EFFICIENT VARIANT OF ALGORITHM FASTICA FOR INDEPENDENT COMPONENT ANALYSIS ATTAINING THE CRAMÉR-RAO LOWER BOUND

Zbyněk Koldovský[1,2] *and Petr Tichavský*[1]

[1]Institute of Information Theory and Automation,
P.O.Box 18, 182 08 Prague 8, Czech Republic

[2] Faculty of Nuclear Sciences and Physical Engineering
Czech Technical University, Trojanova 13, 120 00 Prague 2

ABSTRACT

We propose an improved version of algorithm FastICA which is asymptotically efficient, i.e., its accuracy attains the Cramér-Rao lower bound provided that the probability distribution of the signal components belongs to the class of generalized Gaussian distribution. Its computational complexity is only slightly (about three times) higher than that of ordinary symmetric FastICA. Simulation section shows superior performance of the algorithm compared with JADE, and of non-parametric ICA.

1. INTRODUCTION

Recently, blind techniques became very popular in the signal processing community, in particular, Independent Component Analysis (ICA). The underlying task of ICA is an instantaneous linear mixing model. In this case, the mixing process can be expressed as

$$\mathbf{X} = \mathbf{AS}, \tag{1}$$

where $\mathbf{X}$ is an $d \times N$ matrix with the $(k,\ell)-$th element denoted $x_{k\ell}$, d is the number of mixed signals and N is the number of samples. Similarly, $\mathbf{S}$ denotes a matrix of samples of the original signals s_{ij}. $\mathbf{A}$ is an unknown regular $d \times d$ mixing matrix. It is assumed that s_{ij} are mutually independent i.i.d. random variables with probability density functions (pdf) $f_i(s_{ij})$ $i = 1, \ldots, d$.

There are several algorithms which have been developed to solve the task. For instance, JADE [4] is based on the approximation of kurtosis via cumulants, NPICA [1] uses a nonparametric model of the density functions. Algorithm FastICA [5] is based on Hyvärinen's entropy approximation which depends on a choice of a nonlinear function $G(\cdot)$.

This work is based on our recently published asymptotic performance analysis of the algorithm FastICA [11] and the Cramér-Rao lower bound (CRB) for ICA [7, 9, 15, 10, 13], where under certain conditions the accuracy of FastICA is shown to be very close to the CRB provided that the function $G(\cdot)$ is a primitive function of the score function of the estimated original signal, i.e., $G(x) = \int \psi(x)dx = -\int \frac{f_i'(x)}{f_i(x)}dx$. Here, using the analysis, we derive an improved algorithm which is under the same condition asymptotically efficient, i.e., its asymptotic performance achieves the CRB.

In the next Section, the results of our previous work are summarized. In Section 3, the improved algorithm is derived and its properties are described. Section 4 presents computer simulations in order to validate the efficiency.

2. RECENT RESULTS

2.1. The original FastICA algorithm

The algorithm FastICA was introduced in [5] in two versions: a one-unit approach and a symmetric one. The first step, which is common for both versions and for many other ICA algorithms, consists in removing the sample mean and the decorrelation of the data $\mathbf{X}$, i.e.,

$$\mathbf{Z} \;=\; \widehat{\mathbf{C}}^{-1/2}\,(\mathbf{X} - \overline{\mathbf{X}}) \tag{2}$$

where $\widehat{\mathbf{C}}$ is the sample covariance matrix, $\widehat{\mathbf{C}} = (\mathbf{X}-\overline{\mathbf{X}})(\mathbf{X}-\overline{\mathbf{X}})^T/N$ and $\overline{\mathbf{X}}$ is the sample mean of the measured data. The one-unit ICA is based on minimization/maximization of the criterion $c(\mathbf{w}) = \mathrm{E}[G(\mathbf{w}^T\mathbf{Z}) - G_0]^2$ where $\mathbf{w}$ is the to-be found vector of coefficients that separates a desired independent component from the mixture (one row of the to be found separating matrix), E stands for the sample mean, $G(\cdot)$ is a suitable nonlinear function, called contrast function, and G_0 is the expected value of $G(\eta)$ where η is a standard normal random variable.

The symmetric FastICA estimates all signals in parallel, and each step is completed by a symmetric orthogonaliza-

[0]This work was supported by Ministry of Education, Youth and Sports of the Czech Republic through the project 1M6798555601 and by the Czech Technical University in Prague through the project CTU0508214.

tion:

$$\mathbf{W}^+ \leftarrow g(\mathbf{WZ})\mathbf{Z}^T - \text{diag}[g'(\mathbf{WZ})\mathbf{1}_N]\,\mathbf{W} \quad (3)$$

$$\mathbf{W} \leftarrow (\mathbf{W}^+\mathbf{W}^{+T})^{-1/2}\mathbf{W}^+ \quad (4)$$

where $g(\cdot)$ and $g'(\cdot)$ denote the first and the second derivative of $G(\cdot)$, respectively. Recently, it was proposed to complete the symmetric FastICA by a test of saddle points that eliminates convergence to side minima of the cost function [11].

The separation quality is treated by means of the so called gain matrix, $\mathbf{G} = \widehat{\mathbf{W}}\mathbf{A}$ which characterizes the relative presence of $j-$th original signal component in the estimated $k-$th component, $j, k = 1, \ldots, d$.

2.2. Analysis of FastICA and the CRB

Let $\mathbf{G}^{1U}$ and $\mathbf{G}^{SYM}$, respectively, be the gain matrix obtained by the one-unit and the symmetric variant using a nonlinear function $g(\cdot)$. It was shown in [11] that if g is sufficiently smooth function, the original signals in the mixture have zero mean and unit variance, and the following expectations exist,

$$\mu_k \stackrel{\text{def}}{=} \text{E}[s_k g(s_k)] \quad (5)$$

$$\rho_k \stackrel{\text{def}}{=} \text{E}[g'(s_k)] \quad (6)$$

$$\beta_k \stackrel{\text{def}}{=} \text{E}[g^2(s_k)] \quad (7)$$

then the normalized gain matrix elements $N^{1/2}\mathbf{G}^{1U}_{k\ell}$ and $N^{1/2}\mathbf{G}^{SYM}_{k\ell}$ have asymptotically Gaussian distribution $\mathcal{N}(0, V^{1U}_{k\ell})$ and $\mathcal{N}(0, V^{SYM}_{k\ell})$, where

$$V^{1U}_{k\ell} = \frac{\gamma_k}{\tau_k^2} \qquad V^{SYM}_{k\ell} = \frac{\gamma_k + \gamma_\ell + \tau_\ell^2}{(\tau_k + \tau_\ell)^2}. \quad (8)$$

with $\gamma_k = \beta_k - \mu_k^2$, and $\tau_k = |\mu_k - \rho_k|$.

The variance in (8) is minimized if the function g equals to the score function ψ_k of the corresponding distribution. It is shown to be close but not to coincide with the Cramér-Rao lower bound [7], which is

$$\text{CRB}(\mathbf{G}_{k\ell}) = \frac{1}{N}\frac{\kappa_k}{\kappa_k\kappa_\ell - 1}. \quad (9)$$

where $\kappa_k \stackrel{\text{def}}{=} \text{E}\left[\psi_k^2(s_k)\right]$, $k, \ell = 1, \ldots, d$.

3. EFFICIENT FASTICA

3.1. Generalized symmetric FastICA

Consider a generalized version of the symmetric version of FastICA, where the first step (3) of the iteration will be followed by multiplying each row of $\mathbf{W}^+$ with suitable positive number c_i $i = 1, \ldots, d$ before the symmetric orthogonalization (4). One iteration of such algorithm can be written in three steps:

$$\mathbf{W}^+ \leftarrow g(\mathbf{WZ})\mathbf{Z}^T - \text{diag}[g'(\mathbf{WZ})\mathbf{1}_N]\,\mathbf{W} \quad (10)$$

$$\mathbf{W}^+ \leftarrow \text{diag}[c_1, \ldots, c_d] \cdot \mathbf{W}^+ \quad (11)$$

$$\mathbf{W} \leftarrow (\mathbf{W}^+\mathbf{W}^{+T})^{-1/2}\mathbf{W}^+ \quad (12)$$

The resultant method can be analyzed in the same way as the plain symmetric FastICA in [11]. It can be derived that the non-diagonal normalized gain matrix elements for this method, $N^{1/2}\mathbf{G}^{GS}_{k\ell}$, have asymptotically Gaussian distribution $\mathcal{N}(0, V^{GS}_{k\ell})$, where

$$V^{GS}_{k\ell} = \frac{c_k^2\gamma_k + c_\ell^2(\gamma_\ell + \tau_\ell^2)}{(c_k\tau_k + c_\ell\tau_\ell)^2}. \quad (13)$$

Note that the criterion (13) is invariant with respect to multiplying all parameters $\{c_k\}$ by the same factor.

Assume now that k is fixed and let c_k be chosen, without any loss in generality, as $c_k = 1$. Then, minimization of the expression (13) with respect to c_ℓ, $\ell \neq k$, can be performed analytically, and it gives (using the notation so that instead of c_ℓ, we use $c_{k,\ell}$ to denote the other parameters in the case that $c_k = 1$)

$$c^{OPT}_{k,\ell} = \arg\min_{c_\ell, c_k = 1} V^{GS}_{k\ell} = \frac{\tau_\ell\gamma_k}{\tau_k(\gamma_\ell + \tau_\ell^2)}\,. \quad (14)$$

Note that neither zero nor $+\infty$ minimizes (13) unless $c^{OPT}_{k\ell}$ equals to it.

The optimum value of the criterion is

$$V^{OPT,k}_{k\ell} = \min_{c_\ell, c_k = 1} V^{GS}_{k\ell} = \frac{\gamma_k(\gamma_\ell + \tau_\ell^2)}{\tau_\ell^2\gamma_k + \tau_k^2(\gamma_\ell + \tau_\ell^2)}. \quad (15)$$

Note that the algorithm in (10)-(12) can be made even more general if a different row of $\mathbf{W}$, $\mathbf{W}^+$ in (10) have its own nonlinear function $g_k(\cdot)$.

In the special case that $g_k = \psi_k$, i.e. equals to the score function of the $k-$th signal for each $k = 1, \ldots, d$, we get $\beta_k = \rho_k = \kappa_k$, $\mu_k = 1$, and $\tau_k = \gamma_k = \kappa_k - 1$. Then (15) simplifies to

$$V^{OPT,k}_{k\ell} = \frac{\kappa_k}{\kappa_k\kappa_\ell - 1} = N\,\text{CRB}[\mathbf{G}_{k\ell}].$$

This means that the estimator asymptotically attains the CRB for $N \to \infty$.

3.2. Proposed algorithm

The proposed algorithm, called for brevity "Efficient FastICA" combines the idea of the generalized symmetric FastICA with an adaptive choice of the function g. The algorithm consists of two steps: (1) original symmetric FastICA

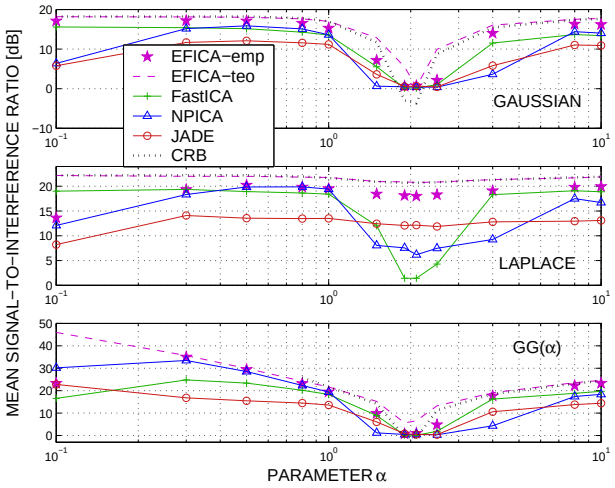

Fig. 1(a) Average SIRs of the three kinds of signals in Example 1.

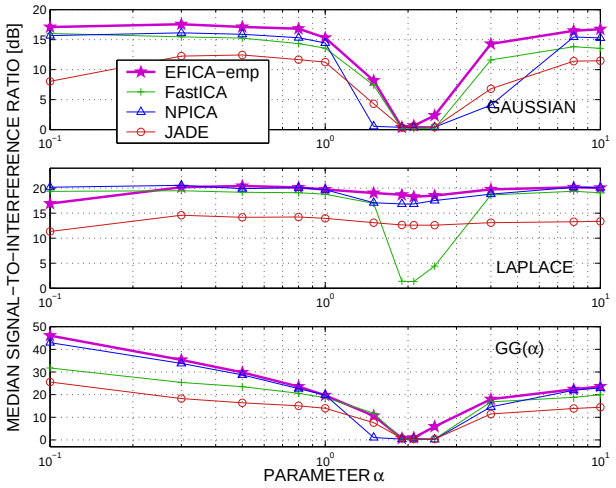

Fig. 1(b) Median SIRs of the same signals.

with the test of saddle points, and (2) a refinement. The refinement utilizes different nonlinearities g adaptively chosen based on outcome of step 1.

The purpose of the step 1 is to quickly and reliably get the primary estimates of the original signals. In this step, therefore, the optional nonlinearity in the original symmetric FastICA $g(x) = \tanh(x)$ is used in this paper due to its universality, but other possibilities seems to give promising results as well, e.g. $g(x) = \frac{x}{1+x^2}$.

3.2.1. Adaptive choice of the nonlinearities

Assume that $\widehat{\mathbf{u}}_k$ is the $k-$th estimated independent signal component obtained in the step 1.

A straightforward solution would be to choose $g_k(\cdot)$ as a score function that belongs to the sample distribution function. This approach would have two flaws. First, the score function is difficult to estimate using the given sample distribution of the component, because it needs not only the probability density function but its derivative, also. The second flaw is that a suitable g needs to be continuous and differentiable.

We suggest an adaptive choice of g_k that works well for the class of generalized Gaussian distribution with parameter α, denoted GG(α) (see Appendix B for details). This class include Standard Gaussian and Laplacean distribution for $\alpha = 2$ and $\alpha = 1$, respectively, and a uniform distribution in the limit $\alpha \to \infty$. The score function of this distribution is proportional to $g(x) = \text{sign}(x) \cdot |x|^{\alpha-1}$ (the scalar multiplicative factor is irrelevant for usage in the FastICA).

Estimation of a suitable parameter α will be done by fitting theoretical fourth order moment of the GG(α) distribution, which is $m_4(\alpha) = \Gamma(\frac{1}{\alpha})\Gamma(\frac{5}{\alpha})\Gamma^{-2}(\frac{3}{\alpha})$ with the sample fourth-order moment,

$$\widehat{m}_{4k} = \widehat{\mathbf{1}}_N^T \widehat{\mathbf{u}}_k^{\cdot 4}/N \tag{16}$$

where $\{\cdot\}^{\cdot 4}$ denotes the elementwise fourth order power. The result of the fitting, valid for $\widehat{m}_{4k} > \lim_{\alpha\to\infty} m_4(\alpha) = 9/5 = 1.8$, is

$$\widehat{\alpha}_k = [\eta_1\sqrt{\widehat{m}_{4k} - 1.8} - \eta_2(\widehat{m}_{4k} - 1.8)]^{-1}$$

with $\eta_1 \approx 0.2906$ and $\eta_2 \approx 0.1851$. If $\widehat{m}_{4k} \leq 1.8$ or $\widehat{\alpha}_k > 15$ maximum power 14 in function g_k is used to avoid numerical problems of the algorithm. Details are deferred to Appendix A.

The problem with the score function of the GG(α) distribution is that it is not continuous for $\alpha \leq 1$ and thus it is not a valid nonlinearity for FastICA. For these α's the statistical efficiency cannot be achieved by the algorithm.

Our ad-hoc choice of g which produces excellent results for supergaussian distributions GG(α) with $\alpha \leq 2$, better than "gauss" in [5], is $g(x) = x\exp(-\eta_3|x|)$ where η_3 is a suitable parameter. The optimum choice of η_3 for the Laplacean distribution can be found to be $\eta_3 \approx 3.348$. In this paper we use this choice for all (supergaussian) components obeying $\widehat{m}_{4k} > 3$.

In summary, the nonlinearity of the choice is

$$g_k(x) = \begin{cases} x\exp(-\eta_3|x|) & \text{for} \quad \widehat{m}_{4k} > 3 \\ \text{sign}(x)\cdot|x|^{\min\{\widehat{\alpha}_k - 1, 14\}} & \text{for} \quad \widehat{m}_{4k} \in [1.8, 3] \\ \text{sign}(x)\cdot|x|^{14} & \text{for} \quad \widehat{m}_{4k} < 1.8 \end{cases} \tag{17}$$

3.2.2. The Refinement

The fact that the asymptotic variance (13) cannot be minimized, in general, for all $k, \ell = 1, \ldots, d$ jointly suggest doing the refinement for each independent component separately [3]. The first step, that is the analogy of (10), is common for all component together:

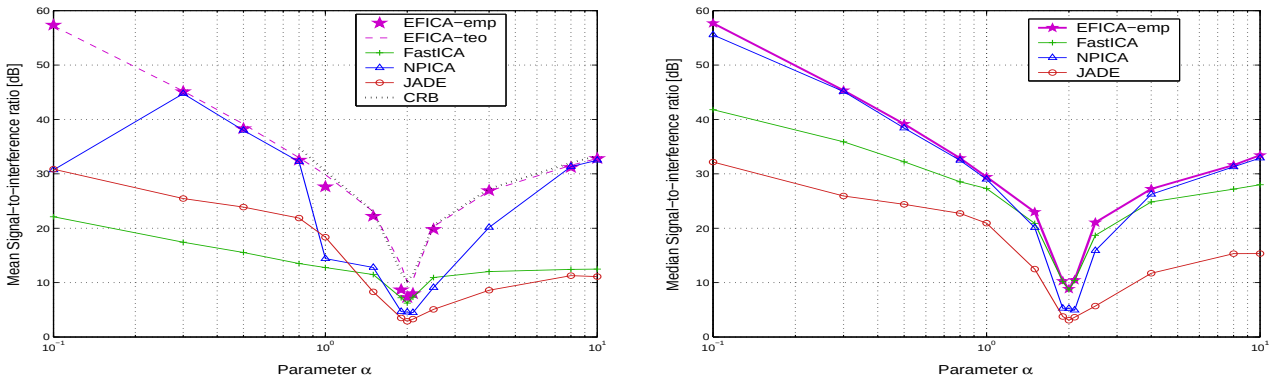

Fig. 2 Quality of separation of 13 GG(α) components with α, respectively, equal to 0.1, 0.3, 0.5, 0.8, 1, 1.5, 1.9, 2, 2.1 ,2.5, 4, 8, and 10. Mean SIRs of the estimated signals are in the left diagram while in the right are median SIRs.

Step R1: For each $k = 1, \ldots, d$, compute $\widehat{m}_{4k}$ in (16) and g_k in (17). Put

$$\widehat{\mathbf{W}}^{+} = [\widehat{\mathbf{w}}_1^{+}, \ldots, \widehat{\mathbf{w}}_d^{+}]^T \,. \tag{18}$$

where

$$\widehat{\mathbf{w}}_k^{+} = \mathbf{Z} g_k(\widehat{\mathbf{u}}_k) - \widehat{\mathbf{w}}_k \, g_k'(\widehat{\mathbf{u}}_k) \mathbf{1}_N \tag{19}$$

for $k = 1, \ldots, d$. (Recall that $\widehat{\mathbf{w}}_k^T \mathbf{Z} = \widehat{\mathbf{u}}_k^T$).
Step R2: For each $k = 1, \ldots, d$, compute

$$\begin{array}{lll} \widehat{\mu}_k &= \widehat{\mathbf{u}}_k^T g_k(\widehat{\mathbf{u}}_k)/N, & \widehat{\tau}_k = |\widehat{\mu}_k - \widehat{\rho}_k| \\ \widehat{\rho}_k &= \widehat{\mathbf{1}}_N^T g_k'(\widehat{\mathbf{u}}_k)/N & \widehat{\gamma}_k = \widehat{\beta}_k - \widehat{\mu}_k^2 \\ \widehat{\beta}_k &= \widehat{\mathbf{1}}_N^T g_k^2(\widehat{\mathbf{u}}_k)/N & \end{array} \tag{20}$$

For each $k, \ell = 1, \ldots, d$, compute

$$c_{k,\ell} = \begin{cases} \frac{\widehat{\tau}_\ell \widehat{\gamma}_k}{\widehat{\tau}_k (\widehat{\gamma}_\ell + \widehat{\tau}_\ell^2)} & \text{for} \quad \ell \neq k \\ 1 & \text{for} \quad \ell = k \end{cases}$$

Next, for each $k = 1, \ldots, d$, compute

$$\widehat{\mathbf{W}}_k^{+} = \text{diag}[c_{k1}, \ldots, c_{kd}] \cdot \widehat{\mathbf{W}}^{+} \tag{21}$$

$$\widehat{\mathbf{W}}_k^{aux} = (\widehat{\mathbf{W}}_k^{+} \widehat{\mathbf{W}}_k^{+T})^{-1/2} \widehat{\mathbf{W}}_k^{+} \tag{22}$$

$$\widehat{\mathbf{w}}_k^{REF} = (\widehat{\mathbf{W}}_k^{aux})_{k,\cdot}^T \tag{23}$$

The resultant $k-$th component estimate is $\widehat{\mathbf{u}}_k^{REF} = \mathbf{Z}^T \widehat{\mathbf{w}}_k^{REF}$. More detailed description of the algorithm will be provided in [8].

4. SIMULATIONS

Three simulation experiments were carried out to approve the efficiency of the proposed algorithm and to provide a comparison with other well-known methods for ICA. For this purpose, we chose the original version of the symmetric FastICA [5] with nonlinearity "tanh", JADE [4], and NPICA [1].

There occur some random convergence failures in the original FastICA and NPICA, which in our method are fixed by the test of saddle points. To eliminate their effect in the comparisons, both mean and median SIRs from each experiment are shown.

Example 1. A Gaussian component, 10 Laplacean, and 10 components with Generalized Gaussian distribution with parameter $\alpha \in [0.1, 10]$ were generated in 100 independent trials. We have chosen the length of data $N = 2000$, which is quite small for separation of 21 signals. In each trial, the signals were mixed with a randomly generated matrix, and separated by the proposed method and forenamed algorithms. In Figure 1(a), the three diagrams show the average SIRs of the Gaussian signal, of the first Laplacean, and of the first with Generalized Gaussian distribution. Respecting the proposed method (designated as EFICA), both the theoretical SIRs estimated via (15) and the empirical (pentagrams) are presented. The results can be compared with those in Figure 1(b) where median SIRs from the same experiment are presented. The theoretical SIRs and the CRB are not included here, since they are derived only for mean SIR.

Example 2. 13 signals with Generalized Gaussian distribution with different values of the parameter α taken from interval $[0.1, 10]$ were mixed with a random mixing matrix and demixed. The experiment was repeated 100 times with fixed length of data $N = 5000$. The results are plotted in Figure 2(a) and 2(b).

[0]The matlab code for the algorithm is made available at http://www.utia.cas.cz/user_data/scientific/SI_dept/Tichavsky.html.

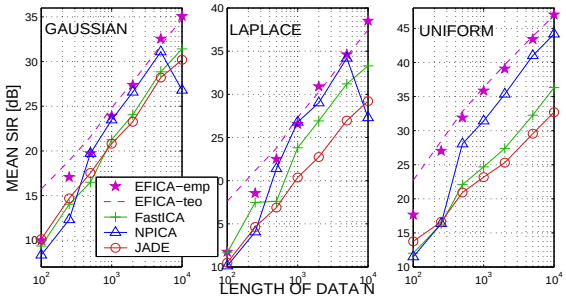

Fig. 3(a) Mean SIRs for various length of data.

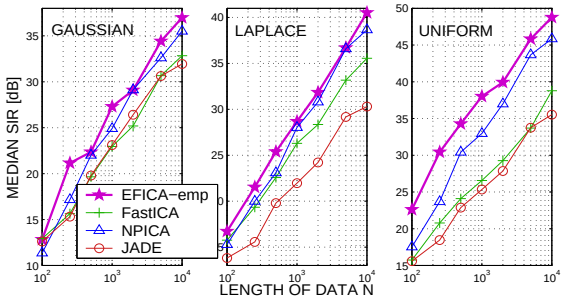

Fig. 3(b): Median SIRs.

Example 3. To demonstrate the performance of the method for different lengths of data three signals with Gaussian, Laplacean, and uniform distribution were mixed with a random mixing matrix and demixed. The average and median SIRs from 100 independent trials for each length of data are plotted in Figure 3(a) and 3(b).

Example 4. In order to demonstrate the computational complexity similar experiment to that in [1] was done. The average CPU time[1] required by the compared methods is shown in Fig. 4(a) for a various length of data when 6 generalized Gaussian signals with a random parameter $\alpha \in [0.1, 10]$ were separated. The results for a fixed length of data $N = 1000$ and a variable number of signals are in Fig. 4(b).

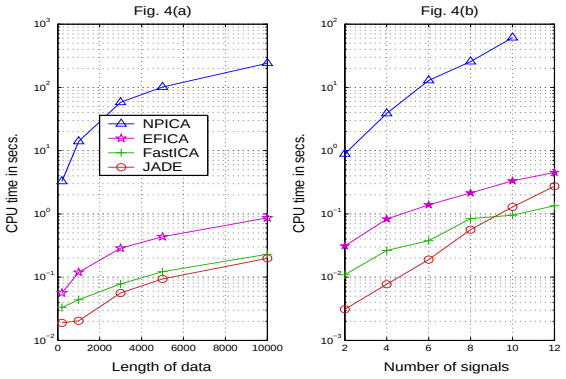

Fig. 4 CPU time needed to separate (a) 6 signals of a various lenght (b) various # of signals of a fixed length $N = 1000$.

[1]The experiment was performed in Matlab on a Pentium IV 2.4 GHz PC with 512 MB of RAM.

Appendix A

To estimate the shape parameter α of Generalized Gaussian variable we will use its fourth moment as a function of α [14]

$$m_4(\alpha) \stackrel{\text{def}}{=} \mathrm{E}[x^4] = \frac{\Gamma(\frac{1}{\alpha})\Gamma(\frac{5}{\alpha})}{\Gamma^2(\frac{3}{\alpha})} \tag{24}$$

We consider that $\alpha \in (2, +\infty)$. Following asymptotic expansion of the third order for $\alpha \to +\infty$ was derived in Maple^{TM}

$$z(\alpha) \stackrel{\text{def}}{=} \sqrt{m_4(\alpha) - \frac{9}{5}} = \frac{A}{\alpha} - \frac{B}{\alpha^2} + o(\frac{1}{\alpha^2}), \tag{25}$$

where $A = \frac{\sqrt{30}\pi}{5}$, $B = \frac{18\sqrt{30}\zeta(3)}{5\pi}$, and $\zeta(3) \doteq 1.202$. To invert the above relation, first, we use a substitution $y = \frac{1}{\alpha}$. Now, we can write

$$\begin{aligned} y &= \frac{1}{A}z + \frac{B}{A}y^2 + o(y^2) = \\ &= \frac{1}{A}z + \frac{B}{A}\left(\frac{1}{A}z + \frac{B}{A}y^2 + o(y^2)\right)^2 + o(y^2) = \\ &= \frac{1}{A}z + \frac{B}{A^3}z^2 + o(z^2) \end{aligned} \tag{26}$$

Using the definition (25) of z follows

$$\alpha \approx \left(\frac{1}{A}\sqrt{m_4(\alpha) - \frac{9}{5}} - \frac{B}{A^3}\left(m_4(\alpha) - \frac{9}{5}\right)\right)^{-1} \tag{27}$$

Note that $m_4(\alpha)$ can be simply estimated as a sample fourth moment of the corresponding variable.

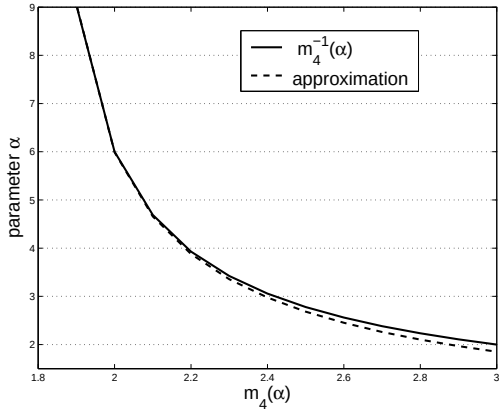

Fig. 3: Comparison of the inversion of (24) with the approximation (27).

Appendix B

Generalized Gaussian distribution family

Introduce the generalized Gaussian density function with parameter α, zero mean and variance one, as

$$f_\alpha(x) = \frac{\alpha\beta_\alpha}{2\Gamma(1/\alpha)} \exp\left\{-(\beta_\alpha|x|)^\alpha\right\} \tag{28}$$

where $\alpha > 0$ is a positive parameter that controls the distribution's exponential rate of decay, and $\Gamma(\cdot)$ is the Gamma function and $\beta_\alpha = \sqrt{\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}}$. This generalized Gaussian family encompasses the ordinary standard normal distribution for $\alpha = 2$, the Laplacean distribution for $\alpha = 1$, and the uniform distribution in the limit $\alpha \to \infty$.

The $k-$th absolute moment for the distribution is

$$\mathrm{E}_\alpha\{|x|^k\} = \int_\infty^\infty |x|^k \, f_\alpha(x)dx = \frac{1}{\beta_\alpha^k} \frac{\Gamma\left(\frac{k+1}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)} \tag{29}$$

The score function of the distribution is

$$\psi_\alpha(x) = -\frac{\frac{\partial f_\alpha(x)}{\partial x}}{f_\alpha(x)} = \frac{|x|^{\alpha-1}\mathrm{sign}(x)}{\mathrm{E}_\alpha[|x|^\alpha]} \tag{30}$$

Then, simple computations give

$$\kappa_\alpha = \mathrm{E}_\alpha[\psi_\alpha^2(x)] = \frac{\mathrm{E}_\alpha[|x|^{2\alpha-2}]}{\{\mathrm{E}_\alpha[|x|^\alpha]\}^2} = \frac{\Gamma\left(2-\frac{1}{\alpha}\right)\Gamma\left(\frac{3}{\alpha}\right)}{\left[\Gamma\left(1+\frac{1}{\alpha}\right)\right]^2} \tag{31}$$

Note that $\kappa_\alpha = +\infty$ for $\alpha \leq \frac{1}{2}$, $\kappa_\alpha = 1$ for $\alpha = 2$ (the distribution is standard Gaussian), and $\kappa_\alpha \to +\infty$ for $\alpha \to +\infty$ (the distribution approaches the uniform distribution on $< -\sqrt{3}, \sqrt{3} >$).

Acknowledgement

The authors acknowledge help of Prof. Oja from Helsinki University of Technology in improving presentation of this paper.

5. REFERENCES

[1] R. Boscolo, H. Pan, and V. P. Roychowdhury, "Independent Component Analysis Based on Nonparametric Density Estimation", *IEEE Trans. on Neural Networks*, vol. 15, no. 1, pp. 55-65, 2004.

[2] J.-F. Cardoso, "Blind signal separation: statistical principles", *Proceedings of the IEEE*, vol. 90, n. 8, pp. 2009-2026, Oct. 98.

[3] J.-F. Cardoso, "On the performance of orthogonal source separation algorithms", *Proc. EUSIPCO*, pp. 776-779, Edinburgh, September 1994.

[4] J.-F. Cardoso, "High-order Contrasts for Independent Component Analysis", *Neural Computat.*, vol. 11, no. 1,pp. 157-192, 1999.

[5] A. Hyvärinen, "A Fast Fixed-Point Algorithm for Independent Component Analysis". *Neural Computation*, 9(7):1483-1492, 1997.

[6] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, Wiley-Interscience, New York, 2001.

[7] Z. Koldovský, P. Tichavský and E. Oja, "Cramér-Rao lower bound for linear Independent Component Analysis", *Proc. ICASSP 2005*, Philadelphia, Vol. III, pp. 581 - 584, March 2005.

[8] Z. Koldovský, P. Tichavský and E. Oja, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound", submitted.

[9] D.T. Pham, Garat, P., "Blind separation of mixture of independent sources through a quasi-maximum likelihood approach", *IEEE Tr. Signal Processing*, Vol. 45, No. 7, July 1997, pp. 1712 - 1725.

[10] O. Shalvi, E. Weinstein, "Maximum likelihood and lower bounds in system identification with non-Gaussian inputs", *IEEE Tr. Information Theory*, Vol. 40, No. 2, March 1994, pp. 328 - 339.

[11] P. Tichavský, Z. Koldovský, and E. Oja, "Asymptotic Performance Analysis of the Fixed-Point Algorithm (Fast-ICA) for Independent Component Analysis", *SSP-2005*, Bordeaux.

[12] P. Tichavský, Z. Koldovský, and E. Oja, "Performance Analysis of the FastICA Algorithm and Cramér-Rao Bounds for Linear Independent Component Analysis", accepted for publication in *IEEE Trans. on Signal Processing*, May 2005.

[13] V. Vigneron, Ch. Jutten, "Fisher information in source separation problems", *Proc. ICA 2004*, Granada, pp. 168-176.

[14] K. Waheed, F. M. Salam, "Blind source recovery using an adaptive generalized Gaussian score function", Circuits and Systems, 2002. MWSCAS-2002. The 2002 45th Midwest Symposium on, -7 Aug. 2002 Vol.3, pp. 656-659.

[15] D. Yellin, B. Friedlander, "Multichannel system identification and deconvolution: performance bounds", *IEEE Tr. Signal Processing*, Vol. 47, No. 5, May 1999, pp. 1410 - 1414.