NOVEL CHARACTERISTIC FUNCTION BASED CRITERIA FOR ICA

J. Eriksson[1,*] *A. Kankainen*[2] *V. Koivunen*[1]

[1]Signal Processing Laboratory
Helsinki Univ. of Technology
P.O. Box 3000, FIN-02015 HUT
Finland

[2]Dept. of Mathematics and Statistics
University of Jyväskylä
P.O. Box 35, FIN-40351 Jyväskylä
Finland

ABSTRACT

We introduce two nonparametric independent component analysis (ICA) criteria based on factorization of characteristic functions. This approach has potential to separate wide class of distributions because characteristic function always exists. A simple criterion allowing for efficient search of the separating matrix and a more advanced criterion possessing desirable consistency property are presented. These criteria may easily be used in orthogonal ICA algorithms. Separating matrix is estimated by establishing pairwise independence among the source signals. Theoretical characteristic functions used in the criteria are replaced by empirical ones. In the examples, the reliable performance of the methods is demonstrated using a variety of source distributions including skewed and heavy-tailed distributions.

1. INTRODUCTION

This paper introduces a new approach to the Blind Source Separation (BSS). It is an Independent Component Analysis (ICA) method that assumes that the source signals are statistically independent. We exploit an alternative definition of the independence expressed in terms of characteristic functions: the joint characteristic function may be factored to a product of the characteristic functions of the independent marginals. Two novel separation criteria based on this factorization are proposed: a simple criterion with low-complexity computational solution and a more advanced criterion with desirable large sample properties. The latter one is a distribution-free, consistent estimator of independence (see [1]).

Conventionally, the factorization of the cumulative distribution function (cdf) is used as a starting point in deriving ICA algorithms. In such approach, cdfs F_1 and F_2 consist of independent components if and only if $F(x_1, x_2) = F_1(x_1)F(x_2)$ for all pairs (x_1, x_2). Naturally this factorization holds for m components. A similar factorization holds for probability density functions (pdf) if the components are assumed to be (absolutely) continuous. Many widely used ICA algorithms are based on the factorization of cdf either explicitly or implicitly through derived contrast functions [2, 3]. In practice, contrast functions are approximated through estimating functions [3]. This is usually done either using parametric models for distributions, or by computing some related statistics. However, such algorithms [4, 5, 6] may even fail in separating certain type of sources encountered in many key application areas, as it is demonstrated in [7, 8], even though these sources are separable in theory. This happens because the sources may not have a good fit in the parametric model, or the computed statistics are not good measures of independence for the sources.

In this paper, the factorization of the characteristic function is used as an objective function (contrast) in blind separation. This approach has significant benefits. First, the source separation is nonparametric since the definition of independence is used directly as a criterion. Moreover, characteristic function always exists and consequently the method has the potential to separate wider class of source distributions than conventional methods. In the process of estimating the separating matrix, the characteristic functions in the definition are just replaced by empirical characteristic functions (ecf). In the case of simple criterion, the joint characteristic function may factorize on a finite interval even though the marginals are not independent. Thus a novel criterion that takes into account the whole real line is proposed.

This paper is organized as follows. Section 2 describes briefly the basic ICA model. In Section 3 the general orthogonal Jacobi algorithm is outlined. In Section 4, the definition of independence is given in terms of characteristic functions. Two objective functions which employ characteristic functions are derived and their properties are discussed in Sections 5 and 6. Optimization of the objective functions is considered as well. Simulations illustrating the reliable performance of the proposed methods are presented in Section 7. Finally, Section 8 concludes the paper.

* This work was funded by the Academy of Finland

2. ICA SYSTEM MODEL

We consider the classical noise-free linear ICA model with instantaneous mixing

$$\mathbf{x} = A\mathbf{s}, \tag{1}$$

where the sources $\mathbf{s} = [s_1, s_2, \ldots, s_m]^T$ are mutually independent real random variables with finite variances and $A_{m\times m}$ is an unknown invertible mixing matrix. The model (1) admits a unique solution up to scaling and permutation indeterminacy

$$\mathbf{y} = W\mathbf{x} \tag{2}$$

such that $WA = \Lambda P$, where Λ diagonal scaling matrix and P is permutation matrix (see [9]). At most one source is allowed to be Gaussian to ensure the identifiability. The goal of a ICA method is to find a separating matrix W such that the recovered sources are as independent as possible.

3. ORTHOGONAL JACOBI APPROACH TO ICA

In finding statistically independent components under the model (1), it suffices that all signals are pairwise independent. Then matrix W in expression (2) is a separating matrix [9]. If a prewhitening step [3] is applied, the matrix W will be orthogonal. This allows for developing an ICA algorithm that employs Jacobi type orthogonal optimization and pairwise processing [9, 10]. Such ICA algorithms were called *Jacobi algorithms*, and they can be summarized as follows [10]:

1. Whitening transform.
2. Sweep. For all pairs $1 \leq i < j \leq m$, do
   (a) Compute the Givens angle $0 \leq \hat{\theta}_{ij} < \pi/2$ that maximize the pairwise independence for the signals y_i and y_j.
   (b) If $\hat{\theta}_{ij} > \hat{\theta}_{\text{min}}$, rotate the pair accordingly.
3. If no pair has been rotated in the previous sweep, end. Otherwise go to step 2.

The constant $\hat{\theta}_{\text{min}}$ above is a threshold value that controls the accuracy of the optimization. It defines the minimum rotation angle that is considered to be significant in finding the separating matrix. In our experiments, we have used the threshold $\hat{\theta}_{\text{min}} = 0.1/\sqrt{n}$, where n is the number of samples, following the guidelines given in [10].

The key point of the algorithm for the ICA problem is the pairwise independence measure in Step 2. Traditionally, cumulant (or moment) based criteria [10] are used. Alternatively, one could use criteria based on factorization of characteristic functions as proposed in Sections 5 and 6.

4. CHARACTERISTIC FUNCTIONS AND INDEPENDENCE

The independence condition for the components of a m-dimensional random vector $\mathbf{x}$ may be written using characteristic functions as follows

$$\psi_{\mathbf{x}}(\mathbf{t}) = \prod_{k=1}^{m} \psi_{\mathbf{e}_k}(t_k) \tag{3}$$

for all $\mathbf{t} = (t_1, t_2, \ldots, t_m)^T \in \mathbb{R}^m$, where ψ is the corresponding m-dimensional characteristic function defined as

$$\psi_{\mathbf{x}}(\mathbf{t}) \triangleq \mathrm{E}\{e^{i<\mathbf{t},\mathbf{x}>}\},$$

where $< \cdot, \cdot >$ denotes the standard vector inner product, and $\psi_{\mathbf{e}_k}$, $i = 1, \ldots, m$, is the characteristic function of the k:th marginal. Denoting the k:th standard unit vector by $\mathbf{e}_k = (0, \ldots, 1, \ldots, 0)^T$ with 1 at the k:th position, the marginals can be written as $\psi(< \mathbf{t}, \mathbf{e}_k >)$ in terms of the joint characteristic function.

The characteristic functions used in the factorization may be estimated from data by means of empirical characteristic functions $\hat{\psi}_n$. For n samples, it is defined as follows

$$\hat{\psi}_n(\mathbf{t}) \triangleq \frac{1}{n} \sum_{l=1}^{n} e^{i<\mathbf{t},\mathbf{x}_l>},$$

where $\mathbf{x}_l$ is the l:th sample $(l = 1, ..., n)$. At every fixed point the estimator converges both in the mean square and almost sure sense [11]. Substituting the empirical characteristic functions into the difference

$$\Delta_{\mathbf{x}}(\mathbf{t}) \triangleq \psi_{\mathbf{x}}(\mathbf{t}) - \prod_{k=1}^{m} \psi_{\mathbf{x}}(<\mathbf{t}, \mathbf{e}_k>), \tag{4}$$

following from the expression (3), we may directly measure the independence instead of using indirect methods (cumulants, approximations of mutual information, etc., see [3]). We use the expression (4) as the basis of constructing the criterion needed in the step 2 of the Jacobi algorithm as described in the following two sections.

5. A SIMPLE ICA CRITERION FOR PAIRWISE INDEPENDENCE

Two scalar random variables X and Y are independent if and only if $\mathrm{Cov}\{f(X)g(Y)\} = 0$ for all f and g ranging over a separating class of functions [12, 13]. A well-known separating class consist of the functions $\cos(tx), \sin(tx)$, $t \geq 0$. Therefore it is in our interest to define the following four functions

$$\begin{aligned}
&\rho_{cc}^{\mathbf{x}}(s,t) \triangleq \mathrm{Cov}\big\{\cos(s <\mathbf{x},\mathbf{e}_1>), \cos(t <\mathbf{x},\mathbf{e}_2>)\big\},\\
&\rho_{cs}^{\mathbf{x}}(s,t) \triangleq \mathrm{Cov}\big\{\cos(s <\mathbf{x},\mathbf{e}_1>), \sin(t <\mathbf{x},\mathbf{e}_2>)\big\},\\
&\rho_{sc}^{\mathbf{x}}(s,t) \triangleq \mathrm{Cov}\big\{\sin(s <\mathbf{x},\mathbf{e}_1>), \cos(t <\mathbf{x},\mathbf{e}_2>)\big\}, \quad (5)\\
&\qquad \text{and}\\
&\rho_{ss}^{\mathbf{x}}(s,t) \triangleq \mathrm{Cov}\big\{\sin(s <\mathbf{x},\mathbf{e}_1>), \sin(t <\mathbf{x},\mathbf{e}_2>)\big\},
\end{aligned}$$

which are all zero for all pairs (s, t) if the components of two dimensional variable $\mathbf{x}$ are independent. It is easy to see the connection to the characteristic function equation (4), namely

$$\Delta_{\mathbf{x}}(s,t) = [\rho_{cc}^{\mathbf{x}}(s,t) - \rho_{ss}^{\mathbf{x}}(s,t)] + i[\rho_{cs}^{\mathbf{x}}(s,t) + \rho_{sc}^{\mathbf{x}}(s,t)].$$

Since $\Delta_{\mathbf{x}}(-s,-t)$ is the complex conjugate of $\Delta_{\mathbf{x}}(s,t)$ and

$$\Delta_{\mathbf{x}}(-s,t) = [\rho_{cc}^{\mathbf{x}}(s,t) + \rho_{ss}^{\mathbf{x}}(s,t)] + i[\rho_{cs}^{\mathbf{x}}(s,t) - \rho_{sc}^{\mathbf{x}}(s,t)],$$

we can define a real-valued function

$$\begin{aligned}
C_{\mathbf{x}}(s,t) \triangleq& \Delta_{\mathbf{x}}(s,t)\Delta_{\mathbf{x}}(-s,-t) + \Delta_{\mathbf{x}}(s,-t)\Delta_{\mathbf{x}}(-s,t)\\
=& |\Delta_{\mathbf{x}}(s,t)|^2 + |\Delta_{\mathbf{x}}(-s,t)|^2\\
=& 2\big(\rho_{cc}^{\mathbf{x}}(s,t)^2 + \rho_{ss}^{\mathbf{x}}(s,t)^2\\
&+ \rho_{cs}^{\mathbf{x}}(s,t)^2 + \rho_{sc}^{\mathbf{x}}(s,t)^2\big).
\end{aligned}$$

This function should be minimized over two dimensional rotations of the variable $\mathbf{x}$ to give the Givens angle needed in the Jacobi optimization. Any 2×2 real orthogonal matrix W may be described as $W_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$. Due to symmetry properties of cosine and sine we have

$$\begin{aligned}
&<W_{(\theta+\pi/2)}\mathbf{x},\mathbf{e}_1> = <W_\theta \mathbf{x},\mathbf{e}_2>\\
&\qquad\qquad \text{and} \qquad\qquad (6)\\
&<W_{(\theta+\pi/2)}\mathbf{x},\mathbf{e}_2> = - <W_\theta \mathbf{x},\mathbf{e}_1> .
\end{aligned}$$

Thus

$$\begin{aligned}
&\rho_{cc}^{W_{(\theta+\pi/2)}\mathbf{x}}(s,t) = \rho_{cc}^{W_\theta \mathbf{x}}(t,s),\\
&\rho_{cs}^{W_{(\theta+\pi/2)}\mathbf{x}}(s,t), = -\rho_{sc}^{W_\theta \mathbf{x}}(t,s),\\
&\rho_{sc}^{W_{(\theta+\pi/2)}\mathbf{x}}(s,t) = \rho_{cs}^{W_\theta \mathbf{x}}(t,s),\\
&\qquad \text{and}\\
&\rho_{ss}^{W_{(\theta+\pi/2)}\mathbf{x}}(s,t) = -\rho_{ss}^{W_\theta \mathbf{x}}(t,s).
\end{aligned}$$

Therefore

$$\begin{aligned}
\hat{C}_{s,t}(\theta) &\triangleq C_{W_\theta \mathbf{x}}(s,t) + C_{W_\theta \mathbf{x}}(t,s) \qquad (7)\\
&= C_{W_{(\theta+\pi/2)}\mathbf{x}}(s,t) + C_{W_{(\theta+\pi/2)}\mathbf{x}}(t,s).
\end{aligned}$$

Since $\hat{C}_{s,t}(\theta)$ is a $\pi/2$-periodic function of θ it has the Fourier series representation

$$\hat{C}_{s,t}(\theta) = a_0 + \sum_{n=1}^{\infty} a_n \cos(4n\theta - \hat{\theta}_n).$$

The empirical estimators corresponding functions (5) are squares of sums of terms in cosine and sine, hence the degree of the representation is low and the estimators should have low variance. Simple golden section search [14] provides convenient tool for fast and accurate minimization.

Because of the scaling indeterminacy in the model (1), we can not give any *a priori* preference on points (s, t) where to test the independence [8]. However, we have found that the simple choice $\hat{C}_{1,1}(\theta)$ applied to data where the sample variances are normalized to one works fine in most cases. Naturally whitening transform performs this normalization.

One could also look for more or better calculation points as is done in some independence tests [11]. However, it seems that the improvements in separation are negligible and the increase in the computational complexity significant.

6. A CONSISTENT ICA CRITERION FOR PAIRWISE INDEPENDENCE

It is not sufficient for two characteristic functions to agree on a finite interval in order them to be the same [11]. The criterion (7) calculated on finite number of points (s, t) (or over a finite interval) may fail to be a good measure of independency for certain signals. Hence, spurious factorizations may result when using that criterion. If, however, two analytic characteristic functions agree on any finite interval around zero, they are the same [11]. Analytic characteristic functions correspond to distributions whose all moments exists. Thus, we may expect the criterion (7) work well if the sources correspond to analytic characteristic functions. However, we would also like to separate reliably other type of sources e.g., distributions where the moments are not necessary defined. This leads us to look for an improved criterion.

Independence tests of the form

$$T = \int_{\mathbb{R}^m} |\Delta_{\mathbf{x}}(\mathbf{t})|^2 g(\mathbf{t}) d\mathbf{t}, \qquad (8)$$

where $g(\cdot)$ is an appropriate real-valued weight function, were considered in [1] (see also [13]). It is shown under a very mild tail condition that this test is consistent, and appears to work better than alternatives found in the literature (see [1]). A particularly good choice of the weight function in (8) is the product of the probability density functions of the standard normal distribution, i.e.

$$g(\mathbf{t}) = \prod_{k=1}^{m} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t_k^2}.$$

In this case, the corresponding test statistic can be written [1] in the form

$$\hat{T} = n\Big[\frac{1}{n^2}\sum_{j=1}^{n}\sum_{l=1}^{n}\prod_{k=1}^{m} e^{-\frac{1}{2}(x_{jk}-x_{lk})^2} - \frac{1}{n^{m+1}}\sum_{j=1}^{n}\prod_{k=1}^{m}\sum_{l=1}^{n} e^{-\frac{1}{2}(x_{jk}-x_{lk})^2} + \frac{1}{n^{2m}}\prod_{k=1}^{m}\sum_{j=1}^{n}\sum_{l=1}^{n} e^{-\frac{1}{2}(x_{jk}-x_{lk})^2}\Big], \quad (9)$$

where m is the dimension and n is the number of samples. The sample variances are assumed to be normalized to unity. Although only the pairwise independence measure is needed, during the sweeps of the Jacobi algorithm the signals are sums of independent random variables. Since the statistic (9) is designed to be a measure of total independence, one might expect it to work better than criteria solely build for two dimensional case. This may also be beneficial for finding the linearly as independent component as possible when the linear model (1) does not hold exactly (there is some non-linear dependence present).

Since the exponent function is monotonic and the criterion (9) is symmetric, it follows from properties (6) that the two dimensional criterion corresponding to (9) is again a $\pi/2$-periodic function of θ, and therefore it can easily be minimized. However, calculation of the criterion is much slower than that of (7), and is some time-critical applications one might opt for the simple criterion.

7. SIMULATIONS

In the simulations, the separation performance is measured in terms of the performance index [15] defined as

$$J \triangleq \sum_{i=1}^{m}\Big(\sum_{j=1}^{m}\frac{|p_{ij}|}{\max_k |p_{ik}|}-1\Big)+\sum_{j=1}^{m}\Big(\sum_{i=1}^{m}\frac{|p_{ij}|}{\max_k |p_{kj}|}-1\Big),$$

where the matrix $P = (p_{ij}) \triangleq WA$, and a zero value indicates perfect separation. In the three signal mixture case the, values of J below 2.5 appear give be meaningful results, and a value below 1 can be considered a successful separation. In addition, the performance is quantitatively measured by the Signal to Interference Ratio (SIR(dB)$= -10\log_{10}$(MSE), where MSE stands for Mean Square Error MSE$= E\left\{(s(n)-y(n))^2\right\}$). To eliminate scaling differences both original signals and extracted signals are normalized to have zero mean and unit variance before the calculation of the SIR. After that source signals are matched to the extracted signals so that the resulting MSE values are as small as possible.

We compared our methods to two widely used ICA algorithms, JADE [5] and Fast-ICA [4] with two different nonlinearities. Three simulation examples with different types of distributions and a variety of sample sizes are presented. All simulation results are averaged over 200 runs.

In the first simulation, three Rayleigh(1) distributed signals were mixed. This type of skewed distributions are extensively employed, for instance, in wireless communications. For each algorithm, the SIR value at different sample sizes is shown in Fig. 1. The boxplot figure corresponding

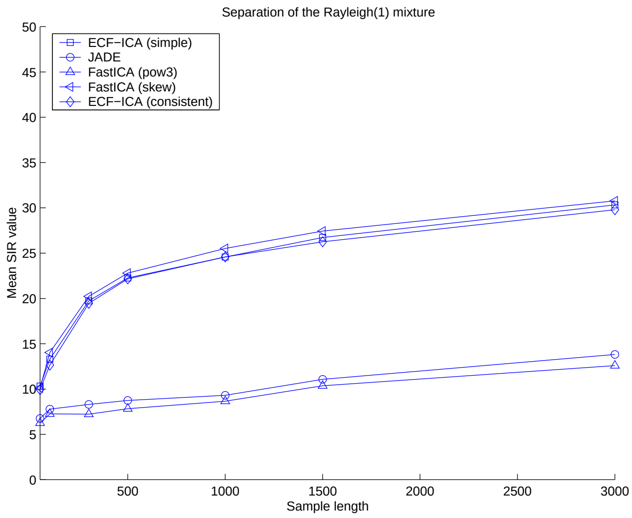

Fig. 1. Obtained SIR values in separating Rayleigh(1) distributed sources with different sample sizes.

to sample size 500 (as all boxplots in this paper) is presented in Fig. 2. The box has lines at the lower quartile, median, and upper quartile values. The lines extending from each end show the extent of the rest of the results with outliers presented as crosses. It can be seen that ECF-ICA methods, and Fast-ICA (skew) can separate signals well already with signal lengths around 200, whereas JADE and Fast-ICA (pow3) fail even with large sample sizes. Many commonly used ICA methods implicitly assume that the source distributions are symmetric and consequently perform poorly in the presence of asymmetric distributions.

The second example is the mixture of three uniformly distributed signals. Results are shown in figures 3 and 4. As noted earlier, the simple characteristic function criterion may fail to find the independent components, since it calculates the difference (4) only in a single point. This can be seen from the simulation, where the simple criterion performance for 2000 observations is comparable to other algorithms with 100 observations. The Fast-ICA (skew) method completely fails in separating the sources.

In the third example, mixtures of three heavy tailed α-stable signals with the characteristic exponent 1.2 (see e.g. [16]) are separated. This type of signals do not possess finite variances, and might therefore create problems for sam-

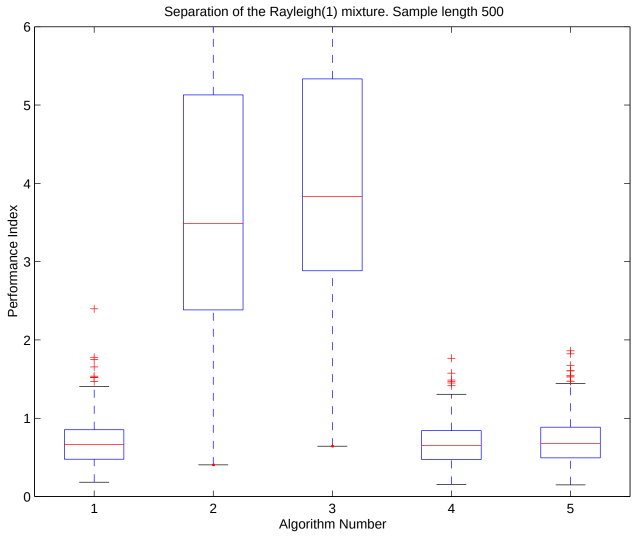

Fig. 2. Separation of Rayleigh(1) sources for the sample length 500. Algorithm numbers correspond (in order): ECF-ICA (simple) (1), JADE (2), Fast-ICA (pow3) (3), Fast-ICA (skew) (4) and ECF-ICA (consistent) (5).

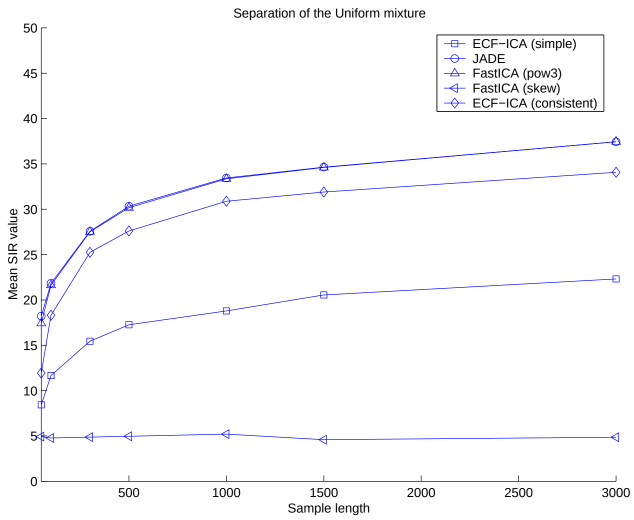

Fig. 3. Obtained SIR values in separating mixtures of Uniformly distributed sources with different sample sizes.

ple moment based methods. It can be seen from Figure 5, however, that all the algorithms perform reliably regardless of the fact that second order moments are not defined. A boxplot illustrating the performance at sample size 500 is given in Figure 6. In overall comparison, ECF-ICA (consistent) method has reliable performance in all situations considered whereas the other methods may fail, in particular if the underlying assumptions on sources are not completely valid. This is due to the fact that ECF-ICA is nonparametric method and it directly minimizes a measure of independence instead of approximating such measure through some statistic.

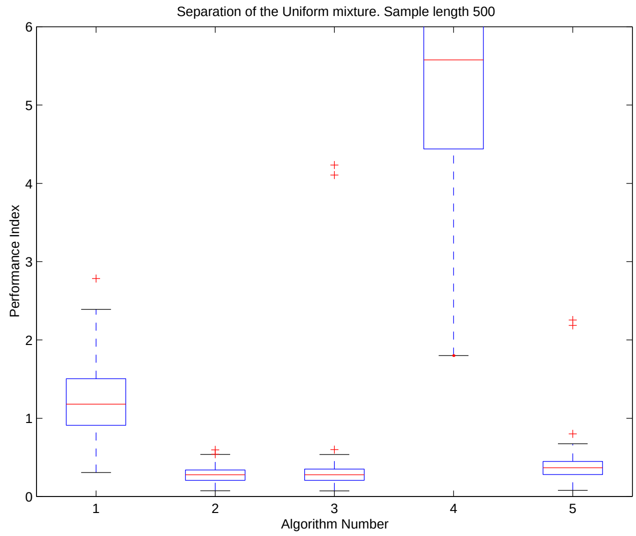

Fig. 4. Separation performance in case of uniformly distributed sources and sample size 500. Algorithm numbers correspond (in order): ECF-ICA (simple) (1), JADE (2), Fast-ICA (pow3) (3), Fast-ICA (skew) (4) and ECF-ICA (consistent) (5).

8. CONCLUSION

We have proposed a new characteristic function based approach to blind source separation. The method assumes that sources are statistically independent and minimizes a criterion that follows directly from the definition of independence in terms of factorization of the characteristic functions. Two such criteria were proposed. The criteria are employed in a Jacobi algorithm that establishes the independence pairwise. The more advanced criterion possess desirable consistency property. It is completely nonparametric and the simulation results indicate that it works reliably for wide variety of source classes even at relatively small sample sizes. The method has a reliable performance even in the cases where the commonly used methods may perform poorly, for example, in the face of skewed source distributions.

9. REFERENCES

[1] A. Kankainen, *Consistent Testing of Total Independence Based on Empirical Characteristic Function*, Ph.D. thesis, University of Jyväskylä, 1995.

[2] T-W Lee, M. Girolami, A. Bell, and T. Sejnowski, “A unifying information-theoretic framework for in-

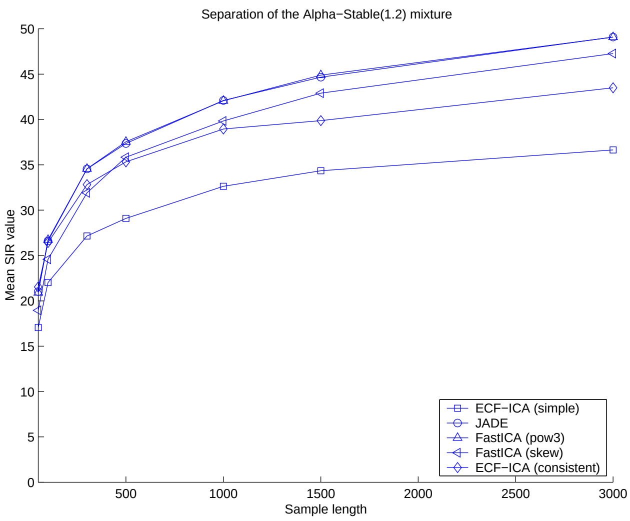

Fig. 5. Obtained SIR values in separating α-stable(1.2) distributed sources with different sample sizes.

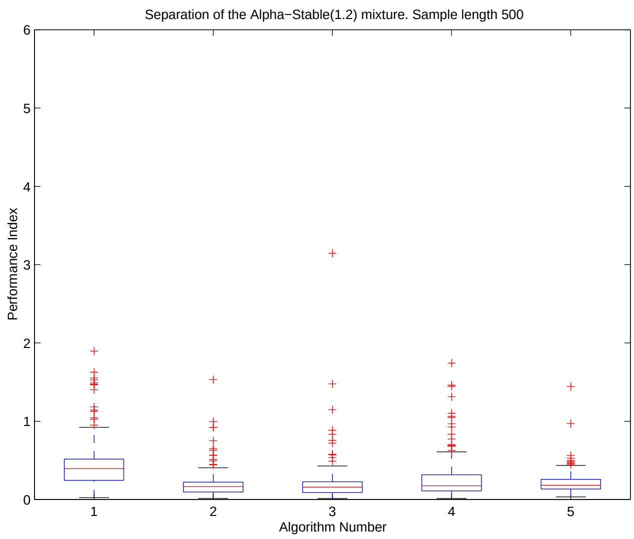

Fig. 6. Separation performance in case of heavy-tailed α-stable sources with characteristic exponent 1.2 and sample size 500. Algorithm numbers correspond (in order): ECF-ICA (simple) (1), JADE (2), Fast-ICA (pow3) (3), Fast-ICA (skew) (4) and ECF-ICA (consistent) (5).

dependent component analysis," *Computers & Mathematics with Applications*, vol. 39, no. 11, pp. 1–21, June 2000.

[3] J-F Cardoso, "Blind signal separation: Statistical principles," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 2009–2025, Oct. 1998.

[4] A. Hyvärinen, "FastICA algorithm," Matlab code with references available at http://www.cis.hut.fi/projects/ica/fastica/.

[5] J-F. Cardoso, "JADE algorithm," Matlab code with references available at http://www.tsi.enst.fr/∼cardoso/stuff.html.

[6] S. Makeig et al., "Infomax algorithm," Matlab code with references available at http://www.cnl.salk.edu/∼scott/.

[7] J. Karvanen, J. Eriksson, and V. Koivunen, "Maximum likelihood estimation of ICA model for wide class of source distributions," in *Proceedings of the 2000 IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing X*, Dec. 2000, vol. 1, pp. 445–454.

[8] J. Eriksson and V. Koivunen, "Blind separation using characteristic function based criterion," in *Proceedings of the 35th Annual Conference on Information Sciences and Systems*, Baltimore, Maryland, Mar. 2001, vol. 2, pp. 781–785.

[9] P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, vol. 36, no. 3, pp. 287–314, Apr. 1994.

[10] J-F. Cardoso, "Higher-order contrast for independent component analysis," *Neural Computation*, vol. 11, no. 1, pp. 157–192, Jan. 1999.

[11] G. Ushakov, *Selected Topics in Characteristic Functions*, Modern Probability and Statistics. VSP VB, Utrect, The Netherlands, 1999.

[12] L. Breiman, *Probability*, Classics in Applied Mathematics. SIAM, 1992.

[13] A. Feuerverger, "A consistent test for bivariate dependence," *International Statistical Review*, vol. 61, no. 3, pp. 419–433, 1993.

[14] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Cawbridge University Press, 1992.

[15] S. Amari, A. Cichocki, and Yang H., "A new learning algorithm for blind signal separation," in *Advances in Neural Information Processing Systems 8*, D. Touretsky, M. Mozer, and H. Hasselmo, Eds., pp. 757–763. MIT Press, 1996.

[16] P. Tsakalides and C. Nikias, "The robust covariation-based MUSIC (ROC-MUSIC) algorithm for bearing estimation in inpulsive noise enviroments," *IEEE Trans. on Signal Processing*, vol. 44, no. 7, pp. 1623–1633, July 1996.