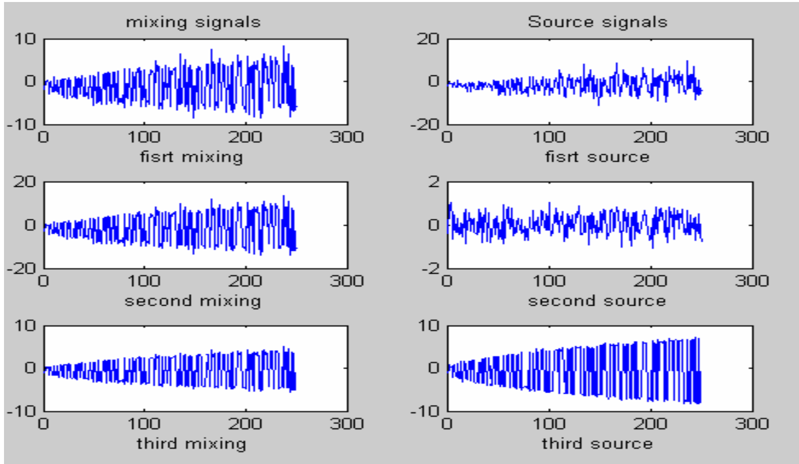


**Figure 4.** The mixing signals in left and original signals in right.

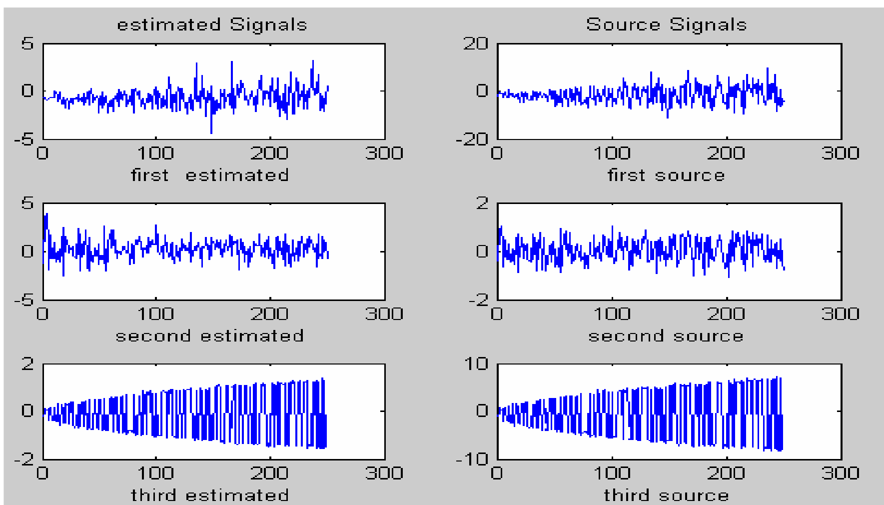


**Figure 5.** The estimated signals in left with scales, permutation and original signals in right**.**

Based on the equation of mixed X = AS we obtain mixed signals as shown in Figure (4) where mixing signals in left and source signals in right. We obtain the source signals by using FastICA and we show the estimated signals in left and original signals in right in Figure (5).

## Conclusions

This paper introduces a new family of score functions based on Generalized Laplace Distribution for BSS in which this family contain Laplace and Gaussian distributions as special cases. To estimate the parameters of these functions, we have chosen to maximize the ML equation with the NM optimization method as alternative method to derive the ML equation. To blindly extract the source signals we resort to FastICA. Simulation results show that the proposed approach is capable of separating mixtures of signals.

## REFERENCES

Agr 'o G (1995). Maximum likelihood estimation for the exponential power function parameters. Comm. Stat. Simul. Comput., 24(2): 523-536.

Box GEP, Tiao GC (1962). A further look at robustness via Bayess theorem. Biometrika, 49: 419-432.

Cardoso JF (1998). Blind signal separation: Statistical principles. Proc. IEEE, 86(10): 2009-2025.

Hyvarinen A, Oja E (1997). A fast fixed-point algorithm for independent component analysis. Computation, 9(7): 1483-1492.

Karvanen J, Koivunen V (2002). Blind separation methods based on Pearson system and its extensions. Signal Process., 82(4): 663-673.

Kokkinakis K, Nandi AK (2006). Flexible score functions for blind separation of speech signals based on generalized gamma probability density functions. Proc. IEEE Int. Conf. Acoustics Speech Signal Process. Toulouse, France, May 1419, 1: 1217-1220.

Shin JW, Chang JH, Kim NS (2005). Statistical modeling of speech signals based on generalized gamma distribution. IEEE Signal Process. Lett., 12(3): 258-261.
Subbotin MT (1923). On the law of frequency of errors. Mat. Sb., 31: 296-300.
Tiao GC, Lund DR (1970). The use of OLUMV estimators in inference robustness studies of the location parameters of a class of symmetric distributions. J. Am. Stat. Assoc., 65: 371-386.
Zeckhauser R, Thompson M (1970). Linear regression with non-normal error terms. Rev. Econ. Stat., 52: 280-286.