# Blind Separation of Impulsive Alpha-Stable Sources Using Minimum Dispersion Criterion

Mohamed Sahmoudi[1], Karim Abed-Meraim[2] and Messaoud Benidir[1].

***Abstract*— This paper introduces a novel Blind Source Separation (BSS) approach for extracting impulsive signals from their observed mixtures. The impulsive signals are modeled as real-valued symmetric alpha-stable $(S\alpha S)$ processes characterized by infinite second and higher order moments. The proposed approach uses the minimum dispersion (MD) criterion as a measure of sparseness and independence of the data. A new whitening procedure by a normalized covariance matrix is introduced. We show that the proposed method is robust, so-named for the property of being insensitive to possible variations in the underlying form of sampling distribution. Algorithm derivation, and simulation results are provided to illustrate the good performance of the proposed approach. The new method has been compared with three of the most popular BSS algorithms; JADE, EASI and Restricted Quasi-Maximum Likelihood (RQML).**

***Index Terms*— $\alpha$-Stable distribution, BSS, Normalized covariance, Minimum dispersion criterion, Robustness.**

## I. Introduction

Heavy-tailed distributions, largely used to model impulsive signals, assign relatively high probabilities to the occurrence of large deviations from the median. A common characteristic property of many heavy-tailed distributions, such as the $\alpha-$stable family, is the nonexistence of finite second or higher order moments. There are several well-known methods for BSS [1], [2], based in general on second or higher-order statistics of the observations and so are inadequate to handle heavy-tailed sources. In that case, fractional lower-order theory is used instead [7]. Only a limited literature was dedicated to BSS of impulsive signals. In [12], the RQML approach is introduced as an extension of the popular Pham's quasi-maximum likelihood approach to the $\alpha$-stable sources case. Other solutions exist in the literature based on the spectral measure [6], the normalized statistics [10] and the order statistics [12]. In this paper, we introduce a new method for $\alpha$-stable source separation from their observed linear mixtures using the minimum dispersion criterion [9]. In the finite variance case, a similar approach for principal components analysis (PCA) that uses the output variances has been proposed in [3].

### A. Why heavy-tailed $\alpha$-stable distributions

The stable distribution is a very flexible modeling tool in that it has a parameter $\alpha$ $(0 < \alpha \leq 2)$, called the *characteristic exponent*, that controls the heaviness of its tails. A small positive value of $\alpha$ indicates severe impulsiveness, while a value of $\alpha$ close to 2 indicates a more Gaussian type of behavior. Stable distributions obey the *Generalized Central Limit Theorem (GCLT)* which states that if the sum of i.i.d random variables with or without finite variance converges to a distribution by increasing the number of variables, the limit distribution must be stable [7]. Thus, non-Gaussian stable distributions arise as sums of random variables in the same way as the Gaussian distribution. Another defining feature of the stable distribution is the so-called *stability property*, which says that the sum of two independent stable random variables with the same characteristic exponent is again stable and has the same characteristic exponent. For these reasons, statisticians [7], economists [8] and other scientists engaged in a variety of disciplines have embraced alpha-stable processes as the model of choice for heavy-tailed data.

### B. The symmetric $\alpha-$stable distributions

No closed form exist for $\alpha$-stable probability density function (pdf) except for the cases $\alpha = 1/2$ (Levy distribution), $\alpha = 1$ (Cauchy distribution) and $\alpha = 2$ (Gaussian distribution) and is best defined by its characteristic function [7].

*Definition 1:* A random variable (RV) is said to have a symmetric alpha-stable distribution $S\alpha S(\gamma,\mu)$ if its characteristic function is of the form $\varphi(t) = \exp\{j\mu t - \gamma \mid t \mid^{\alpha}\}$ where $0 < \alpha \leq 2$ is the *characteristic exponent*, which measures the thickness of the tails of the distribution, $\mu \in \mathbb{R}$ is the *location parameter* and $\gamma(\gamma > 0)$ is the *dispersion* of the distribution. The dispersion parameter $\gamma$ determines the spread of the distribution around its location parameter $\mu$.

The following properties of $S\alpha S(\gamma,\mu)$ laws will be used next for BSS of $\alpha$-stable signals [7].

*Property 1:* Let $X_1$ and $X_2$ be independent RVs with $X_i \sim S\alpha S(\gamma_i,\mu_i), i = 1,2$. Then, $X_1 + X_2 \sim S\alpha S(\gamma,\mu)$ with $\gamma = \gamma_1 + \gamma_2$ and $\mu = \mu_1 + \mu_2$.

*Property 2:* Let $X \sim S\alpha S(\gamma,\mu)$ and $h$ be a real constant. Then, $hX \sim S(\mid h \mid^{\alpha} \gamma, h\mu)$.

*Property 3:* If $X \sim S\alpha S(\gamma,\mu)$ and $\alpha \neq 2$, then $\lim_{t\to\infty} t^{\alpha} P(\mid X \mid > t) = \gamma C_{\alpha}$ where $C_{\alpha}$ is a constant that depends on $\alpha$ only.

A direct consequence of this property is that for $S\alpha S$ RV, $p$th moments are finite if and only if $p < \alpha$.

*Property 4:* The *fractional lower order moments (FLOMs)* of an $\alpha$-stable random variable with zero location parameter and dispersion $\gamma$ are given by $E|X|^p = C(p,\alpha)\gamma^{\frac{p}{\alpha}}$ for $0 < p < \alpha$ where $E(.)$ denotes the expectation operator and $C(p,\alpha)$ is a constant depending only on $p$ and $\alpha$.

[1] Are with the LSS, CNRS-SUPELEC-UPS, 91192 Gif-sur-Yvette, France (E-mail: sahmoudi, benidir@lss.supelec.fr)

[2] Is with TSI, Telecom Paris, 75634 Paris, France (E-mail: abed@tsi.enst.fr)

### C. Problem Formulation

In many situations of practical interest, one has to consider $m$ mutually independent signals whose $n \geq m$ linear combinations are observed. They are formulated as $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$, where $\mathbf{s}(t) = [s_1(t), \cdots, s_m(t)]^T$ is the real valued *impulsive source vector* and $\mathbf{A}$ is a $n \times m$ full rank *mixing matrix*. The source signals $s_i(t), i = 1, \cdots, m$ are assumed to be mutually independent, $S\alpha S(0, \gamma_{s_i})$ processes with the same characteristic exponent $\alpha$ [1]. The purpose of blind source separation is to find a separating matrix, i.e., an $m \times n$ matrix $\mathbf{B}$ such that $\mathbf{z}(t) = \mathbf{B}\mathbf{x}(t)$ is an estimate of the source signals up to a permutation and scaling factors [1].

## II. Source Separation

### A. Whitening by normalized covariance matrix

The first step consists in whitening the observations (orthogonalizing the mixture matrix $\mathbf{A}$). For finite variance signals, the whitening matrix $\mathbf{W}$ is computed as the inverse square root of the signal covariance matrix. At a first glance, this should not be applied to $\alpha$-stable sources. However, we have been proved in [11] that a properly normalized covariance matrix converges to a finite matrix with the appropriate structure when the sample size $N$ tends to infinity. Note that for limiting space, all proofs are omitted in this letter and can be found in [11].

*Theorem 1:* Let $\mathbf{x} \triangleq \mathbf{A}\mathbf{s}$ be a data vector of an $\alpha$-stable process mixture and $\hat{\mathbf{R}} \stackrel{\text{def}}{=} \frac{1}{N}\sum_{t=1}^{N} \mathbf{x}(t)\mathbf{x}(t)^T$ it's sample covariance matrix. Then, the normalized covariance matrix of $\mathbf{x}$ defined by

$$\hat{\mathcal{R}} \stackrel{\text{def}}{=} \frac{\hat{\mathbf{R}}}{Trace(\hat{\mathbf{R}})} \quad (1)$$

converges asymptotically to the finite matrix $\mathbf{A}\mathbf{D}\mathbf{A}^T$, where $\mathbf{D} = diag(d_1, \cdots, d_m)$ with $d_i = \frac{\gamma_{s_i}}{\sum_{j=1}^{m} \gamma_{s_j} \|\mathbf{a}_j\|^2}$ where $\|.\|$ denotes the Frobenius norm.

*Proposition 1:* Let $\hat{\mathcal{R}}$ be the normalized covariance matrix defined above in (1) of the considered $\alpha$-stable mixture. Then the inverse square root matrix of $\hat{\mathcal{R}}$ is a data whitening matrix.

### B. Minimum Dispersion Criterion

The *minimum dispersion (MD)* criterion is a common tool in linear theory of stable processes as the dispersion of a stable RV plays a role analogous to the variance. In addition, we should note that the MD criterion is a direct generalization of the MMSE criterion in the Gaussian case [7]. Let $\mathbf{z}(t) \triangleq \mathbf{B}\overline{\mathbf{x}}(t)$ where $\mathbf{B}$ is unitary, $\overline{\mathbf{x}}$ denotes the whitened data, i.e, $\overline{\mathbf{x}} = \mathbf{W}\mathbf{x}$ and $\mathbf{B}$ is a separating matrix to be estimated. Let us consider the global MD criterion given by the sum of dispersions of all entries of $\mathbf{z}$, i.e.

$$J(\mathbf{B}) \triangleq \sum_{i=1}^{m} \gamma_{z_i} \quad (2)$$

where $\gamma_{z_i}$ denotes the dispersion of $z_i(t)$ the i-th entry of $\mathbf{z}(t)$. In this letter we prove that the MD criterion defines a contrast function in the sense that the global minimization of the objective function given in (2) leads to a separating solution. The $p$th order moment of an $\alpha$-stable RV and its dispersion are related through only a constant (see property 4). Therefore, the MD criterion is equivalent to least $l_p$-norm estimation where $0 < p < \alpha$. Although the most widely used contrast functions for BSS are based on the second and fourth-order cumulants [1], we believe however that there are good reasons to extend the class of contrast functions from cumulants to fractional moments, as we argue next. Mutual information (MI) is usually chosen to measure the degree of independence. Because the direct estimation of MI is very difficult, one can then derive approximative contrast functions, often based on cumulant expansions of the densities. However, one can approximate the Shannon entropy (that is closely related to the MI) using the $l_p$-norm concept ([5]) and hence use it to approximate the MI. For example, in [4] the author uses the $l_p$-norm concept to approximate the MI and then to find the optimal contrast function for exponential power family of density $f_p(x) = k_1 \exp(k_2|x|^p)$. Thus we propose the MD criterion for measuring independence of alpha-stable distributed data as shown by the following result.

*Theorem 2:* The minimum dispersion criterion in (2) is a contrast function under orthogonality constraint for separating an instantaneous mixture of $S\alpha S$ sources.

The proposed method requires no or little *a priori* knowledge of the input signals. The dispersion as well as the characteristic exponent $\alpha$ are estimated according to [7] where the proposed estimator is proved to be consistent and asymptotically normal.

### C. Separation Algorithm

Theorem 2 proves that under orthogonal transform the signal has minimum dispersion if its entries are mutually independent. The problem now is to minimize a cost function under orthogonal constrained. Different approaches exist to perform this constrained optimization problem. We chose here to estimate $\mathbf{B}$ as a product of Givens rotations according to $\mathbf{B} = \prod_{\#sweeps} \prod_{1 \leq p < q \leq m} \mathbf{\Omega}_{pq}(\theta)$ Where $\mathbf{\Omega}_{pq}(\theta)$ is the elementary Givens rotation defined as orthogonal matrix where all diagonal elements are 1 except for the two elements $c = \cos(\theta)$ in rows (and columns) $p$ and $q$. Likewise, all off-diagonal elements of $\mathbf{\Omega}_{pq}(\theta)$ are 0 except for the two elements $s = \sin(\theta)$ and $-s$ at positions $(p, q)$ and $(q, p)$, respectively. The minimization of $J(\mathbf{\Omega}_{pq}(\theta))$ is done numerically by searching $\theta$ using a fine grid into $[0, \pi/2]$ [2]. The so called MD algorithm can be summarized as follows:

[1]Note that the proposed method retains its performance under at least mild violation of the assumption under which it is derived. Indeed, the robustness against model deviations is assessed next by simulation experiments.

[2]Here, we consider $[0, \pi/2]$ instead of $[0, \pi]$ because $\Omega_{pq}(\theta + \pi/2)$ is equal to $\Omega_{pq}(\theta)$ up to a generalized permutation matrix.

**Minimum Dispersion Algorithm**

**Step 1.** *Whitening transform.*

**Step 2.** *Sweep. For all pairs* $1 \leq p < q \leq m$*, do*

- *Compute the Givens angle* $0 \leq \hat{\theta}_{pq} < \pi/2$ *that maximize the pairwise independence for* $z_p$ *and* $z_q$ *by minimizing the global dispersion* $J(\mathbf{\Omega}_{pq}(\theta))$.
- *If* $\hat{\theta}_{pq} \geq \theta_{min}$[a]*, rotate the pair accordingly.*
- *If no pair has been rotated in the previous sweep, end. Otherwise perform another sweep.*

[a]The constant $\theta_{min}$ is a threshold value that defines the minimum rotation angle that is significant in estimating $\mathbf{B}$. In our simulation, it is set equal to $\pi/100$ the value of the chosen angle grid resolution.

## III. Performance Evaluation

Consider $n = 3$ mixtures of $m = 3$ i.i.d. impulsive standard $S\alpha S$ ($\mu = 0$ and $\gamma = 1$) source signals. The statistics are evaluated over 100 Monte Carlo runs and the mixing matrix as well as the sources are generated randomly at each run. The performance of our MD method is compared to three widely used BSS algorithms; JADE [1], EASI [2] and RQML [12]. To measure the quality of source separation, we did use the generalized rejection level criterion defined below.

### A. Generalized mean rejection level (GMRL) criterion

To evaluate the performance of the separation method, we propose to define the rejection level $I_{perf}$ as the mean value of the interference signal dispersion over the desired signal dispersion. This criterion generalizes the existing one [9] based on signal powers[3] which represents the mean value of interference to signal ratio. If source $k$ is the desired signal, the related generalized rejection level would be:

$$I_k \stackrel{\text{def}}{=} \frac{\gamma(\sum_{l \neq k} C_{kl} s_l)}{\gamma(C_{kk} s_k)} = \frac{\sum_{l \neq k} |C_{kl}|^\alpha \gamma_l}{|C_{kk}|^\alpha \gamma_k} \tag{3}$$

where $\gamma(x)$ denotes the dispersion of a $S\alpha S$ RV $x$. Therefore, the averaged rejection level is given by $I_{perf} = \frac{1}{m} \sum_{i=1}^{m} I_i$.

### B. Experimental results

In Figure 1, the GMRL of the MD, EASI, JADE and RQML algorithms versus the characteristic exponent is plotted. The sample size is set to $N = 1000$. It appears that the parameter $\alpha$ is of crucial importance as it has a major influence on the separation performance. Two important features are observed: the mean rejection level increases when the sources are very impulsive ($\alpha$ close to zero) or when they are close to the Gaussian case ($\alpha$ close to two). In the latter case (i.e. $\alpha = 2$), the source separation is not possible. Moreover, we observe that the MD algorithm outperforms the other existing algorithms for most values of $\alpha$. In Figure 2, the simulation study shows that estimation errors of the characteristic exponent $\alpha$ of sources distribution have little influence on the performance of the algorithm. In Figure 3, for our proposed MD algorithm, two different scenarios lead to similar performance. In the first scenario, we consider a mixture of three $\alpha$-stable sources with same characteristic exponent $\alpha = 1.5$ and in the second one, we assume wrongly three $S\alpha S$ sources with $\alpha = 1.5$ while, in reality, the sources are $S\alpha S$ with different characteristic exponents $\alpha_1 = 1.5$, $\alpha_2 = 1$ (Cauchy pdf) and $\alpha_3 = 2$ (Gaussian pdf). It can be observed that the algorithm can separate sources from their mixtures even though we deviate from the assumptions under which it is derived. Consequently, the MD algorithm is robust to possible sources modelization errors. Figure 4 shows the performance realized by each of the four BSS algorithms as a function of the sample size $N$ for $\alpha = 1.5$. One can observe that good performance is reached by the MD algorithm for relatively small/medium sample sizes. This figure demonstrates also that EASI fails to separate $\alpha$-stable signals and that JADE is sub-optimal in this context. This is due to the fact that EASI and JADE are not specifically designed for heavy-tailed signals. Moreover, we observe a certain performance gain in favor of the MD algorithm compared to RQML. This is due to the fact that truncating observations, in RQML procedure, created by large source signal values is not optimal because these observations must be very informative. In the sixth experiment, we consider the case where the observation is corrupted by an additive white gaussian noise. The GMRL versus noise power is depicted in Figure 5 for $\alpha = 1.5$ and $N = 1000$. In this experiment, the noise level $\sigma^2$ is varied between 0 dB and -30 dB. As can be seen, the performance degrades significantly when the noise power is high. This might be explained by the fact that the theory does not take into consideration additive noise. Improving robustness against noise is still an open problem under investigation. It can be seen from Figure 6, however, that the proposed MD method has reliable performance and outperforms RQML algorithm in the low or moderate noise power situation.

[3]For $S\alpha S$ processes the variance (power) is replaced by the dispersion.

## IV. Conclusion

We have introduced a two step procedure for $\alpha$-stable source separation. A first whitening step allows to orthogonalize the mixing matrix using a normalized covariance matrix of the observation. In the second step, the remaining orthogonal matrix is estimated by minimizing a global dispersion criterion. The proposed method is robust to modelization errors of the sources pdf. Numerical examples are presented to illustrate the effectiveness of the proposed method that is shown to perform better than the RQLM method. Moreover, they confirm that existing BSS methods, which are not designed specifically for impulsive signals, fail to provide good separation quality.

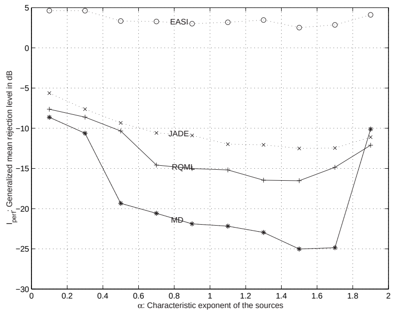
## References


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

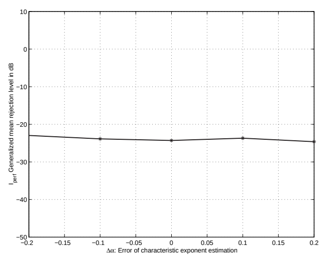
[2] J. F. Cardoso and B. Laheld ”Equivariant adaptive source separation”, *IEEE Tran. on Sig. Proc. (44)*, pp. 3017-3030, 1996.

[3] D. Erdogmus, Y. N. Rao, J. C. Principe, J. Zaohao and K. E. Hild II, ”Simultaneous Extraction of Principal Components Using Givens Rotations and Output Variances” *In Proceeding of ICASSP’2002 Conference*, pp. 1069-1072.

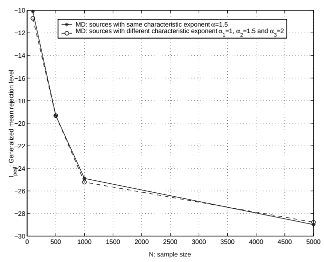
[4] A. Hyvarinen, ”Fast and robust fixed-point algorithms for independent component analysis” *IEEE Transactions on Neural Networks*, Volume: 10 , Issue: 3, Pages: 626 - 634, May 1999.



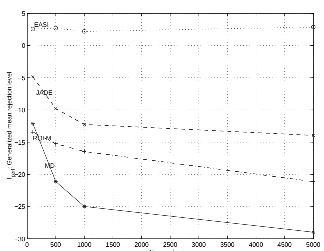

(a) Fig. 2 GMRL vs. $\alpha$ for $N = 1000$.



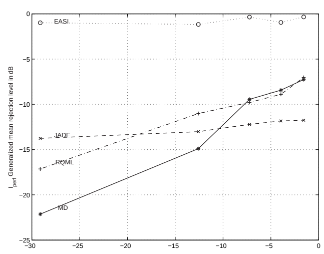

(b) Fig. 3 GMRL vs. the estimation error $\Delta\alpha$.




(c) Fig. 4 GMRL vs. sample size N.




(d) Fig. 5 GMRL vs. sample size N for $\alpha = 1.5$.




(e) Fig. 6 GMRL vs. the additive noise power for $\alpha = 1.5$.


[5] J. Karvanen and A. Cichocki, ”Measuring Sparseness of Noisy Signals”, *In Proc. of ICA’2003 Conference*, Japan.

[6] P. Kidmose, “Blind Separation of Heavy Tail Signals”, *PhD Thesis, Technical University of Denmark (Lyngby)*, Sep. 2001.

[7] C. L. Nikias and M. Shao, “Signal Processing with $\alpha$-Stable Distributions and Applications”, *John Wiley*, 1995.

[8] S. T. Rachev, *Handbook of Heavy Tailed Distributions in Finance*. *Amsterdam: Elsevier*, 2003.

[9] M. Sahmoudi, K. Abed-Meraim and M. Benidir, ”Blind Separation of Instantaneous Mixtures of Impulsive alpha-stable sources”, *3rd International Symposium on Image and Signal Processing and Analysis (ISISPA 2003)*, Rome, Italy, 2003.

[10] M. Sahmoudi, K. Abed-Meraim and M. Benidir “Blind Separation of Heavy-Tailed Signals Using Normalized Statistics”, *In Proceeding of ICA’2004 Conference*, Granada Spain, Sept. 22-24, 2004.

[11] M. Sahmoudi, *Robust Separation and Estimation of Non-stationnary and/or Non-Gaussian Signals*, PhD Thesis, LSS, Université Paris-Sud, France, 2004.

[12] Y. Shereshevski, A. Yeredor and H. Messer,“Super-efficiency in blind signal separation of symmetric heavy-tailed sources”, *In Proc. of 11th IEEE Workshop on Statistical Signal Processing*, pp: 78 -81, Aug. 2001