BLIND SOURCE SEPARATION OF NOISY MIXTURES USING A SEMI-PARAMETRIC APPROACH WITH APPLICATION TO HEAVY-TAILED SIGNALS

M. Sahmoudi[1]*, K. Abed-Meraim*[1]*, M. Lavielle*[2]*, E. Kuhn* [2] *and Ph. Ciblat*[1]

[1]TSI, Telecom-Paris, 37-39, rue Dareau, 75014 Paris, France
[2] Laboratoire de Mathématiques, Paris-Sud University, France

ABSTRACT

In this paper, we propose a new semi-parametric approach for blind source separation (BSS) of noisy mixtures with application to heavy-tailed signals. The semi-parametric statistical principle is used to formulate the BSS problem as a maximum likelihood (ML) estimation. More precisely, this approach consists of combining the logspline model for sources density approximation with a stochastic version of the EM algorithm for mixing matrix estimation. The proposed method is truly blind to the particular underlying distribution of the mixed signals and performs simultaneously the estimation of the unknown probability density functions (pdf) of the source signals and the estimation of the mixing matrix. The application of logspline density approximation also enables the algorithm to be robust to modelization errors of the sources. In addition, it is robust against outliers and impulsive effect. Computer simulations are provided to illustrate the effectiveness of the proposed separation method comparatively with classical ones.

1. INTRODUCTION

Blind source separation is one of the most attractive research topics nowadays in the field of signal processing and its applications [7]. The noisy case has not received much attention in the BSS literature, maybe because it was felt that dealing explicitly with noise is useless in a high SNR context and hopeless in a low SNR context while in not-so-bad SNR situations, processing noisy data using noise-free models yields good enough results. Another reason maybe that, in the standard approaches to BSS, the sources are modeled as i.i.d. sequences, with the effect that including noise in the model very much changes the structure of the estimation problem and makes it more difficult to tackle.

• **Noisy linear instantaneous mixtures.** In this paper, we consider the classical noisy linear BSS model with instantaneous mixtures given by:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \boldsymbol{\varepsilon}(t), \;\; t = 1 \dots T \tag{1}$$

where $\mathbf{A}$ is a $n \times m$ *unknown* full column rank mixing matrix. The sources $s_1(t), \cdots, s_m(t)$ are collected in a $m \times 1$ vector denoted $\mathbf{s}(t)$ and are assumed to be i.i.d. signals: the joint distribution density π is factorized as $\pi = \prod_{j=1}^{m} \pi_{s_j}$. The noise vector $\boldsymbol{\varepsilon}(t)$ (independent with $\mathbf{s}(t)$) has independent components $\varepsilon_1(t), \dots, \varepsilon_n(t)$ with zero mean and unknown variance σ^2. The goal of a BSS method is to find a *separating matrix* i.e. an $m \times n$ matrix $\mathbf{B}$ such that the recovered sources $\mathbf{B}\mathbf{x}(t)$ are as independent as possible. In the noiseless case, (1) admits a unique solution up to scaling and permutation indeterminacy $\mathbf{y}(t) = \mathbf{B}\mathbf{x}(t)$ such that $\mathbf{C} \triangleq \mathbf{B}\mathbf{A} = \mathbf{P}\boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is a diagonal scaling matrix and $\mathbf{P}$ is a permutation matrix (see [7]). At most one source is allowed to be Gaussian to ensure the identifiability. Another problem is that if one or more sources do not have finite second or higher moments (e.g. heavy-tailed distributions) then prewhitening or criteria optimization would cause a breakdown [6, 9, 10].

• **Heavy-tailed α-stable sources.** The popular class of symmetric α-stable (SαS) statistical model of heavy-tailed signals has been proposed for signal processing applications [1]. An $S\alpha S$ distribution is best defined by its characteristic function $\varphi(\omega) = \exp\left(j\mu\omega - \gamma|\omega|^{\alpha}\right)$, where $\alpha \in]0; 2]$ is the *characteristic exponent* that determines the shape of the distribution. The smaller α is, the heavier the tails of the α-stable density. $\mu \in \mathbb{R}$ is the *location parameter*, and $\gamma > 0$ is the *dispersion* index. No closed-form expressions exist for α-stable density other than the cases of $\alpha = 2$ (Gaussian distribution), $\alpha = 1$ (Cauchy distribution) and of $\alpha = 1/2$ (Levy distribution). Alpha-stable densities obey three important properties which further justify their role in data modeling [1]:

1) ***Stability***: A weighted sum of independent α-stable random variables is α-stable with the same α.

2) ***Generalized central limit theorem (GCLT)***: Without limitation of finite variance, stable models are the *only* distribution that can be the limit in distributions of i.i.d. random variables.

3) ***Heavy-tailed asymptotic behavior***: Let X be an α-stable r.v. with $\alpha < 2$. Then : $P(X > x) \sim \gamma C_\alpha x^{-\alpha}$ as $x \to \infty$ where C_α is a positive constant depending only on α.

An important consequence of this latter property is the *non-existence* of the second and higher order moments of stable distributions. For this reason, most classical BSS methods are inadequate in this context and divergence behaviors may be observed [6, 9, 10].

• **Maximum-likelihood source separation.** To apply the ML principle, it is not difficult to derive the likelihood function using a parametric model of source densities. However, the distribution model mismatch between the output pdf and the chosen underlying distribution model is a serious problem in such approaches. Incorrect assumptions on the source distributions can result in poor estimation performance or in a complete failure to achieve the source separation [7]. Alternative methods that employ a non-parametric density estimation have been introduced (e.g. [4]). These methods usually consist in a density estimation technique that alternates with a cost function optimization step in an iterative approximation framework. Although these approaches do not require the definition of a specific model for the density functions, their capability

This work was in part supported by the CNRS within a MathSTIC project entitled 'Estimation and detection techniques for impulsive signals'.

of separation arbitrarily distributed sources, not have been fully assessed. Thus, finding a compromise between performances and the robustness to the source's pdf mismatch in a blind signal separation framework is still an open and challenging problem. In this paper, we propose a new semi-parametric BSS method using the ML approach and an approximation of the sources densities by the logspline model in order to avoid any assumption of the source distribution. This method can be applied in particular to heavy-tailed signals, for which few algorithms exist [6, 9, 10]. This estimate exhibits better performance than standard BSS methods in a variety of simulation contexts.

2. SEMI-PARAMETRIC BSS

Any BSS problem can be seen as an usual missing data problem. Indeed, the observed data are the observations $\{\mathbf{x}(t)\}_{1\leq t\leq T}$, whereas the random sources $\{\mathbf{s}(t)\}_{1\leq t\leq T}$ are the unobserved data. Then, the complete data of the model is $\{\mathbf{x}(t),\mathbf{s}(t)\}_{1\leq t\leq T}$. We suppose that the unobserved sources are related to the observations through the density functions h of $\mathbf{x}$ conditionally to $\mathbf{s}$ [1]. Our purpose is to estimate the sources density $\pi=\prod_{j=1}^{m}\pi_{s_j}$, the mixing matrix $\mathbf{A}$ and the noise-variance σ^2. For that we propose a semi-parametric approach which consists of combining the logspline model for sources density approximation with a stochastic version of the EM algorithm. We use logspline models for two reasons: on one hand, they have good functional approximation properties, on the other hand, they are well-adapted to the implementation of the SAEM (Stochastic Approximation version of the Expectation Maximization) algorithm [8] allowing to compute easily our estimator. Indeed, the first assumption of the used SAEM algorithm is equivalent to suppose that the complete data likelihood $f(\mathbf{x},\mathbf{s},\boldsymbol{\eta})$ belongs to the curved *exponential family* and can be written:

$$f(\mathbf{x},\mathbf{s},\boldsymbol{\eta})=\exp\left\{-\Psi(\boldsymbol{\eta})+\langle\tilde{S}(\mathbf{x},\mathbf{s}),\Phi(\boldsymbol{\eta})\rangle\right\} \qquad (2)$$

where $\langle .,.\rangle$ denotes the scalar product, $\boldsymbol{\eta}$ denotes the unknown global parameters vector to be estimated and $\tilde{S}(\mathbf{x},\mathbf{s})$ is known as the minimal sufficient statistics (MSS) of the complete model. In this case of unknown density functions following model (2), a good approximation which satisfies this latter condition is given by the logspline model. Moreover, it is was shown that this estimation technique is inherently robust against outliers and impulsiveness effects [11]. For this reason, we apply this method to impulsive random variables with possibly *heavy-tailed distributions* characterized by infinite second and higher order moments. We define now precisely the logspline model which will be used.

2.1. Estimating a density function by B-spline approximations

In order to get a non parametric estimate of the source density function π, we propose to use the logspline model. Let $\mathcal{I}$ be equal to $[a,b]$ where $-\infty<a<b<+\infty$ and consider a given knots sequence $\tau=(t_l)_{1\leq l\leq K+1}$ with $a=t_1$ and $b=t_{K+1}$. Consider now the space $\mathcal{S}_{q,\tau}$ of spline functions of positive order q on $\mathcal{I}$, namely piecewise polynomial functions of degree $q-1$ associated to this knots sequence. Then the dimension of $\mathcal{S}_{q,\tau}$ is equal to $J=q+K-1$ and there exists a B-splines basis denoted $B_1,\cdots,B_J$ for $\mathcal{S}_{q,\tau}$ [3]. The logspline density estimation method models a log-density function as a spline function:

$$\forall s\in\mathcal{I},\quad \pi_\theta(s)=\exp\left[\sum_{j=1}^{J}\theta_j B_j(s)-c(\theta)\right] \qquad (3)$$

$$\text{where } c(\theta)=\log\left(\int_{\mathcal{I}}\exp\left[\sum_{j=1}^{J}\theta_j B_j(s)\right]ds\right)$$

is a normalization factor and $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_J)\in\mathbb{R}^J$. We choose the dimension J of the logspline model in function of the sample size T such that $J=o(\sqrt{T})$ (see [8] for more details). We define now the observed log-likelihood corresponding to the logspline model of the observations defined as follow:

$$\mathcal{L}_T(\theta)=\frac{1}{T}\sum_{t=1}^{T}\log\int_{\mathcal{I}}h\left(\mathbf{x}(t)|s\right)\pi_\theta(s)ds \qquad (4)$$

Then we consider the maximum likelihood estimator $\pi_{\hat{\theta}_{T,J}}$ of the density π in the logspline model given by:

$$\hat{\theta}_{T,J}=\arg\max_{\theta\in\Theta_J}\mathcal{L}_T(\theta) \qquad (5)$$

This family is not identifiable since we have for all a real: $c(\theta+a)=c(\theta)+a$ implying that $\pi_{\theta+a}=\pi_\theta$. We set systematically $\theta_J=0$ in order to get an identifiable family of log-density functions and we denote Θ_J the subspace of $\mathbb{R}^J$ composed of vectors having zero as last coordinate and $\mathcal{M}_{q,\tau}$ the set of associated densities, i.e. $\{\pi_\theta,\theta\in\Theta_J\}$. We describe briefly some properties of the B-splines detailed in de Boor's book [3]:
1) ***B-spline***: For all $1\leq j\leq J$, the function B_j takes values in the interval $[0,1]$. Moreover, we have $\sum_{j=1}^{J}B_j(s)=1\ \forall s\in\mathcal{I}$.
2) ***Approximation property of the logspline model***: We define $\delta_J=\inf_{\theta\in\Theta_J}\|\log f-\log\pi_\theta\|_\infty$. For some positive continuous density function f on $\mathcal{I}$, δ_J tends to zero when J goes to infinity. See [3] for more details on the links between the convergence rate and the regularity of f. The particular properties of the logspline model let us think that $\pi_{\hat{\theta}_{T,J}}$ will have remarkable properties when T tend to infinity. In a first time, we explain how we compute this estimator in practice simultaneously with the mixing matrix and the noise variance.

2.2. The SAEM algorithm

To compute the unknown parameters $\boldsymbol{\eta}=(\boldsymbol{\theta}^T,vec(\mathbf{A})^T,\sigma^2)^T$, we use the SAEM algorithm coupled with a MCMC (Markov Chain Monte-Carlo) procedure presented in [8]. Here we apply this algorithm for estimating the mixing matrix $\mathbf{A}$ and the variance σ^2 using the logspline model to approach the estimate $\pi_{\hat{\theta}_{T,J}}$. The complete log-likelihood corresponding to the logspline model has the following expression:

$$\mathcal{L}_T^{com}(\boldsymbol{\eta})=\frac{1}{T}\sum_{t=1}^{T}\log h\left(\mathbf{x}(t)|\mathbf{s}(t)\right)+\frac{1}{T}\sum_{t=1}^{T}\log\pi_\theta\left(\mathbf{s}(t)\right) \qquad (6)$$

So we apply the SAEM algorithm to this parametric model in order to approach the estimator $\hat{\boldsymbol{\eta}}_{T,J}$ of $\boldsymbol{\eta}$, that maximizes the observed log-likelihood. To put out the minimal sufficient statistics of the

[1] The distribution of $\mathbf{x}$ conditionally to $\mathbf{s}$, denoted by h, corresponds in the fact to the distribution of the additive noise in the BSS model (1) with the same variance and a non-zero mean value equal to $\mathbf{As}$.

model, we write the developed expression of the complete log-likelihood:

$$\mathcal{L}_T^{com}(\boldsymbol{\eta}) = \frac{1}{T}\sum_{i=1}^{T} \log h\left(\mathbf{x}(t)|\mathbf{s}(t)\right) + \sum_{j=1}^{J} \theta_j \left[\frac{1}{T}\sum_{t=1}^{T} B_j(\mathbf{s}(t))\right] - c(\theta)$$

We choose as MSS $\tilde{S}(\mathbf{x},\mathbf{s}) = (\frac{1}{T}\sum_{i=1}^{T} B_j(s_i), 1 \leq j \leq J)$ and we implement the k-th iteration of the SAEM algorithm as:

- **S-step**: Generate a realization $\mathbf{s}'$ using as proposal distribution the prior distribution π_{θ_k} and take $\mathbf{s}_k$ equal to $\mathbf{s}'$ or to $\mathbf{s}_{k-1}$ according to the value of the acceptance probability.
- **A-step**: Update the minimal sufficient statistics $\tilde{S}_k$ according to the stochastic approximation:

$$\tilde{S}_k = \tilde{S}_{k-1} + \beta_{k-1} \quad \tilde{S}(\mathbf{x},\mathbf{s}_k) - \tilde{S}_{k-1} \tag{7}$$

  where β_k is a positive step-sizes sequence decreasing to 0.
- **M-step**: Update $\boldsymbol{\eta}_k$ by maximizing the complete log-likelihood of the model evaluated in the observations and in the current value of the minimal sufficient statistics.

This algorithm converges a.s. toward a local maximum of the log-likelihood of the observations under very general regularity conditions (see [8] for convergence results). In practice, the algorithm is easy to implement and has a relatively low computational cost.

3. PERFORMANCE EVALUATION & COMPARISON

3.1. Some existing BSS methods

We briefly describe here three BSS approaches for comparison with the new semi-parametric approach introduced above.

1) ***FastICA algorithm*** **[7].** Under the whitened zero-mean demixing model $\mathbf{y} = \mathbf{W}\mathbf{z}$, the FastICA algorithm finds the extrema of a generic cost function $\mathbb{E}\{G(\mathbf{w}^T\mathbf{z}\}$, where $\mathbf{w}^T$ is one of the rows of the demixing matrix $\mathbf{W}$. The cost function can be e.g. a normalized cumulant or an approximation of the marginal entropy which is minimized in order to find maximally nongaussian projections $\mathbf{w}^T\mathbf{z}$. This algorithm is facing three problems. First, some sources may not have zero means in which case the mean values must be explicitly included in the analysis. Second, in FastICA, the derivative of the even function G is assumed to be an odd function. If this condition fails to be satisfied, the FastICA as such may not work. Third, FastICA is not robust to heavy-tailed effect.

2) ***JADE algorithm*** **[5].** This algorithm operates on cumulants as a measure of independence. It seeks to approach independence through the maximizing of the higher order cumulants. However, one major weakness of this algorithm is that higher order cumulants are extremely vulnerable to outlier effects. Besides being sensitive to outliers, JADE also fails to separate certain source distribution, i.e. skewed zero-kurtotic signals generated by the power distribution. This is because by minimizing only the 4-th order cumulants, third order effects like the skewness are ignored.

3) ***Minimum Dispersion (MD) algorithm*** **[10].** This approach is a two-step parametric algorithm for heavy-tailed source separation.

Step 1: Robust whitening. In the case of α-stable signals, it is proven in [9] that the normalized covariance matrix of $\mathbf{x}$ defined by $\hat{\mathbf{R}}_x^n = \frac{\hat{\mathbf{R}}_x}{Trace(\hat{\mathbf{R}}_x)}$ with $\hat{\mathbf{R}}_x = \frac{1}{T}\sum_t \mathbf{x}(t)\mathbf{x}(t)^T$ converges asymptotically (i.e. when T tends to infinity) to the finite matrix $\mathbf{A}\mathbf{D}\mathbf{A}^T$, where $\mathbf{D}$ is a positive diagonal matrix. Hence, the normalized covariance matrix has the appropriate structure and the whitening problem becomes standard.

Step 2: MD criterion. Let $\mathbf{z}(t) = \mathbf{B}\overline{\mathbf{x}}(t)$ where $\mathbf{B}$ is an orthogonal separating matrix to be estimated and $\overline{\mathbf{x}}$ denotes the whitened data. It is shown in [9] that under orthogonality constraint, the MD criterion given by $J(\mathbf{B}) = \sum_{i=1}^{m} \gamma_{z_i}$, where γ_{z_i} denotes the dispersion of $z_i(t)$ the i-th entry of $\mathbf{z}(t)$, is a contrast function.

The essential limitation of this method is that it can be used only for heavy-tailed sources with α-stable distribution.

3.2. Parametric versus semi-parametric approaches

The MD method is said to be parametric in the sense that it relies on the a priori knowledge of the exact source pdf. In this case, we have a finite set of parameters to estimate. On the other hand, the SAEM method is said to be semi-parametric in the sense that the source pdf is unknown and need to be jointly estimated with the desired parameters (i.e. mixing matrix) [2]. Clearly, estimating a pdf is a difficult problem as the number of parameters to be estimated is infinite. In the semi-parametric approach, we estimate a limited number of parameters by replacing the estimation problem by an approximation one. The parametric approach is preferred whenever a reliable a priori knowledge on the source pdf is available. In the situations where the pdf is only partially or inaccurately known, semi-parametric methods should be used because of their robustness against modelization errors as shown next by simulation results.

3.3. Computer simulation experiments

Here, we compare our proposed semi-parametric method SAEM to JADE, FastICA and to the parametric MD algorithm. In all simulation experiments the results are averaged over 100 iterations and the mixing matrix $\mathbf{A}$ is generated randomly at each run. The stepsize sequence (β_k) used for SAEM was $\beta_k = 1/k$. For the choice of the size J of the logspline model in SAEM, we have tested some values for J lower than 10 since we have at least 100 observations. The best estimation seems to be given for $q = 4$ and $J = 5$, so we will hold these values for the following experiments. We choose as initial value θ_0, such that the logspline density estimate is initialized with the uniform distribution on $\mathcal{I} = [-50, 50]$.

To measure the quality of separation we will use Amari's error criterion as a performance index (PI) defined as

$$PI = \sum_{i=1}^{m}\left(\sum_{j=1}^{m}\frac{|C_{i,j}|}{\max_k |C_{i,k}|} - 1\right) + \sum_{j=1}^{m}\sum_{i=1}^{m}\frac{|C_{i,j}|}{\max_k |C_{k,j}|} - 1\Bigg)$$

where $\mathbf{C} = (C_{i,j})_{1\leq i,j\leq m} = \mathbf{B}\mathbf{A}$ is the global system.

• **Experiment 1: Robustness against outliers.** First, we test the robustness against outliers. We mix two sources, one of Gaussian distribution and the second of uniform distribution with randomly chosen mixing matrices. The data set contains 1000 points. Without outliers, the performances of SAEM, JADE and FastICA are all excellent (PI ≈ 0.05). To test for outlier-robustness, we replace 50 data point with outliers, i.e. uniformly distributed data points within a disc of radius 500 around the origin (the norm of the original data points is roughly within the range from 0 to 100). As expected, SAEM still works fine. In fact, typically it does not even change its solution, because it simply ignores the outliers in the B-spline adjustment stage. JADE and FastICA however, produce arbitrary results because they employ higher-order statistics which are highly sensitive to outliers.

• **Experiment 2: Asymptotic consistency.** Figure 1 shows some simulation results in case of noiseless three mixtures ($n = 3$ observations) of three sources ($m = 3$) with, respectively, a uniform distribution on $[0, 1]$, a Gaussian distribution with zero mean and unit variance and standard $S\alpha S$ with $\alpha = 1.5$. To detect whether BSS algorithms can obtain consistent estimates in such situation, the sample size was increased from (1): $T = 1000$ to (2): $T = 5000$. We compare SAEM and two other famous BSS algorithms, JADE and FastICA. Similarly to [6], we present the boxplots based on quartils to assess the consistency of our method.

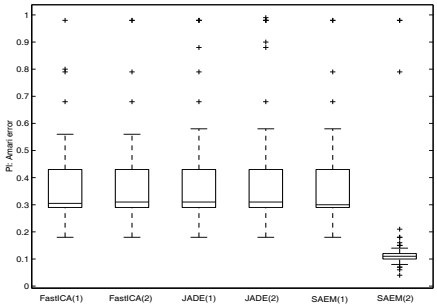

Fig. 1. Consistency of different BSS algorithms. The sample sizes were 1000 for case (1) and 5000 for case (2).

From the boxplots, we can see that as the sample size increases, the estimation error (PI) for SAEM decreases more significantly toward zero than for JADE and FastICA.

• **Experiment 3: Robustness against impulsive noise.** In this experiment we add impulsive noise to the above mixtures (considered in the experiment 2) according to $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \sigma\epsilon(t)$ with $\epsilon(t)$ being a n-dimensional Gaussian noise of unit variance. We track the evolution of the performance index as a function of the noise level σ for kurtotic (super-Gaussian) noise: we used multidimensional Gaussian noise, where we change the absolute value to the power of 5.

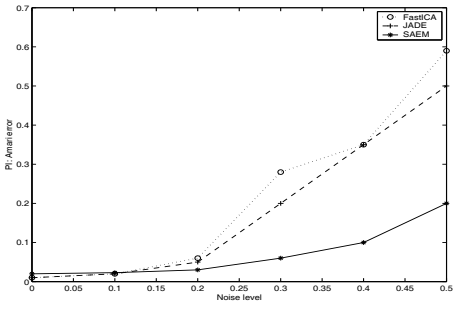

Fig. 2. The performance index versus noise level.

Figure 2 shows that JADE and FastICA start to fail at a certain noise level, whereas SAEM continues to produce good BSS solutions. Note that we have chosen the median over 100 runs because the PI depend strongly on the actual realization of the noise.

• **Experiment 4: Robustness against error modelization.** Here, we consider $m = 3$ impulsive sources with generalized gaussian distribution of parameter $p = 1.5$ (i.e. the source pdf is proportional to $\exp(-|x|^p)$). In that case, the signals are of finite variances and $n = 4$ noise free mixtures are considered.

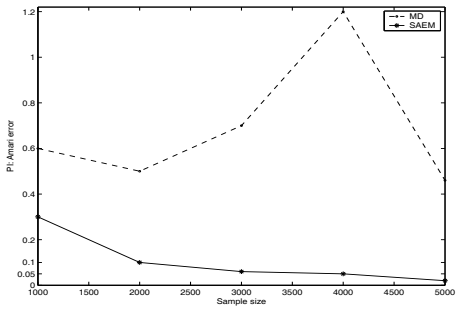

Fig. 3. The performance index versus sample size.

As can be observed from figure 3, the MD method fails to separate correctly the sources as it relies on the SαS source pdf assumption that is not verified in this example. This illustrate the robustness of the SAEM compared to the MD method with respect to the pdf modelization errors.

4. CONCLUSION

In this work, we developed a new semi-parametric BSS method using the SAEM algorithm. The proposed method is applied for the blind separation of noisy linear instantaneous mixtures of possibly heavy-tailed sources. The SAEM based method is compared with the JADE, FastICA and the minimum dispersion (MD) methods and shown to be more general (as it can be applied to a larger class of source signals and in different scenario). The proposed SAEM algorithm outperforms JADE and FastICA in terms of consistency and robustness against the outliers and impulsive noise and outperforms the MD method in terms of robustness against modelization errors.

5. REFERENCES

[1] R.J. Adler, R. Feldman and M.S. Taqqu, Eds *A practical Guide to Heavy tails : Statistical Techniques and Applications*, Birkhauser, Boston, 1998.

[2] P. Bickel, et al. *Efficient and Adaptive Estimation for Semiparametric Models*, Springer, 1998.

[3] C. de Boor, *A practical guide to splines*, Applied Mathematical Sciences, Springer-Verlag, New York, 1978.

[4] R. Boscolo et al, "Independent Component Analysis Based on Nonparametric Density Estimation", *IEEE Trans. on Neural Networks*, vol. 15, No. 1, Jan. 2004.

[5] J.F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals", *Radar and Sig. Proc., IEE Proc. F*, Dec. 1993.

[6] A. Chen and P.J. Bickel, "Robustness of Prewhitening Against Heavy-Tailed Sources", *In Proc. of ICA 2004*, Granada, Spain, 2004.

[7] A. Hyvarinen, J. Karhunen and E. Oja, *Independent Component Analysis* , Wiley, 2001.

[8] E. Kuhn, and M. Lavielle, "Coupling a stochastic approximation version of EM with a MCMC procedure", *ESAIM P&S, France*, Vol. 8, PP. 115-131, Aug. 2004.

[9] M. Sahmoudi, K. Abed-Meraim and M. Benidir "Blind Separation of Heavy-Tailed Signals Using Normalized Statistics". *In Proc. of ICA 2004*, Granada, Spain, Sept. 2004.

[10] M. Sahmoudi, K. Abed-Meraim and M. Benidir, "Blind Separation of Impulsive α-stable Sources Using Minimum Dispersion Criterion" *in IEEE Signal Processing Letters*, Vol. 12, pp: 281-284, April 2005.

[11] Takada, Teruko, "Nonparametric density estimation: A comparative study." Economics Bulletin, Vol. 3, No. 16 pp. 1-10, 2001.