An Efficient Pairwise Kurtosis Optimization Algorithm for Independent Component Analysis

Fei Ge and Jinwen Ma*

Department of Information Science, School of Mathematical Sciences and LMAM, Peking University, Beijing, 100871, China
felixge@math.pku.edu.cn, jwma@math.pku.edu.cn

Abstract. In the framework of Independent Component Analysis (ICA), kurtosis has been used widely in designing source separation algorithms. In fact, the sum of absolute kurtosis values of all the output components is an effective objective function for separating arbitrary sources. In this paper, we propose an efficient ICA algorithm via a modified Jacobi optimization procedure on the kurtosis-sum objective function. The optimal rotation angle for any pair of the output components can be solved directly. It is demonstrated by numerical simulation experiments that our proposed algorithm can be even more computationally efficient than the FastICA algorithm under the same separation performance.

Keywords: Independent Component Analysis, kurtosis, pairwise optimization, Jacobi algorithm.

1 Introduction

In statistical signal processing and data analysis, finding a new set of coordinates of the observed multi-dimensional data such that the components are as independent as possible may help to discover the underlying structure of the data. This problem is known as blind source separation (BSS) [1] or independent component analysis (ICA) [2]. Regardless of the temporal dependence, the observed data are assumed to be independently generated via a linear transformation of certain independent sources. This model is mathematically expressed by $\mathbf{x} = \mathbf{A}\mathbf{s}$, where $\mathbf{x}$ and $\mathbf{s}$ are random vectors denoting the observations and sources respectively, while $\mathbf{A}$ is a constant mixing matrix of full column rank.

Under this model, any matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{\Lambda P}$ ($\mathbf{\Lambda}$ denotes a diagonal matrix for scaling and $\mathbf{P}$ denotes a permutation matrix) is a feasible solution to the ICA problem, i.e., it is a separating matrix. By optimizing some contrast function [2] of the transformed observations $\mathbf{y} = \mathbf{W}\mathbf{x}$, a separating matrix can be obtained iteratively.

Many criteria can lead to a feasible solution, for example, information maximization [3], minimum mutual information [4] and maximum likelihood [1]. However, there is still a practical difficulty that the knowledge of the observation distribution is limited: only a sample set is available, and no density model

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2010, CCIS 93, pp. 94–101, 2010.

can be employed in general. One possible way is to resort to third- and fourth-order cumulants and approximate the probability density function (pdf) by its truncated Edgeworth expansion [2] or Gram-Charlier expansion [4], as sample cumulants can be easily computed. These treatments lead to the high-order ICA methods [5].

Despite the inaccuracy of the pdf's approximation, the high-order methods are applicable since all cross cumulants vanish as the source components are mutually independent. Comon [2] proved that the sum of squares of fourth-order marginal standardized cumulants (kurtoses) is a contrast function. Remarkably, by maximizing the square of fourth order marginal cumulant, Delfosse and Loubaton [6] proposed a deflationary separation scheme to sequentially extract the source signals, with theoretically guaranteed convergence to correct solutions. Unfortunately, the correct convergence in the deflation mode does not guarantee the same result in symmetrical (simultaneous separation) form.

Practically, it is more significant to consider the symmetrical separation of the sources. In our previous work [7], we analyzed the sum of absolute values of kurtoses of the output components as a contrast or objective function for whitened observations. We referred it as kurtosis-sum objective function and proved that it has no spurious maxima for two-source mixing problem. Actually, Moreau and Macchi [8] also studied this objective function and proposed a self-adaptive maximization algorithm, which parameterized the demixing matrix via planar rotations but was too complicated for implementation.

In this paper, we propose an efficient pairwise Jacobi optimization algorithm for ICA via a modified Jacobi procedure on the kurtosis-sum objective function. For maximizing the sum of absolute values of kurtoses of any pair of output components, the problem is reduced to finding an optimal rotation angle, which can be solved directly from data. Following the Jacobi procedure [5], the kurtosis-sum objective function can be maximized by a series of planar rotations. To save the computation cost, we make some modifications on it. The simulation experiments show that our proposed ICA algorithm is more computationally efficient than the FastICA [9] algorithm under the same separation performance.

In the sequel, maximizing the kurtosis-sum objective function in the pairwise processing mode is introduced in Section 2. In Section 3, we present the pairwise kurtosis optimization algorithm for ICA in the form of the modified Jacobi optimization procedure. Then in Section 4 the computation complexity is analyzed and compared with the FastICA algorithm with third power nonlinearity via simulation experiments. We finally give a brief conclusion in Section 5.

2 Pairwise Optimization of the Kurtosis-Sum

To perform source separation from whitened observation $\mathbf{x}$, we previously proposed a kurtosis switching algorithm [7] to maximize the kurtosis-sum objective function:

$$J(\mathbf{W}) = \sum_{i=1}^{n} |\mathrm{kurt}\{y_i\}| = \sum_{i=1}^{n} |E\{(\mathbf{w}_i^T \mathbf{x})^4 - 3\}| \ , \tag{1}$$

under the constraint $\mathbf{W}^T\mathbf{W} = \mathbf{I}$, where kurt$\{\cdot\}$ and $E\{\cdot\}$ denote the kurtosis and expectation of a random variable, respectively. Our previous algorithm was gradient-style, and extra step must be utilized to retain the orthogonality of $\mathbf{W}$. Since $\mathbf{W}$ is orthogonally constrained, it is an alternative to consider pairwise optimization, i.e., only two rows of $\mathbf{W}$ are updated in each step by a Givens rotation [5]. This procedure is also known as Jacobi algorithm, and all pairs of rows must be processed repeatedly, until an optimal state is reached.

Thanks to the simple additive structure of Eq. (1), when only two rows of $\mathbf{W}$ are updated under the orthogonal constraint, that is, when the l-th and k-th rows are rotated in a 2-D plane:

$$\begin{cases} \mathbf{w}_l' = \mathbf{w}_l \cos\theta + \mathbf{w}_k \sin\theta \ , \\ \mathbf{w}_k' = -\mathbf{w}_l \sin\theta + \mathbf{w}_k \cos\theta \ , \end{cases} \tag{2}$$

the $n-2$ terms $|\text{kurt}\{y_i\}|$ $(i \neq l,\, i \neq k)$ in Eq. (1) will remain unchanged. Thus for processing only the (l,k) pair, maximizing Eq. (1) is reduced to maximize

$$\begin{aligned} J_{l,k}(\theta) &= |E\{(\mathbf{w}_l'^T\mathbf{x})^4 - 3\}| + |E\{(\mathbf{w}_k'^T\mathbf{x})^4 - 3\}| \\ &= |E\{(y_l\cos\theta + y_k\sin\theta)^4\} - 3| + |E\{(-y_l\sin\theta + y_k\cos\theta)^4\} - 3|. \end{aligned} \tag{3}$$

This function is obviously piecewise smooth. To find an analytic solution of its maximum, we define

$$J_{\text{I}}(\theta) = E\{(y_l\cos\theta + y_k\sin\theta)^4 + (y_k\cos\theta - y_l\sin\theta)^4\} - 6 \ , \tag{4}$$

$$J_{\text{II}}(\theta) = E\{(y_l\cos\theta + y_k\sin\theta)^4 - (y_k\cos\theta - y_l\sin\theta)^4\} \ , \tag{5}$$

and consider an alternative form:

$$\hat{\theta} = \arg\max_{\theta} \max\{|J_{\text{I}}(\theta)|, |J_{\text{II}}(\theta)|\} \tag{6}$$

$$= \begin{cases} \arg\max_\theta |J_{\text{I}}(\theta)|, & \text{if } \max_\theta |J_{\text{I}}(\theta)| > \max_\theta |J_{\text{II}}(\theta)|, \\ \arg\max_\theta |J_{\text{II}}(\theta)|, & \text{otherwise} \ . \end{cases} \tag{7}$$

It turns out that $J_{\text{I}}(\theta)$ and $J_{\text{II}}(\theta)$ are just sinusoidal functions as follows:

$$J_{\text{I}}(\theta) = A\sin(4\theta + \alpha) + c \ , \tag{8}$$

$$J_{\text{II}}(\theta) = B\sin(2\theta + \beta) \ , \tag{9}$$

with $A \geq 0, B \geq 0, c, \alpha, \beta$ being the parameters related to y_l and y_k. Note that when the expectations in Eq. (3) are replaced by the sample averages, these expressions are still valid. One convenient way of determining these parameters is to connect Eqs. (4) (5) with Eqs. (8) (9) and set $\theta = 0$ to obtain

$$J_{\text{I}}(\theta)|_{\theta=0} = E\{y_l^4 + y_k^4\} - 6 = A\sin(\alpha) + c \ , \tag{10}$$

$$J_{\text{II}}(\theta)|_{\theta=0} = E\{y_l^4 - y_k^4\} = B\sin(\beta) \ , \tag{11}$$

$$J_{\text{I}}'(\theta)|_{\theta=0} = E\{4y_l^3 y_k - 4y_k^3 y_l\} = 4A\cos(\alpha) \ , \tag{12}$$

$$J_{\text{II}}'(\theta)|_{\theta=0} = E\{4y_l^3 y_k + 4y_k^3 y_l\} = 2B\cos(\beta) \ , \tag{13}$$

$$J_{\text{I}}''(\theta)|_{\theta=0} = E\{24y_l^2 y_k^2 - 4y_k^4 - 4y_l^4\} = -16A\sin(\alpha) \ . \tag{14}$$

The above five equations are enough to determine $J_{\mathrm{I}}(\theta)$ and $J_{\mathrm{II}}(\theta)$. The parameters can be solved as follows:

$$c = \frac{3}{4}E\{y_l^4 + y_k^4\} + \frac{3}{2}E\{y_l^2 y_k^2\} - 6 \ , \tag{15}$$

$$A = \sqrt{(E\{y_l^4 + y_k^4\} - 6 - c)^2 + (E\{y_l^3 y_k - y_k^3 y_l\})^2} \ , \tag{16}$$

$$B = \sqrt{(E\{y_l^4 - y_k^4\})^2 + (2E\{y_l^3 y_k + y_k^3 y_l\})^2} \ , \tag{17}$$

$$\alpha = \begin{cases} \sin^{-1}((E\{y_l^4 + y_k^4\} - 6 - c)/A), & \text{if } \cos\alpha > 0, \\ \pi - \sin^{-1}((E\{y_l^4 + y_k^4\} - 6 - c)/A), & \text{otherwise}, \end{cases} \tag{18}$$

$$\beta = \begin{cases} \sin^{-1}(E\{y_l^4 - y_k^4\}/B), & \text{if } \cos\beta > 0, \\ \pi - \sin^{-1}(E\{y_l^4 - y_k^4\}/B), & \text{otherwise}. \end{cases} \tag{19}$$

Finally we need to compare $J_{\mathrm{I}}(\theta)$ and $J_{\mathrm{II}}(\theta)$, and then locate the maximum of Eq. (3). According to Eq. (7), if $|c| + A > B$, we should choose

$$\hat{\theta} = \arg\max_{\theta} |J_{\mathrm{I}}(\theta)| = \begin{cases} (\frac{\pi}{2} - \alpha)/4, & \text{if } c \geq 0, \\ (-\frac{\pi}{2} - \alpha)/4, & \text{if } c < 0; \end{cases} \tag{20}$$

otherwise we should choose

$$\hat{\theta} = \arg\max_{\theta} |J_{\mathrm{II}}(\theta)| = (\frac{\pi}{2} - \beta)/2, \tag{21}$$

then $\hat{\theta}$ must be the optimal rotating angle for the sub-problem. In summary, it requires the five fourth-order moments

$$E\{y_l^4\},\ E\{y_k^4\},\ E\{y_l^3 y_k\},\ E\{y_l y_k^3\} \text{ and } E\{y_l^2 y_k^2\} \tag{22}$$

to calculate a closed-form solution of $\hat{\theta}$. Practically the sample moments will be used instead. In the following we denote the corresponding sample moments of Eq. (22) by $\hat{\mu}_{4,0}$, $\hat{\mu}_{0,4}$, $\hat{\mu}_{3,1}$, $\hat{\mu}_{1,3}$ and $\hat{\mu}_{2,2}$, respectively.

The angles α and β can also be calculated by the inverse tangent function. If $\hat{\theta}$ is determined via Eq. (20), the solution is equivalent to the EML estimator [10]. If $\hat{\theta}$ is determined via Eq. (21), it is equivalent to the AEML estimator [11]. However, EML and AEML estimators were derived under two-source mixing setting. Our solution happens to be a hybrid of these two estimators, and Eq. (7) serves as some switching mechanism. Unlike our previous algorithm [7], there is no need to evaluate the "switching coefficients" separately.

To demonstrate, we sketch in Fig. 1 the objective function (3) as well as $J_{\mathrm{I}}(\theta)$, $J_{\mathrm{II}}(\theta)$ and the calculated optimal angle, from a pairwise step of a four-source mixing problem. The sample moments were $\hat{\mu}_{4,0} = 2.6166$, $\hat{\mu}_{0,4} = 3.0539$, $\hat{\mu}_{3,1} = 0.061506$, $\hat{\mu}_{1,3} = 0.22054$ and $\hat{\mu}_{2,2} = 1.0809$, calculated using all 4000 samples.

3 Pairwise Kurtosis Optimization Algorithm

For n output components, a full Jacobi sweep consists in $n(n-1)/2$ rotation attempts for all pairs of components. If $\hat{\theta}$ is not sufficiently close to 0 or $\pm\pi/2$,

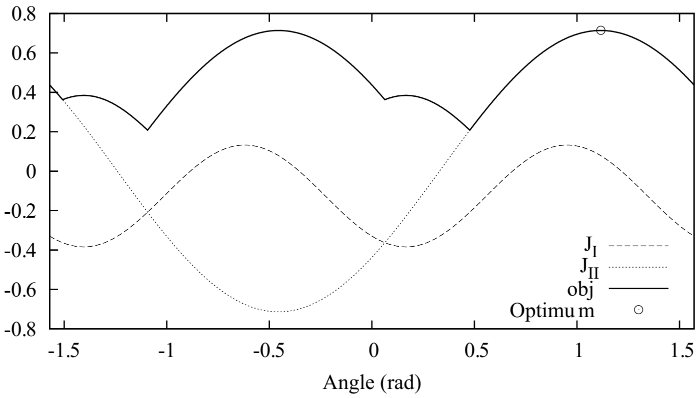

Fig. 1. Example of pairwise sub-problem solution

this pair is then rotated. A standard Jacobi algorithm performs Jacobi sweeps one after another, until no pair has been rotated in one sweep.

For maximizing the kurtosis-sum objective function via pairwise processing in batch mode, the overall computation cost is almost the cost of calculating the sample moments and making changes (rotations) to the data. The standard Jacobi algorithm has an obvious drawback: when optimizing a particular pair (l, k), if the neither the data $y_l(t)$ nor the data $y_k(t)$ has been changed since the latest rotation of this pair, this pair is already optimal and calculating $\hat{\theta}$ must yield 0, $\pi/2$ or $-\pi/2$, which is useless.

To avoid such unnecessary calculations, we can keep a flag $F(l, k)$ for each pair (l, k) and check it before a rotation attempt. If the flag is already marked as "done", do not calculate $\hat{\theta}$; otherwise mark this flag as "done" and continue as usual. Whenever a rotation is applied to a pair (p, q), clear all flags $F(p, i)$ and $F(j, q)$ (for all $i \neq p$, $j \neq p$) because this rotation affects some other pairs. The terminating criterion for the Jacobi algorithm becomes when the flags of all pairs are marked as "done". This modification does not change the behaviour of the Jacobi algorithm.

In simulation experiments we also observed that a slight rotation is not likely to affect other pairs that are at optimal state, thus we added an heuristic rule to further save some computation cost, without significant degradation of separation quality. This rule is that if $\hat{\theta}$ is not far from 0, $\pi/2$ or $-\pi/2$ (but rotation is still needed), do not clear any flags. To summarize, our pairwise kurtosis optimization algorithm via the modified Jacobi procedure can be listed as follows:

INPUT: whitened observation data $\{\mathbf{y}(t)\}$, $t = 1, \ldots, N$.

(i) *Initialize*: set the flags $F(i, j) = 0$, set $\mathbf{h} = (h_1, \ldots, h_n)^T = \mathbf{0}$.

(ii) Perform a Jacobi sweep though each pair in the order $(1, 2), (1, 3), \ldots, (1, n), (2, 3), \ldots, (2, n), \ldots, (n-1, n)$. For each pair (p, q), do

(a) If $F(p, q) = 1$, return to step (ii) and select the next pair.

(b) Calculate $a(t) = y_p(t)^2$, $b(t) = y_q(t)^2$, $c(t) = y_p(t)y_q(t)$ $(t = 1, \ldots, N)$.
(c) If $h_p = 0$, calculate $\hat{\mu}_{4,0} = \sum_{t=1}^{N} a(t)^2/N$, otherwise set $\hat{\mu}_{4,0} = h_p$.
If $h_q = 0$, calculate $\hat{\mu}_{0,4} = \sum_{t=1}^{N} b(t)^2/N$, otherwise set $\hat{\mu}_{0,4} = h_q$.
Calculate $\hat{\mu}_{2,2} = \sum_{t=1}^{N} c(t)^2/N$, $\hat{\mu}_{3,1} = \sum_{t=1}^{N} a(t)c(t)/N$,
$\hat{\mu}_{1,3} = \sum_{t=1}^{N} b(t)c(t)/N$.
(d) Obtain $\hat{\theta}$ according to Eqs. (15)-(21). Adjust $\hat{\theta}$ such that $|\hat{\theta}| \leq \pi/4$.
(e) If $|\hat{\theta}| < \theta_{\min}$, go to (h).
(f) Rotate the data: $\{y_p(t), y_q(t)\}' = \{y_p(t)\cos\hat{\theta} + y_q(t)\sin\hat{\theta},$
$-y_p(t)\sin\hat{\theta} + y_q(t)\cos\hat{\theta}\}$ $(t = 1, \ldots, N)$.
Calculate $h_p = \hat{\mu}_{4,0}\cos^4\hat{\theta} + 4\hat{\mu}_{3,1}\cos^3\hat{\theta}\sin\hat{\theta} + 6\hat{\mu}_{2,2}\cos^2\hat{\theta}\sin^2\hat{\theta}$
$+4\hat{\mu}_{1,3}\cos\hat{\theta}\sin^3\hat{\theta} + \hat{\mu}_{0,4}\sin^4\hat{\theta} - 3$,
and $h_q = \hat{\mu}_{4,0}\sin^4\hat{\theta} - 4\hat{\mu}_{3,1}\sin^3\hat{\theta}\cos\hat{\theta} + 6\hat{\mu}_{2,2}\cos^2\hat{\theta}\sin^2\hat{\theta}$
$-4\hat{\mu}_{1,3}\sin\hat{\theta}\cos^3\hat{\theta} + \hat{\mu}_{0,4}\cos^4\hat{\theta} - 3$.
(g) If $|\hat{\theta}| > \theta_{\text{tol}}$, clear all flags $F(p, i)$ $(i \neq q)$ and $F(j, q)$ $(j \neq p)$.
(h) Set $F(p, q) = 1$.

(iii) If $F(i, j) = 1$ for all $j > i$, stop. Otherwise go to (ii).

In the above algorithm, $\theta_{\min}$ is the threashold parameter for any Givens rotation, and $\theta_{\text{tol}} > \theta_{\min}$ enables our heuristic rule. The vector $\mathbf{h}$ stores the calculated marginal fourth order sample moments, which can be computed with minimum cost using older values of the moments, by expanding

$$E\{y_p'^4\} = E\{(y_p\cos\hat{\theta} + y_q\sin\hat{\theta})^4\}, \quad E\{y_q'^4\} = E\{(-y_p\sin\hat{\theta} + y_q\cos\hat{\theta})^4\} \ . \quad (23)$$

4 Computation Complexity and Simulation

Now we analyze the computation complexity of our proposed algorithm. The computation cost is evaluated in flops (floating point operation, defined as a multiplication followed by an addition [2]). Considering $N \gg n$, then one full Jacobi sweep approximatively needs $5Nn(n-1)$ flops: $3N$ flops in step (ii.b), another $3N$ flops in step (ii.c) only except for several initial steps requiring $5N$ or $4N$, and $4N$ flops in step (ii.f). But our pairwise kurtosis optimization algorithm will skip some computation steps, so the total number of flops varies in each run and must be checked by simulation.

For comparison, we also consider the famous FastICA algorithm [9] in symmetric mode with the third power nonlinearity. The computation cost of FastICA is almost spent on the fixed-point iteration step:

$$\mathbf{w}_i^+ = \frac{1}{N}\sum_{t=1}^{N}[\mathbf{x}(t)y_i(t)^3] - 3\mathbf{w}_i \ , \quad (24)$$

which needs $2nN + 2N$ flops. So updating the whole matrix $\mathbf{W}$ requires $n(2n + 2)N$ flops. The FastICA algorithm usually can converge in less than twenty iterations, which is very efficient.

Table 1. Source separation results from 100 Monte Carlo simulation runs

	Algorithm	ISR			Flops
		0.25 quantile	median	0.75 quantile	(average)
$n = 4$	FastICA	-27.60 dB	-24.77 dB	-21.76 dB	8.92×10^5
	pairwise	-27.66 dB	-24.79 dB	-21.84 dB	6.635×10^5
$n = 8$	FastICA	-21.38 dB	-19.74 dB	-18.51 dB	4.356×10^6
	pairwise	-21.48 dB	-19.88 dB	-18.51 dB	3.927×10^6
$n = 16$	FastICA	-16.86 dB	-15.76 dB	-14.43 dB	2.614×10^7
	pairwise	-16.95 dB	-15.94 dB	-14.88 dB	2.207×10^7

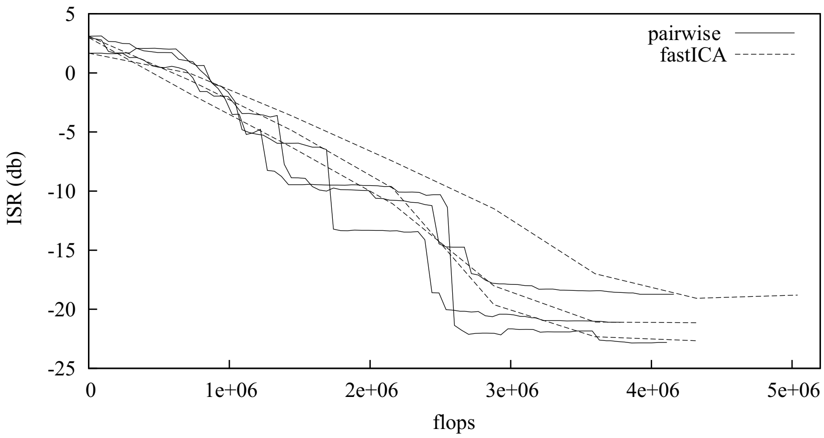

Fig. 2. Typical performance curves with 8 arbitary sources

We conducted simulation experiments using randomly generated data as sources. For each source used in one simulation, $N = 5000$ samples was generated from (arbitrarily chosen) one of the following distributions: (1) uniform distribution, (2) binary $\{-1, 1\}$ with equal probability, (3) Beta distribution Be$(2, 2)$, (4) a simple Gaussian mixture with $p(x) = \frac{1}{2\sqrt{\pi}} e^{-(x-\sqrt{2}/2)^2} + \frac{1}{2\sqrt{\pi}} e^{-(x+\sqrt{2}/2)^2}$, (5) Laplacian distribution, (6) hyperbolic secant distribution, (7) t-distribution with freedom 5, (8) t-distribution with freedom 13. The sources were normalized and the mixing matrix was a randomly generated orthogonal matrix, in order to bypass the influence of prewhitening step.

We tested three different dimensions when $n = 4$, $n = 8$ and $n = 16$, respectively. In each simulation, the pairwise algorithm and the FastICA algorithm were tested with the same mixture data. 100 Monte Carlo runs were performed for each dimension, and Table 1 summarizes the averaged performances. The separation quality is measured by average Interference-to-Signal power Ratio (ISR), which is a function of $\mathbf{R} = \mathbf{WA}$ in the simulation:

$$\mathrm{ISR}(\mathbf{R}) = 10 \log_{10} \left\{ \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\sum_{j=1}^{n} r_{ij}^2}{\max_{1 \le j \le n} r_{ij}^2} - 1 \right) \right\} . \tag{25}$$

Typical performance curves were sketched in Fig. 2, for three different simulations with $n = 8$.

These experimental results were obtained when $\theta_{\min} = 0.0025$ and $\theta_{\mathrm{tol}} = 0.025$. The separation quality of the two algorithms were almost identical. Comparing the flops in each run we can find that our pairwise algorithm is more numerically efficient.

5 Conclusions

We have proposed a pairwise kurtosis optimization approach through the Jacobi optimization procedure to maximize the kurtosis-sum objective function for ICA. This is an extension to our previous work. The closed-form solution of the optimal rotation angle makes the pairwise optimization possible and effective. The modification to the standard Jacobi scheme can save a portion of the computation cost. Simulation results confirm that our proposed algorithm is numerically efficient, comparable to the FastICA and maybe faster than it.

Acknowledgements

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

References

1. Cardoso, J.F.: Blind Signal Separation: Statistical Principles. Proceedings of the IEEE 86, 2009–2025 (1998)
2. Comon, P.: Independent Component Analysis – a New Concept? Signal Processing 36, 287–314 (1994)
3. Bell, A., Sejnowski, T.: An Information-Maximization Approach to Blind Separation and Blind Deconvolution. Neural Computation 7, 1129–1159 (1995)
4. Amari, S.I., Cichocki, A., Yang, H.: A New Learning Algorithm for Blind Separation of Sources. Advances in Neural Information Processing 8, 757–763 (1996)
5. Cardoso, J.F.: High-order contrasts for Independent Component Analysis. Neural Computation 11, 157–192 (1999)
6. Delfosse, N., Loubaton, P.: Adaptive Blind Separation of Independent Sources: a Deflation Approach. Signal Processing 45, 59–83 (1995)
7. Ge, F., Ma, J.: Analysis of the Kurtosis-Sum Objective Function for ICA. In: Sun, F., Zhang, J., Tan, Y., Cao, J., Yu, W. (eds.) ISNN 2008, Part I. LNCS, vol. 5263, pp. 579–588. Springer, Heidelberg (2008)
8. Moreau, E., Macchi, O.: High-order Contrast for Self-adaptive Source Separation. International Journal of Adaptive Control and Signal Processing 10, 19–46 (1996)
9. Hyvärinen, A.: Fast and Robust Fixed-point Algorithms for Independent Component Analysis. IEEE Trans. Neural Networks 10, 626–634 (1999)
10. Zarzoso, V., Nandi, A.K.: Blind Separation of Independent Sources for Virtually any Source Probability Density Function. IEEE Trans. Signal Processing 47, 2419–2432 (1999)
11. Zarzoso, V., Nandi, A.K., Herrmann, F., Millet-Roig, J.: Combined Estimation Scheme for Blind Source Separation with Arbitary Source PDFs. Electronic Letters 37, 132–133 (2001)