Fei Ge and Jinwen Ma*

100871,

jwma@math.pku.edu.cn

Abstract.

Keywords:

Independent Component Analysis (ICA) provides a powerful statistical tool for signal processing and data analysis. It aims at decomposing a random vector which is an instantaneous linear combination of several independent random variables. Thus, the decomposed components should be mutually as independent as possible. One major application of ICA is Blind Signal Separation (BSS), where simultaneous observations $\mathbf{x}(t) = [x_1(t), \ldots, x_m(t)]^T$ are linear mixtures of independent signal sources $\mathbf{s}(t) = [s_1(t), \ldots, s_n(t)]^T$ via a mixing matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ such that $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$. Typically, we can consider the case $m = n$ and the purpose of ICA is to solve or learn an $n \times n$ matrix $\mathbf{W}$ such that $\mathbf{WA}$ has one and only one non-zero entry in each row and in each column. In fact, a such $\mathbf{W}$, being called a separating matrix or demixing matrix, corresponds to a feasible solution of the ICA problem. Clearly, the independence assumption on these estimated components is the key to solve the ICA problem. That is, if $\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t$

Actually, the independence measure among the estimated components can serve as a good objective or contrast function for ICA. Supposing that $p_i(y_i)$ is the marginal probability density function (pdf) of the i-th component of $\mathbf{y} = \mathbf{Wx} = \mathbf{WAs}$, and $p(\mathbf{y})$ is the joint pdf of $\mathbf{y}$, we can use the Kullback divergence to set up the following Minimum Mutual Information (MMI) criterion [1]:

$$I(\mathbf{y}) = \int p(\mathbf{y}) \log \frac{p(\mathbf{y})}{\prod_{i=1}^{n} p_i(y_i)} d\mathbf{y} \ . \tag{1}$$

Clearly, $I(\mathbf{y})$ is nonnegative and vanishes to zero only when all y_i are mutually independent. Moreover, this MMI criterion is equivalent to the Maximum Likelihood (ML) criterion [2] if $p_i(\cdot)$ coincides with the pdf of each source.

Since the pdfs of the sources are unknown in advance, we generally utilize some predefined or model pdfs to substitute the real pdfs in the mutual information. In such a way, however, the MMI approach works only in the cases where the components of $\mathbf{y}$ are either all super-Gaussians [3] or all sub-Gaussians [4]. For the cases where sources contain both super-Gaussian and sub-Gaussian signals in an unknown manner, it was conjectured that these model pdfs $p_i(y_i)$ should keep the same kurtosis signs of the source pdfs. This conjecture motivated the proposal of the so-called one-bit matching condition [5], which can be basically stated as "all the sources can be separated as long as there is a one-to-one same-sign-correspondence between the kurtosis signs of all source pdf's and the kurtosis signs of all model pdf's". Along the one-bit matching condition, Liu, Chiu, and Xu simplified the mutual information into a cost function and proved that the global maximum of the cost function correspond to a feasible solution of the ICA problem [6]. Ma, Liu, and Xu further proved that all the maxima of the cost function corresponds to the feasible solutions in two-source mixing setting [7]. Recently, this cost function was further analyzed in [8] and an e cient learning algorithm was constructed with it in [9]. However, the one-bit matching condition is not su cient for the MMI criterion because Vrins and Verleysen [10] have already proved that spurious maxima exist for it when the sources are strongly multimodal.

On the other hand, there have been many ICA algorithms that explicitly or implicitly utilize certain flexible pdfs to fit di erent types of sources. Actually, these methods learn the separating matrix as well as the parameters in the flexible model pdfs, or nonlinear functions, or switching functions, simultaneously. From the simple switching or parametric functions (e.g., [11,12,13]) to the complex mixture densities (e.g., [5,14,15]), these flexible functions have enabled the algorithms to successfully separate the sources in both simulation experiments and applications. However, there is still an essential issue whether all the local optima of the objective function in each of these methods can correspond to the feasible solutions. Clearly, if all the local optima correspond to the feasible solutions, any gradient-type algorithm can be always successful on solving the ICA problem. Otherwise, if there exists some optimum which does not correspond to a feasible solution, any gradient-type algorithm may be trapped in such a local optimum and lead to a spurious solution. Thus, for an objective function, it is

It follows from Eq.(2) that:

$$\text{kurt}\{\alpha x\} = \alpha^4 \text{kurt}\{x\}, \quad \alpha \in \mathbb{R}; \tag{3}$$

and if x_1 and x_2 are independent, we certainly have

$$\text{kurt}\{x_1 + x_2\} = \text{kurt}\{x_1\} + \text{kurt}\{x_2\} \ . \tag{4}$$

2.1 Kurtosis-Sum Objective Function

We consider the ICA problem with n sources and n observations. Without loss of generality, we assume that the sources have zero mean and unit variance. Moreover, the observed signals can be further pre-whitened such that $E\{\mathbf{x}\} = 0$, and $E\{\mathbf{x}\mathbf{x}\}^T = \mathbf{I}$. Then, for any orthogonal transformation matrix $\mathbf{W}$, the estimated signals $\mathbf{y} = \mathbf{W}\mathbf{x}$ are always whitened.

The kurtosis-sum objective function is defined by

$$J(\mathbf{W}) = \sum_{i=1}^{n} |\text{kurt}\{y_i\}| = \sum_{i=1}^{n} |\text{kurt}\{\mathbf{w}_i^T \mathbf{x}\}|, \tag{5}$$

where $\mathbf{x}$ is the (pre-whitened) observed signal (as a random vector), and $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_n]^T$ is the orthogonal de-mixing matrix to be estimated.

Since the two transformations are linear, $\mathbf{y} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{A}\mathbf{s} = \mathbf{R}\mathbf{s}$, where $\mathbf{R}$ is another orthogonal matrix. Because $\mathbf{A}$ is constant, we consider $\mathbf{R}$ instead of $\mathbf{W}$ and have

$$\begin{aligned} J(\mathbf{W}) = J(\mathbf{R}) &= \sum_{i=1}^{n} |\text{kurt}\{\sum_{j=1}^{n} r_{ij} s_j\}| = \sum_{i=1}^{n} |\sum_{j=1}^{n} r_{ij}^4 \text{kurt}\{s_j\}| \\ &= \sum_{i=1}^{n} |\sum_{j=1}^{n} r_{ij}^4 \kappa_j| = \sum_{i=1}^{n} k_i \sum_{j=1}^{n} r_{ij}^4 \kappa_j, \end{aligned} \tag{6}$$

where κ_j denotes the kurtosis of the j-th source signal, and

$$k_i = \text{sign}\{\sum_{j=1}^{n} r_{ij}^4 \kappa_j\}. \tag{7}$$

In the above equations, κ_j is unknown. Moreover, $\mathbf{R}$ is related with $\mathbf{W}$, but also unknown. However, with the samples of x we can directly estimate kurt$\{y_i\}$ and the kurtosis objective function. Since the absolute value of a function cannot be di erentiable at zero, we set k_i as a ± 1 coe cient, which leads to a kurtosis switching function.

2.2 Kurtosis Switching Algorithm

We further construct a kurtosis switching algorithm to maximize the kurtosis-sum objective function. Before doing so, we give an estimate of kurt$\{y_i\}$ with the samples from the observation. Actually, with a set of samples $\mathcal{D} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$, it is quite reasonable to use the following statistic:

$$f(\mathbf{w}_i|\mathcal{D}) = \frac{1}{N}\sum_{l=1}^{N}(\mathbf{w}_i^T\mathbf{x}_l)^4 - 3 \tag{8}$$

to estimate kurt$\{\mathbf{w}_i^T\mathbf{x}\}$.

With the above preparations, we can construct the kurtosis switching algorithm as follows.

(1) Initialization. The mixed signal $\mathbf{x}$ should be pre-whitened. $\mathbf{W}$ is initially set to be an orthogonal matrix, and k_i is set to be either 1 or -1.
(2) Select a sample data set $\mathcal{D}$ from the mixed signals.
(3) Evaluate the kurtosis values of the current estimated components, $f(\mathbf{w}_i|\mathcal{D})$ and update $k_i :=$ sign$\{f(\mathbf{w}_i|\mathcal{D})\}$, for $i = 1, \ldots, n$. (Note that this update is not always active in each iteration.)
(4) Calculate the gradient. Compute $\partial f(\mathbf{w}_i|\mathcal{D})/\partial\mathbf{w}_i$ for $i = 1, \ldots, n$, and set

$$\nabla J = \left[k_1\frac{\partial f(\mathbf{w}_1|\mathcal{D})}{\partial\mathbf{w}_1}, \cdots, k_n\frac{\partial f(\mathbf{w}_n|\mathcal{D})}{\partial\mathbf{w}_n}\right]. \tag{9}$$

(5) Obtain the constraint gradient. Project ∇J onto Stiefel manifold by

$$\hat{\nabla} J = \mathbf{W}\mathbf{W}^T\nabla J - \mathbf{W}\nabla J^T\mathbf{W}. \tag{10}$$

(6) Update $\mathbf{W} := \mathbf{W} + \eta\hat{\nabla}J$. Certain regularization process may be implemented on $\mathbf{W}$ if $\mathbf{W}$ is far from orthogonal.
(7) Repeat step (2) through (6), until $||\hat{\nabla}J|| < \varepsilon$, where $||\cdot||$ is the Euclidean norm and $\varepsilon (> 0)$ is a pre-selected threshold value for stopping the algorithm.

In this algorithm, the absolute value operator $|\cdot|$ is replaced by multiplying a switch coe cient $k_i = \pm 1$, which guarantees the maximization of the original kurtosis-sum objective function, because the kurtosis signs are always checked. Meanwhile, we utilize a modified gradient of the objective function w.r.t. $\mathbf{W}$, which automatically keeps the constraint $\mathbf{W}\mathbf{W}^T = \mathbf{I}$ satisfied after each update of $\mathbf{W}$, for small η.

With the kurtosis switching algorithm, we can lead to a local maximum of the kurtosis-sum objective function. We now analyze the no spurious solution property of the kurtosis-sum objective function for two-source case in the asymptotic sense. The two sources are required to have zero kurtosis.

Clearly, in the two-source case, $\mathbf{R}$ is a 2×2 orthogonal matrix, and can be parameterized by

$$\mathbf{R} = \begin{bmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{bmatrix} \text{ or } \begin{bmatrix}\cos\theta & \sin\theta \\ \sin\theta & -\cos\theta\end{bmatrix}. \tag{11}$$

Thus, we have

$$J(\mathbf{W}) = J(\mathbf{R}) = J(\theta) = |\kappa_1 \cos^4\theta + \kappa_2 \sin^4\theta| + |\kappa_1 \sin^4\theta + \kappa_2 \cos^4\theta|. \quad (12)$$

Below we analyze the local maxima of $J(\theta)$ for di erent signs of κ_1 and κ_2.
Case 1. If $\kappa_1 > 0$ and $\kappa_2 > 0$, or $\kappa_1 < 0$ and $\kappa_2 < 0$, we have

$$J(\theta) = (|\kappa_1| + |\kappa_2|)(\cos^4\theta + \sin^4\theta) = |\kappa_1 + \kappa_2|(\frac{3}{4} + \frac{1}{4}\cos 4\theta).$$

In this case the kurtosis of each source component of $\mathbf{s}$ is always positive. It is easily verified that $J(\theta)$ has local maxima only at $\theta \in \{m\pi/2\}, m \in \mathbb{K}$ which lead $\mathbf{R}$ to the following forms:

$$\mathbf{R} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ or } \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

where $\lambda_i \in \{\pm 1\}, i = 1, 2$. Certainly, all these $\mathbf{R}$, i.e., the local maxima, correspond to the feasible solutions of the ICA problem.

Case 2. If $\kappa_1 < 0$ and $\kappa_2 > 0$, the kurtosis signs of the two source components of $\mathbf{s}$ are di erent. In this case, $J(\theta)$ becomes a piecewise function as follows.

$$J(\theta) = \begin{cases} (\kappa_1 + \kappa_2)(\sin^4\theta + \cos^4\theta), & \text{if } \frac{\sin^4\theta}{\cos^4\theta} \geq -\frac{\kappa_1}{\kappa_2} \text{ and } \frac{\sin^4\theta}{\cos^4\theta} < -\frac{\kappa_2}{\kappa_1} \\ (-\kappa_1 - \kappa_2)(\sin^4\theta + \cos^4\theta), & \text{if } \frac{\sin^4\theta}{\cos^4\theta} < -\frac{\kappa_1}{\kappa_2} \text{ and } \frac{\sin^4\theta}{\cos^4\theta} \geq -\frac{\kappa_2}{\kappa_1} \\ (\kappa_1 - \kappa_2)(\cos^4\theta - \sin^4\theta), & \text{if } \frac{\sin^4\theta}{\cos^4\theta} \geq -\frac{\kappa_1}{\kappa_2} \text{ and } \frac{\sin^4\theta}{\cos^4\theta} \geq -\frac{\kappa_2}{\kappa_1} \\ (\kappa_2 - \kappa_1)(\cos^4\theta - \sin^4\theta), & \text{if } \frac{\sin^4\theta}{\cos^4\theta} < -\frac{\kappa_1}{\kappa_2} \text{ and } \frac{\sin^4\theta}{\cos^4\theta} < -\frac{\kappa_2}{\kappa_1} \end{cases}$$

For convenience, we define $\alpha = -\frac{\kappa_1}{\kappa_2}$ and $\phi = \tan^{-1}(\sqrt[4]{\min(\alpha, 1/\alpha)}) \leq \frac{\pi}{4}$. Then, the range of θ can be divided into three non-overlapping sets:

$$S_1 = \{\theta | \tan^4\theta \geq \max(\alpha, 1/\alpha)\} = \bigcup_{m=-\infty}^{+\infty} [m\pi + \frac{\pi}{2} - \phi, m\pi + \frac{\pi}{2} + \phi];$$

$$S_2 = \{\theta | \tan^4\theta < \min(\alpha, 1/\alpha)\} = \bigcup_{m=-\infty}^{+\infty} (m\pi - \phi, m\pi + \phi);$$

$$S_3 = \{\theta | \min(\alpha, 1/\alpha) \leq \tan^4\theta < \max(\alpha, 1/\alpha)\} = \mathbb{R} \setminus (S_1 \bigcup S_2).$$

We now consider θ in the three sets, respectively, as follows.

(a). If $\theta \in S_1$, $J(\theta) = (\kappa_1 - \kappa_2)(\cos^4\theta - \sin^4\theta) = -(\kappa_2 - \kappa_1)\cos 2\theta$ has local maxima only at $\{m\pi + \frac{\pi}{2}\}, m \in \mathbb{K}$, and $\inf_{\theta \in S_1} J(\theta) = -(\kappa_2 - \kappa_1)\cos(\pi - 2\phi) = (\kappa_2 - \kappa_1)\cos 2\phi$.

(b). If $\theta \in S_2$, $J(\theta) = (\kappa_2 - \kappa_1)(\cos^4\theta - \sin^4\theta) = (\kappa_2 - \kappa_1)\cos 2\theta$ has local maxima only at $\{m\pi\}, m \in \mathbb{K}$. And $\inf_{\theta \in S_2} J(\theta) = (\kappa_2 - \kappa_1)\cos 2\phi$

(c). If $\theta \in S_3$, $J(\theta) = (\kappa_1 + \kappa_2)(\sin^4\theta + \cos^4\theta)$ if $-\kappa_1 < \kappa_2$; or $J(\theta) = (-\kappa_1 - \kappa_2)(\sin^4\theta + \cos^4\theta)$ if $-\kappa_1 > \kappa_2$. So $J(\theta) = |\kappa_1 + \kappa_2|(\sin^4\theta + \cos^4\theta) =$

$|\kappa_1+\kappa_2|(\frac{3}{4}+\frac{1}{4}\cos 4\theta)$. It is easy to see that $J(\theta)$ has no local maximum within S_3, and $\sup_{\theta\in S_3} J(\theta) = |\kappa_1+\kappa_2|(\frac{3}{4}+\frac{1}{4}\cos 4\phi)$.

According to the above analysis, we have

$$
\begin{aligned}
\inf_{\theta\in S_1} J(\theta) = \inf_{\theta\in S_2} J(\theta) &= (\kappa_2-\kappa_1)\cos 2\phi = (\kappa_2-\kappa_1)(1-\tan^4\phi)\cos^4\phi \\
&= (\kappa_2-\kappa_1)(1-\min(-\frac{\kappa_1}{\kappa_2},-\frac{\kappa_2}{\kappa_1}))\cos^4\phi \\
&= (\kappa_2-\kappa_1)\frac{|\kappa_2+\kappa_1|}{\max(-\kappa_1,\kappa_2)}\cos^4\phi; \qquad (13)\\
\sup_{\theta\in S_3} J(\theta) &= |\kappa_1+\kappa_2|(\frac{3}{4}+\frac{1}{4}\cos 4\phi) = |\kappa_1+\kappa_2|(1+\tan^4\phi)\cos^4\phi \\
&= |\kappa_1+\kappa_2|(1+\min(-\frac{\kappa_1}{\kappa_2},-\frac{\kappa_2}{\kappa_1}))\cos^4\phi \\
&= |\kappa_1+\kappa_2|\frac{\kappa_2-\kappa_1}{\max(\kappa_2,-\kappa_1)}\cos^4\phi. \qquad (14)
\end{aligned}
$$

Because $\mathbb{R} = S_1\bigcup S_2\bigcup S_3$, and $\inf_{\theta\in S_1} J(\theta) = \inf_{\theta\in S_2} J(\theta) = \sup_{\theta\in S_3} J(\theta)$, $J(\theta)$ cannot reach any local maximum at the boundary points of S_3. Thus, $J(\theta)$ can have local maxima only at $\{m\pi/2\}, m\in\mathbb{K}$.

For the case $\kappa_1 > 0$ and $\kappa_2 < 0$, it can be easily verified that $J(\theta)$ behaves in the same way as in Case 2. Summing up all the analysis results, we have proved that in the two-source case, $J(\mathbf{W}) = J(\mathbf{R})$ can only have the local maxima that correspond to the feasible solutions of the ICA problem. That is, $J(\mathbf{W})$ is locally maximized only at a separation matrix $\mathbf{W}$ which leads $\mathbf{R}$ to a permutation matrix plus sign ambiguities.

From the above analysis, we can find that when the sources with positive kurtosis and negative kurtosis co-exist, the range of $\mathbf{R}$ (corresponding to a unit circle of θ) can be divided into some non-overlapping sets and on each of them, the kurtosis signs of y_i does not change. Thus, the update of the kurtosis sign of y_i in each iteration is not necessarily active. In fact, a real switching operation happens only when the parameter moves across the boundaries of two such sets.

In order to substantiate our theoretical results and test the kurtosis switching algorithm, we conducted two experiments on real and artificial signals. We also compared the results of our algorithm with those of the Extended Infomax algorithm [11] and FastICA algorithm [20].

Firstly, we utilized two audio recordings as independent source signals. Each of these two signals contain 4000 samples and their sample kurtosis are 0.6604 and 0.1910, respectively. The observation signals were generated as two linear mixtures of these two audio signals through a random matrix. We implemented the kurtosis switching algorithm on the observation signals. After the kurtosis switching algorithm stopped, it was found that the two sources were separated

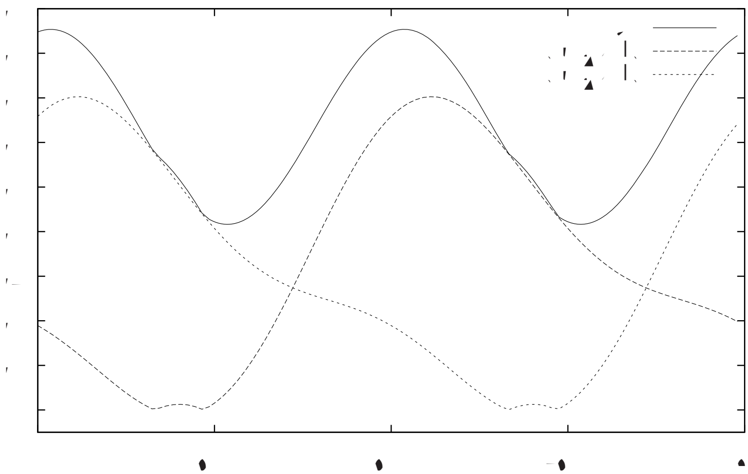

Fig. 1. ... of ... $J(\theta)$ ... Abso... of $\mathbf{y}$ ... for θ fro... π

with

$$\mathbf{R} = \begin{bmatrix} 1.0021 & -0.0280 \\ -0.0233 & -1.0020 \end{bmatrix}.$$

Actually, the performance index (refer to [4]) of this separation result was 0.1024. In the same situation, the Extended Infomax algorithm could arrive only at a performance index 0.4658. For the FastICA algorithm, the symmetric approach was selected, and the performance index was 0.1215 by using “tanh” as nonlinearity, but improved to 0.0988 by using “power 3”. As a result, on the correctness of the ICA solution, the kurtosis switching algorithm could be as good as the FastICA algorithm, although it required more iterations and took much longer time than FastICA.

For illustration, we further show the sketches of the kurtosis-sum objective function and the absolute kurtosis values of the two estimated components of $\mathbf{y}$ in the above two-source experiment for θ from zero to π in Fig. 1. Theoretically, as the two sources are super-Gaussian, their mixtures should have positive kurtosis. However, the estimated kurtosis of y_i could be negative at some θ or $\mathbf{W}$. Besides, our analysis indicates that either of $|\mathrm{kurt}\{y_i\}|$ is the maximum at $\theta = n\pi/2$, but it is not so for finite data. Actually, the maxima of the kurtosis-sum objective function were not exactly at $\theta = n\pi/2$, due to the errors from the estimation.

We further conducted another experiment on seven synthetic sources: random samples generated from (a). Laplacian distribution, (b). Exponential distribution which is not symmetric, (c). Uniform distribution, (d). Beta distribution $\beta(2,2)$, (e). A Gaussian mixture (bimodal): $\frac{1}{2}N(-1.5, 0.25) + \frac{1}{2}N(1.5, 0.25)$, (f). A Gaussian mixture (unimodal): $\frac{1}{2}N(0, 0.25) + \frac{1}{2}N(0, 2.25)$, (g). A Gaussian mixture (trimodal): $\frac{1}{3}N(-2, 0.25) + \frac{1}{3}N(0, 0.25) + \frac{1}{3}N(2, 0.25)$. Three of them ((a),

(b) and (f)) were super-Gaussian while the rest four sources were sub-Gaussian. All the sources were normalized before mixing. For each source, there were 1000 samples. The observation signals were generated as seven linear mixtures of these seven independent synthetic signals through a random matrix. We implemented the kurtosis switching algorithm on these observation signals and obtained a successful separation matrix with $\mathbf{R}$ being given as follows:

$$\mathbf{R} = \begin{bmatrix} -0.0205 & 0.0141 & 0.0139 & -0.0223 & -0.0361 & 0.0397 & 1.0231 \\ -0.0178 & 0.0008 & -1.0108 & 0.0049 & 0.0015 & 0.0106 & 0.0697 \\ 1.0103 & 0.0121 & -0.0088 & -0.0071 & -0.0688 & 0.0019 & 0.0103 \\ -0.0333 & -0.0398 & 0.0034 & -1.0165 & -0.0057 & -0.0130 & -0.0106 \\ -0.0740 & 0.0320 & -0.0114 & 0.0392 & -1.0097 & -0.0138 & -0.0378 \\ -0.0179 & 1.0062 & 0.0085 & -0.0614 & 0.0200 & 0.0430 & -0.0017 \\ -0.0232 & 0.0575 & -0.0059 & 0.0466 & -0.0407 & -1.0112 & 0.0111 \end{bmatrix}.$$

According to $\mathbf{R}$, we obtained that the performance index of the kurtosis switching algorithm was 2.0003. In the same situation, the FastICA algorithm's performance index was 1.9542 when using "power 3" as nonlinearity, but became 1.3905 when using "tanh". However, the Extended Infomax algorithm did not separate all the sources, with a performance index of 15.4736. Therefore, in this complicated case with seven sources, the kurtosis switching algorithm achieved a separation result almost as good as the FastICA algorithm though it required more steps to converge, but outperformed the Extended Infomax algorithm. Moreover, this experimental result also demonstrated that our theoretical results on the kurtosis-sum objective function can be extended to the cases with more than 2 sources.

Besides the two demonstrations above, we have conducted many simulations, with various types of signal sources. All the experimental results conformed to the theoretical analysis and no spurious solutions have been encountered.

We have investigated the ICA problem through the kurtosis-sum objective function which is just the sum of absolute kurtosis values of the estimated components. Actually, we prove that for two-source case, the maxima of this kurtosis-sum objective function all correspond to the feasible solutions of the ICA problem, as long as the sources have non-zero kurtosis. Moreover, in order to maximize the kurtosis-sum objective function, a kurtosis switching algorithm is constructed. The experimental results show that the kurtosis-sum objective function works well for solving the ICA problem and apart from the convergence speed, the kurtosis switching algorithm can arrive at a solution as good as the FastICA algorithm.

Acknowledgements. This work was supported by the Ph.D. Programs Foundation of Ministry of Education of China for grant 20070001042.

1. 36, 287 314 (1994)
2. 4, 112 114 (1997)
3. 7, 1129 1159 (1995)
4. 8, 757 763 (1996)
5. 22, 69 80 (1998)
6. 16, 383 399 (2004)
7. 17, 331 334 (2005)
8. 2006. 3889, pp. 838 845. (2006)
9. 2006. 3889, pp. 173 180. (2006)
10. 12, 190 193 (2005)
11. 11, 417 441 (1999)
12. 15, 233 243 (2004)
13. 2005. 3801, pp. 915 920. (2005)
14. 13, 677 689 (2001)
15. 15, 55 64 (2004)
16. 10, 1345 1351 (1997)
17. 44, 3017 3030 (1996)
18. 45, 59 83 (1995)
19. 9, 1483 1492 (1997)
20. 10, 626 634 (1999)