



# Hybrid linear and nonlinear complexity pursuit for blind source separation<sup>☆</sup>

Zhenwei Shi<sup>a,\*</sup>, Hongjuan Zhang<sup>b</sup>, Zhiguo Jiang<sup>a</sup>

<sup>a</sup> Image Processing Center, School of Astronautics, Beihang University, Beijing 100191, PR China

<sup>b</sup> Department of Mathematics, Shanghai University, Shanghai 200444, PR China

## ARTICLE INFO

### Article history:

Received 22 October 2010

Received in revised form 8 February 2012

### Keywords:

Blind source separation (BSS)

Independent component analysis (ICA)

Linear predictability

Nonlinear predictability

## ABSTRACT

Blind source separation (BSS) is an increasingly popular data analysis technique with many applications. Several methods for BSS using the statistical properties of original sources have been proposed; for a famous case, non-Gaussianity, this leads to independent component analysis (ICA). In this paper, we propose a hybrid BSS method based on linear and nonlinear complexity pursuit, which combines three statistical properties of source signals: non-Gaussianity, linear predictability and nonlinear predictability. A gradient learning algorithm is presented by minimizing a loss function. Simulations verify the efficient implementation of the proposed method.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

Blind source separation (BSS) [1,2] is an emerging data analysis technique used in many practical applications such as speech and image processing, biomedical signal processing, wireless telecommunication systems, economic data analysis, data mining, etc. The main objective of BSS is to recover unknown original source signals from their mixtures without knowing the mixing channels, using some statistical properties of the original sources. Suppose that  $n$  unknown sources are mixed simultaneously in a linear mixing channel modeled as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad (1)$$

where  $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T$  denotes the  $n$ -dimensional observation vector,  $\mathbf{A}$  is the  $n \times n$  unknown nonsingular constant mixing matrix, and  $\mathbf{s}(t) = (s_1(t), \dots, s_n(t))^T$  is the  $n$ -dimensional vector of unknown zero-mean and unit-variance original sources. The task of BSS is to recover the sources  $\mathbf{s}(t) = (s_1(t), \dots, s_n(t))^T$  from the mixtures  $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T$ .

The BSS problem has been studied by researchers in applied mathematics, neural networks and statistical signal processing. Several methods for BSS using the statistical properties of original sources have been proposed; areas encompassed have included non-Gaussianity (or equivalently, independent component analysis, ICA) [1–10], and time-structure information, such as linear predictability or smoothness [1,11], linear autocorrelation [12–14], coding complexity [15–19], temporal predictability [20], nonstationarity [21–23], energy predictability [24], nonlinear innovation [25], nonlinear autocorrelation [26–28], etc.

<sup>☆</sup> The work was supported by the National Natural Science Foundation of China under Grants 60975003, 11126057 and 91120301, the 973 Program under Grant 2010CB327904, the Fundamental Research Funds for the Central Universities under Grants YWF-10-01-A10 and YWF-11-03-Q-066, the Beijing Municipal Natural Science Foundation (Non-negative Component Analysis for Hyperspectral Imagery Unmixing) under Grant 4112036, Shanghai Leading Academic Discipline Project (J50101), and the Key Disciplines of Shanghai Municipality (Operations Research and Cybernetics, S30104).

\* Corresponding author. Tel.: +86 10 823 16 502; fax: +86 10 823 38 798.

E-mail address: [shizhenwei@buaa.edu.cn](mailto:shizhenwei@buaa.edu.cn) (Z. Shi).

In this paper, we present a hybrid technique for BSS based on linear and nonlinear complexity pursuit, which combines three statistical properties of source signals: non-Gaussianity, linear predictability and nonlinear predictability. First, we propose a contrast function for BSS based on the hybrid technique, and we perform the optimization using a gradient descent algorithm (Section 2). Then, we show how the contrast function is connected to other BSS contrast functions (Section 3). Simulation results show that the model separates sources in cases where existing methods are not able to do so (Section 4), and finally we conclude the paper (Section 5).

## 2. The proposed algorithm

Assume that the measured sensor signals  $\mathbf{x}$  have already been followed by an  $n \times n$  whitening matrix  $\mathbf{V}$  such that the components of  $\tilde{\mathbf{x}}(t) = \mathbf{V}\mathbf{x}(t)$  are of unit variance and uncorrelated. Furthermore, assume that we want to estimate a source signal; for this purpose we design a single processing unit described as

$$\begin{aligned}\tilde{y}_i(t) &= \mathbf{w}_i^T \tilde{\mathbf{x}}(t) \\ \tilde{y}_i(t - \tau) &= \mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau),\end{aligned}\quad (2)$$

where  $\mathbf{w}_i = (w_{i1}, \dots, w_{in})^T$  is the weight vector which corresponds to the estimate of one row of  $(\mathbf{V}\mathbf{A})^{-1}$ ,  $\tilde{y}_i(t)$  is the output signal which corresponds to the estimate of the source signal  $s_i$ , and  $\tau$  is some lag constant, often equal to 1.

We present the following constrained minimization problem based on the contrast function with the non-Gaussianity, the linear predictability and the nonlinear predictability of the desired source:

$$\begin{aligned}\min_{\|\mathbf{w}_i\|=1} \Psi(\mathbf{w}_i) &= \lambda E\{G(\tilde{y}_i(t) - \alpha \tilde{y}_i(t - \tau))\} + (1 - \lambda)E\{G(f(\tilde{y}_i(t)) - \beta f(\tilde{y}_i(t - \tau)))\} \\ &= \lambda E\{G(\mathbf{w}_i^T \tilde{\mathbf{x}}(t) - \alpha \mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))\} + (1 - \lambda)E\{G(f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t)) - \beta f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau)))\}.\end{aligned}\quad (3)$$

In this optimization problem,  $\tilde{y}_i(t) - \alpha \tilde{y}_i(t - \tau)$  and  $f(\tilde{y}_i(t)) - \beta f(\tilde{y}_i(t - \tau))$  define the linear and nonlinear innovation functions of the desired source.  $f$  defines the nonlinear predictability of the desired source, and examples of choices are  $f(u) = u^2$  or  $f(u) = \frac{1}{\gamma} \log(\cosh(\gamma u))$  (where  $\gamma \geq 1$  is a constant).  $\alpha$  and  $\beta$  define autoregressive coefficients.  $G$  is a given (usually convex) loss function; generally, we can choose  $G(u) = u^2$ ,  $G(u) = |u|$ , or  $G(u) = \frac{1}{\gamma} \log(\cosh(\gamma u))$  (where  $\gamma \geq 1$  is a constant). In this paper, we adopt the loss function  $G(u) = \log(\cosh(u))$ , due to its better analysis properties and robustness against outliers, such as those found in ICA [1,2,11].  $\lambda$  defines the coefficient of trade-off between linear and nonlinear predictability, which measures the degrees of linear and nonlinear predictability of the desired source and includes only linear or quadratic predictability as special cases ( $\lambda = 0$  for nonlinear predictability and  $\lambda = 1$  for linear predictability).

In fact, assuming that  $\lambda = 1$ , the contrast function (3) reduces to

$$\min_{\|\mathbf{w}_i\|=1} \Psi(\mathbf{w}_i) = E\{G(\mathbf{w}_i^T \tilde{\mathbf{x}}(t) - \alpha \mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))\}, \quad (4)$$

which is the complexity pursuit contrast function for BSS presented in the paper [15,16] when the autoregressive model has just one predicting term, which combines non-Gaussianity and linear predictability for BSS. Thus, in fact, we adopt a hybrid technique for BSS based on linear and nonlinear complexity pursuit.

To perform the optimization in (3), we can use a simple gradient descent. The gradients of  $\Psi(\mathbf{w}_i)$  with respect to  $\mathbf{w}_i$ ,  $\alpha$ , and  $\beta$  are obtained as

$$\begin{aligned}\frac{\partial \Psi(\mathbf{w}_i)}{\partial \mathbf{w}_i} &= \lambda E\{(\tilde{\mathbf{x}}(t) - \alpha \tilde{\mathbf{x}}(t - \tau))g(\mathbf{w}_i^T \tilde{\mathbf{x}}(t) - \alpha \mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))\} \\ &\quad + (1 - \lambda)E\{(f'(\mathbf{w}_i^T \tilde{\mathbf{x}}(t))\tilde{\mathbf{x}}(t) - \beta f'(\mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))\tilde{\mathbf{x}}(t - \tau))g(f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t)) - \beta f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau)))\},\end{aligned}\quad (5)$$

$$\frac{\partial \Psi(\mathbf{w}_i)}{\partial \alpha} = -\lambda E\{(\mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))g(\mathbf{w}_i^T \tilde{\mathbf{x}}(t) - \alpha \mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))\}, \quad (6)$$

$$\frac{\partial \Psi(\mathbf{w}_i)}{\partial \beta} = -(1 - \lambda)E\{f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))g(f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t)) - \beta f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau)))\}, \quad (7)$$

where the function  $g$  is the derivative of  $G$  (when  $G(u) = \log(\cosh(u))$ ,  $g(u) = \tanh(u)$ ) and the function  $f'$  is the derivative of  $f$ . Thus, a hybrid technique based on linear and nonlinear complexity pursuit for BSS (hybrid complexity BSS: HCBSS) is obtained as follows:

*Algorithm outline: HCBSS (estimating one source)*

- (1) Center the data to make the mean zero and whiten the data to give  $\tilde{\mathbf{x}}(t)$ . Choose initial values for  $\mathbf{w}_i$ ,  $\alpha$  and  $\beta$ , and suitable learning rates  $\mu_{\mathbf{w}_i}$ ,  $\mu_\alpha$  and  $\mu_\beta$ .

(2) Update the weight vector by using

$$\begin{aligned}\mathbf{w}_i &\leftarrow \mathbf{w}_i - \mu_{\mathbf{w}_i}(\lambda E\{(\tilde{\mathbf{x}}(t) - \alpha \tilde{\mathbf{x}}(t - \tau))g(\mathbf{w}_i^T \tilde{\mathbf{x}}(t) - \alpha \mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))\} \\ &\quad + (1 - \lambda)E\{(f'(\mathbf{w}_i^T \tilde{\mathbf{x}}(t))\tilde{\mathbf{x}}(t) - \beta f'(\mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))\tilde{\mathbf{x}}(t - \tau))g(f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t)) - \beta f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau)))\}) \\ \alpha &\leftarrow \alpha + \mu_\alpha \lambda E\{(\mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))g(\mathbf{w}_i^T \tilde{\mathbf{x}}(t) - \alpha \mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))\}, \\ \beta &\leftarrow \beta + \mu_\beta (1 - \lambda) E\{f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))g(f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t)) - \beta f(\mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau)))\}, \\ \mathbf{w}_i &\leftarrow \mathbf{w}_i / \|\mathbf{w}_i\|.\end{aligned}$$

(3) If convergence is not achieved, go back to step (2).

To estimate the separating matrix  $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_n)^T$ , one can simply use a deflation scheme (one-by-one estimation) or symmetric orthogonalization [1,2].

### 3. Connection to other BSS contrast functions

In this section, we show how the contrast function (3) is connected to other BSS contrast functions.

First, assume that the signal has no time dependences and the function  $f(u)$  is a linear function; then our contrast function (3) reduces to

$$\min_{\|\mathbf{w}_i\|=1} \Psi(\mathbf{w}_i) = E\{G(\mathbf{w}_i^T \tilde{\mathbf{x}}(t))\}. \quad (8)$$

This is in fact the well-known one-unit contrast function of ICA [2]. In particular, if  $G$  is chosen as the negative log density of the component, which captures the non-Gaussianity of the component, we can obtain one of the independent components by performing optimization. Most well-known algorithms for approximating ICA for non-Gaussian sources are closely related to the contrast function (8) [2].

Assume that  $G(u) = u^2$  and the coefficient  $\lambda = 1$ ; the contrast function (3) reduces to

$$\min_{\|\mathbf{w}_i\|=1} \Psi(\mathbf{w}_i) = E\{(\mathbf{w}_i^T \tilde{\mathbf{x}}(t) - \alpha \mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))^2\}. \quad (9)$$

From the contrast function, we can obtain a similar blind source extraction (BSE) algorithm [11] based on the linear predictability. Using the linear predictability, one can perform BSE when sources have linear temporal autocorrelations.

As another special case, assume that  $\lambda = 1$ ; the contrast function (3) reduces to

$$\min_{\|\mathbf{w}_i\|=1} \Psi(\mathbf{w}_i) = E\{G(\mathbf{w}_i^T \tilde{\mathbf{x}}(t) - \alpha \mathbf{w}_i^T \tilde{\mathbf{x}}(t - \tau))\}, \quad (10)$$

which is the complexity pursuit contrast function for BSS presented in the paper [15,16] when the autoregressive model has just one predicting term.

Also, assume that  $\lambda = 0$  and  $\beta = 1$ ; the contrast function (3) reduces to

$$\min_{\|\mathbf{w}_i\|=1} \Psi(\mathbf{w}_i) = E\{G(f(\tilde{y}_i(t)) - f(\tilde{y}_i(t - \tau)))\}, \quad (11)$$

which is the contrast function using the difference (innovation) of sources, i.e., the MINDIFF algorithm for BSS presented in the paper [25].

Also, assume that  $\lambda = 0$  and  $G(u) = f(u) = u^2$ ; the contrast function (3) reduces to

$$\min_{\|\mathbf{w}_i\|=1} \Psi(\mathbf{w}_i) = E\{(\tilde{y}_i^2(t) - \beta \tilde{y}_i^2(t - \tau))^2\}, \quad (12)$$

which is the energy (square) predictability contrast function for BSS presented in the paper [24].

In the following simulations, the cases show that the existing BSS methods are not able to separate the sources; only the HCBSS algorithm presented in the paper performs well.

### 4. Experimental results

In order to verify the effectiveness of the proposed algorithm, we carried out many simulations using artificial data, images, artificial ECG data and real-world ECG data [29]. The accuracy of separation is measured by the following index:

$$\text{SNR}_i = -10 \log E\{(s_i(t) - y_i(t))^2\}, \quad (13)$$

where  $s_i$  is the original signal, and  $y_i$  is the recovered corresponding signal (both are normalized to have zero mean and unit variance). The higher  $\text{SNR}_i$  is (e.g., over 20), the better the performance is.

Besides that, we also calculated the performance index

$$\text{PI} = \frac{1}{n^2} \left\{ \sum_{i=1}^n \left( \sum_{j=1}^n \frac{|p_{ij}|}{\max_k |p_{ik}|} - 1 \right) + \sum_{j=1}^n \left( \sum_{i=1}^n \frac{|p_{ij}|}{\max_k |p_{kj}|} - 1 \right) \right\}, \quad (14)$$

where  $p_{ij}$  is the  $ij$ th element of the  $n \times n$  matrix  $\mathbf{P} = \mathbf{WV}\mathbf{A}$ .

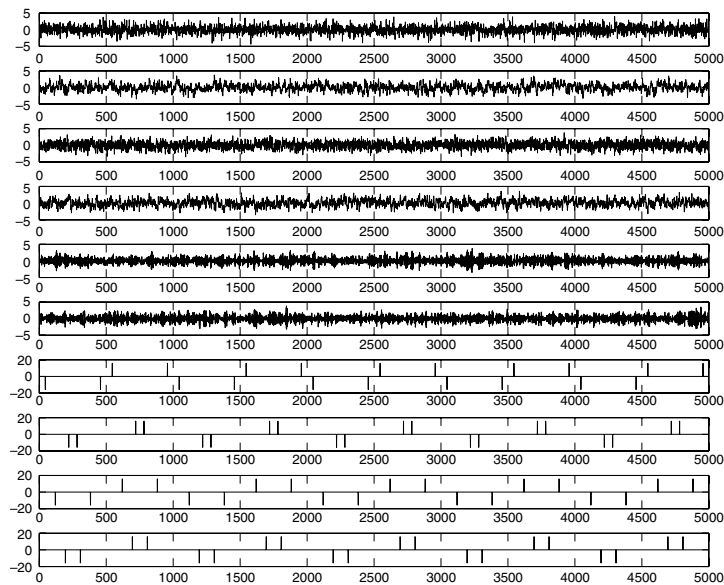


Fig. 1. Ten source signals.

#### 4.1. Experiments on artificial data

To validate the algorithm, we performed blind source separation experiments with artificial data. In each trial, we created ten source signals. For the first six sources (taken from [30]), first, we created four signals using a first-order autoregressive model with constant variances of the innovations, with 5000 time points. Of these four, signals 1 and 2 were created with super-Gaussian innovations, and the signals 3 and 4 with Gaussian innovations. All of these innovations had constant unit variance. The signals 1 and 3 had identical autoregressive coefficients (0.33), and therefore identical autocovariances; the signals 2 and 4 had identical coefficients (0.75) as well. Finally, we created signals 5 and 6 such that they had smoothly changing variances as follows: we created two Gaussian signals with the same autoregressive method as above (except that the coefficient was 0.9), and then completely randomized the signs of the signals by multiplying the signals by two binary i.i.d. signals that took the values  $\pm 1$  with equal probabilities. For the last four sources, we adopted some of the data ACvsparse10 (taken from <http://www.bsp.brain.riken.jp/ICALAB/ICALABSignalProc/benchmarks/>); these were very sparse sources.

The waveforms of ten original signals are shown in Fig. 1. These ten signals were then mixed as in ICA, using random mixing  $10 \times 10$  matrices. The waveforms of the ten mixed signals are shown in Fig. 2. The HCBSS algorithm is used to estimate the separating matrix (where  $G(u) = f(u) = \log(\cosh(u))$ , the learning rates  $\mu_{w_i} = 1$ ,  $\mu_\alpha = 0.1$  and  $\mu_\beta = 0.1$ , and  $\lambda = 0.2$ ). Ordinary ICA methods based on non-Gaussianity would be able to separate only signals 1 and 2. Methods based on second-order correlations would not be able to separate any of the signals, since there was no signal with a unique autocorrelation. Methods based on nonstationary variances would be able to separate only signals 5 and 6. Methods combining non-Gaussianity with autocorrelations would be able to separate only the first four signals. Also, a unifying algorithm for blind separation [30] cannot separate the four sparse sources. We ran our algorithm on 100 data sets generated as described above. Symmetric orthogonalization was used.

Fig. 3 shows the average performance indexes over 100 independent trials against iteration numbers for the HCBSS algorithm with the symmetric orthogonalization. The algorithm correctly estimated the independent components, in around 200 iterations. Note that a single generic nonlinearity that corresponds to super-Gaussian innovations was able to separate both Gaussian and super-Gaussian signals, which indicates that the method is robust with respect to the choice of nonlinearity in much the same way as ICA.

Moreover, this experiment was independently repeated 100 times and the averaged SNRs are listed in Table 1. The following BSS algorithms are included in the comparison: (a) FastICA [2] based on non-Gaussianity; (b) complexity pursuit (CP) [15], which combines non-Gaussianity and linear predictability for BSS; (c) energy (square) predictability (EP) [24] to BSS; (d) nonlinear innovation (MINDIFF) [25] to BSS; (e) the cumulant-based fixed-point approach using the nonstationarity of variance (FPNSV) [21]; (f) the JADE algorithm proposed by Cardoso and Souloumiac [31]; (g) an algorithm for an ICA problem based on an efficient entropy estimator (RADICAL) [32]; (h) the proposed HCBSS algorithm (where  $G(u) = f(u) = \log(\cosh(u))$ , and the learning rates  $\mu_{w_i} = 1$ ,  $\mu_\alpha = 0.1$  and  $\mu_\beta = 0.1$ ).

From Table 1, we can see that the FastICA, JADE and RADICAL algorithms separate the last four signals well, due to their linear autocorrelations, but fail to separate the first six signals; in contrast to the linear predictability methods, EP, MINDIFF and FPNSV separate the last six signals well, but all fail to separate the first four signals; CP can successfully separate each

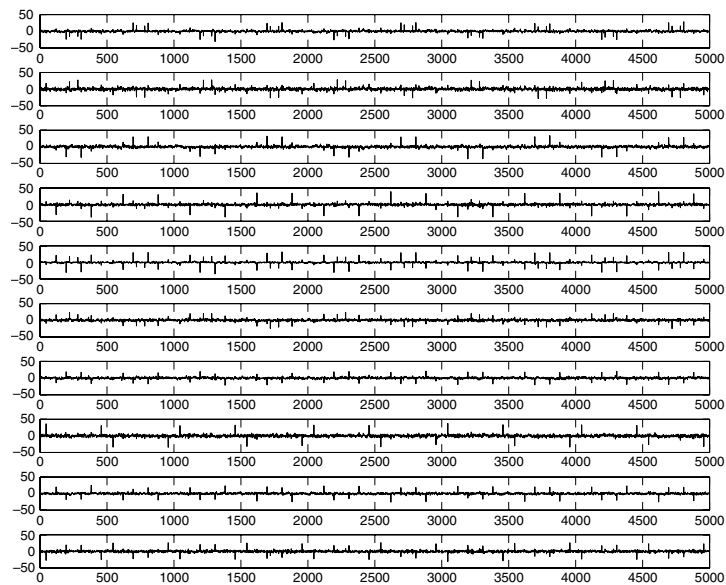


Fig. 2. Ten mixed signals.

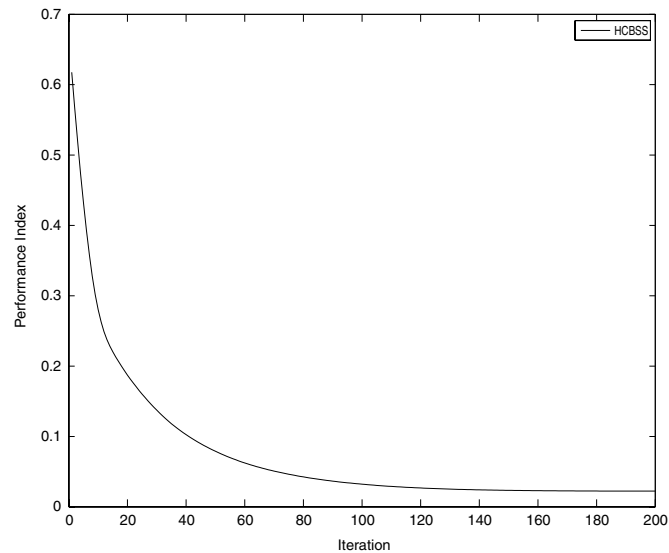


Fig. 3. Average performance indexes over 100 independent runs for ten sources, using the HCBSS algorithm.

**Table 1**  
The average SNR indexes over 100 independent trials for the BSS algorithms.

	HCBSS	FastICA	CP	EP	MINDIFF	FPNSV	JADE	RADICAL
<b>SNR<sub>1</sub></b>	32.49	26.57	35.99	7.97	8.62	19.07	11.05	25.66
<b>SNR<sub>2</sub></b>	39.63	15.39	39.77	9.57	10.72	13.10	8.87	9.40
<b>SNR<sub>3</sub></b>	25.76	9.99	22.79	9.22	10.10	9.45	7.68	8.06
<b>SNR<sub>4</sub></b>	31.59	7.92	30.13	8.95	10.63	10.72	7.27	9.37
<b>SNR<sub>5</sub></b>	25.15	10.60	13.49	34.28	42.94	32.63	7.96	7.60
<b>SNR<sub>6</sub></b>	24.66	9.26	13.47	34.00	36.24	31.96	6.47	9.13
<b>SNR<sub>7</sub></b>	48.10	49.42	32.64	75.27	28.85	34.20	99.77	50.04
<b>SNR<sub>8</sub></b>	48.68	49.70	32.66	76.07	28.56	34.26	99.86	50.04
<b>SNR<sub>9</sub></b>	48.10	49.43	32.66	75.27	28.54	34.22	99.70	50.06
<b>SNR<sub>10</sub></b>	48.08	49.39	32.62	75.28	28.54	34.18	99.73	50.08

source signal, except for signals 5 and 6. Thus, one cannot separate all of the source signals using source separation methods based on non-Gaussianity, linear predictability, coding complexity, nonstationarity of variance, and energy predictability.

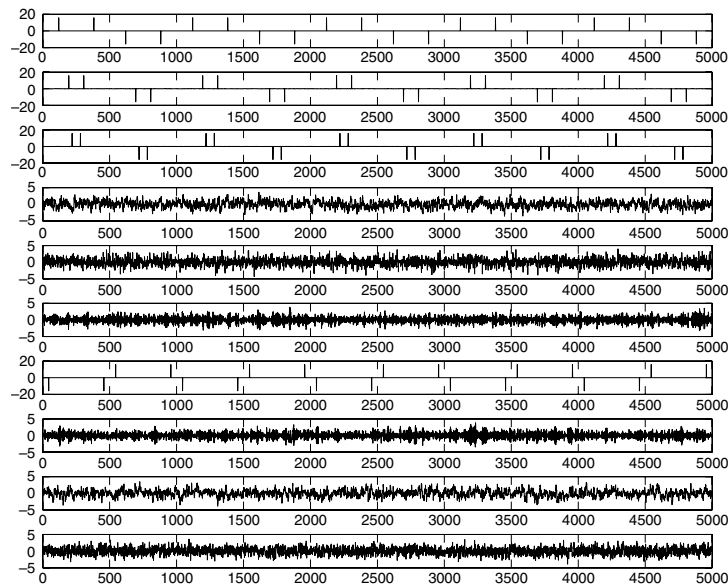
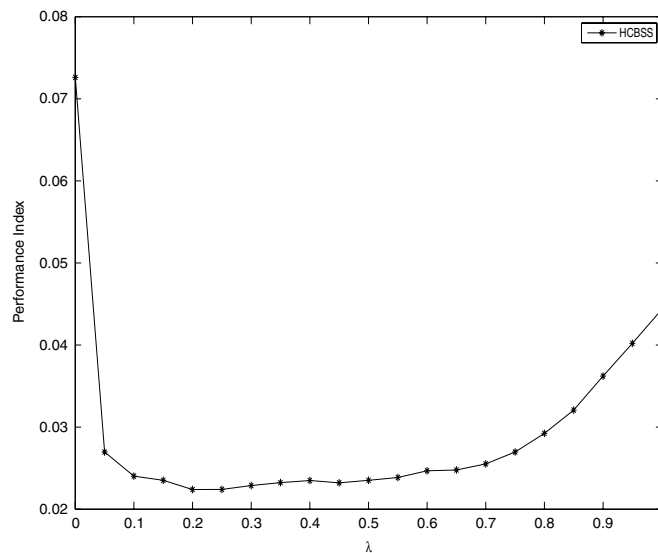


Fig. 4. Ten separated signals.

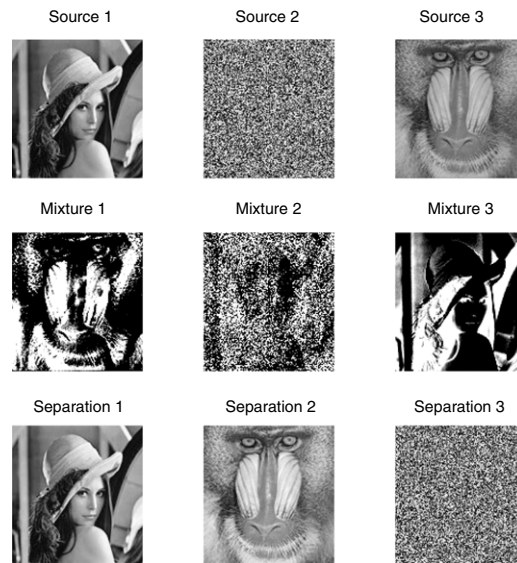
Fig. 5. The comparison of PI with  $\lambda$  for artificial data.

However, the proposed HCBSS algorithm, based on linear and nonlinear complexity pursuit, separates all source signals well, showing that the three statistical properties of the source signals—the non-Gaussianity, the linear predictability and the nonlinear predictability—can be the BSS principles. Fig. 4 shows the ten signals recovered by the HCBSS algorithm, which are very close to the original sources.

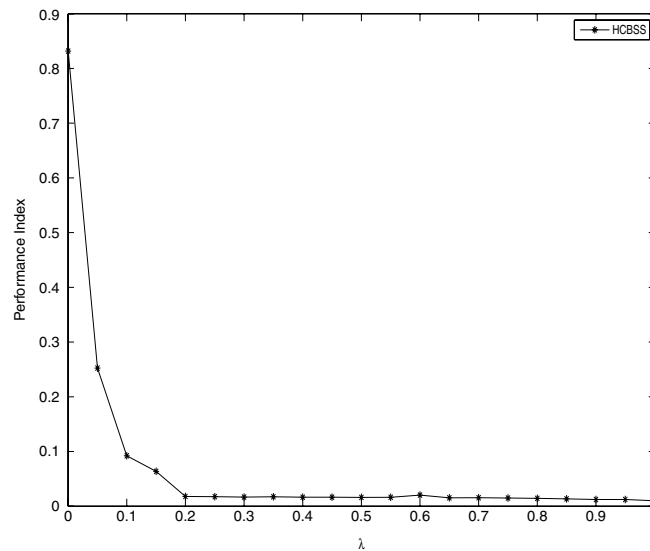
In our algorithm, the parameter  $\lambda$  aims to balance linear and nonlinear predictability. By adjusting this parameter, we obtain an optimal solution. Fig. 5 gives the comparison of PI with  $\lambda$  for artificial data. It is shown that the performance of HCBSS is robust for most of the parameter values between 0 and 1. Therefore, we are able to choose some constant between 0 and 1, which corresponds to the lowest PI value. Therefore in this experiment  $\lambda = 0.2$  is used. Note that the performance is estimated as the mean of the PI values of 100 independent trials for each value of  $\lambda$ . In every trial, eight algorithms are run with 200 iterations, which seems to always be enough for convergence.

#### 4.2. Experiments on image data

Two  $128 \times 128$  images and one i.i.d. Gaussian noise were used in the second simulation. From the top down, Fig. 6 gives the original images sources 1–3 and the mixtures 1–3. The images separated by the HCBSS algorithm are presented



**Fig. 6.** Simulation results for the mixture of two images and one Gaussian signal. From top to bottom: the original images sources 1–3, the mixtures 1–3, and the separations 1–3 obtained by the HCBSS algorithm.



**Fig. 7.** The comparison of PI with  $\lambda$  for image data.

at the bottom of Fig. 6. The performance indexes of the separations are 44.11 (separation 1), 55.52 (separation 2) and 30.98 (separation 3), respectively.

For the comparison of performances, we tested the previous eight BSS algorithms ((a)–(h)) in the image experiment. The average performance indexes of 100 independent trials were computed. The results are shown in Table 2.

From Table 2, we can see that FastICA, FPNV and JADE cannot perfectly separate all images; other algorithms, such as CP, EP and MINDIFF, have the ability to separate all images, but their performances of some separations are not better than those of the HCBSS and RADICAL algorithms. Obviously, for the whole performance, HCBSS is identical to the RADICAL algorithm and more efficient than the other algorithms.

It is worth noting that the same parameter selection procedure is shown in Fig. 7, which gives the comparison of PI with  $\lambda$  for image data. From Fig. 7, we can see that the performance of HCBSS is actually robust when the parameter value is larger than 0.2. In this experiment, we chose  $\lambda$  as 0.9 because this corresponds to the optimal PI value.

#### 4.3. Experiments on ECG data

The extraction of the fetal ECG (FECG) using a non-invasive technique is an important challenge in biomedical signal processing and analysis. The FECG contains important information about the health and condition of the fetus. However,

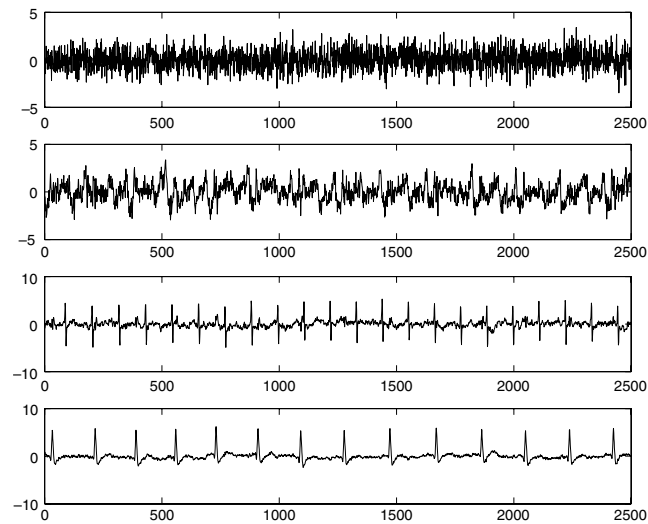


Fig. 8. Four artificial ECG signals.

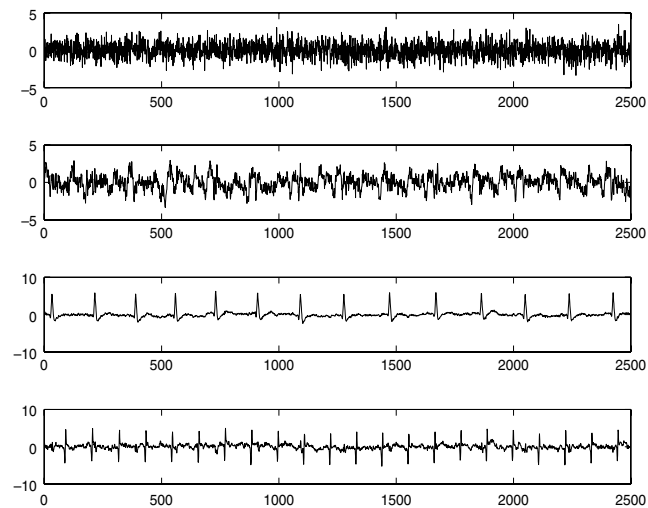


Fig. 9. Separations for artificial ECG signals obtained by the HCBSS algorithm.

there are some problems; for example, the FECG is always corrupted by various kinds of noise, such as the maternal ECG (MECG) with an extremely high amplitude, respiration and stomach activity, thermal noise, etc. Therefore, how to extract the FECG clearly has become a vital issue. In the following, we made many experiments on artificial ECG data and real-world ECG data [29], to which we added some additive Gaussian noise, in order to extract a clearer FECG signal and demonstrate the validity of the proposed algorithm.

#### 4.3.1. Experiments on artificial ECG data

We adopted four zero-mean and unit-variance source signals (2500 samples), as shown in Fig. 8. From the top down, they are an electrode artifact, the FECG, the MECG and one Gaussian signal. The observed signals were generated by a  $4 \times 4$  random mixing matrix. We ran the HCBSS algorithm, and Fig. 9 shows the separations. The accuracies of these, which were measured by means of the SNR index, are 37.67, 45.07, 34.90 and 44.62, respectively. Note that Fig. 10 shows the averaged **PI** values over 100 independent trials. Obviously, the proposed algorithm has the lowest **PI** values and better convergence than the other five algorithms.

In order to further verify the performance of the proposed algorithm, we independently repeated the SNR comparison experiment 100 times and the averaged SNRs are listed in Table 3. It shows that the HCBSS algorithm is the best among the eight algorithms. Certainly, the performances of the CP, FPNSV and JADE algorithms are also better than those of the other four algorithms, that is the FastICA, EP, MINDIFF and RADICAL algorithms. We also know that the FastICA algorithm is almost comparable to RADICAL, which only separates the last two signals successfully.



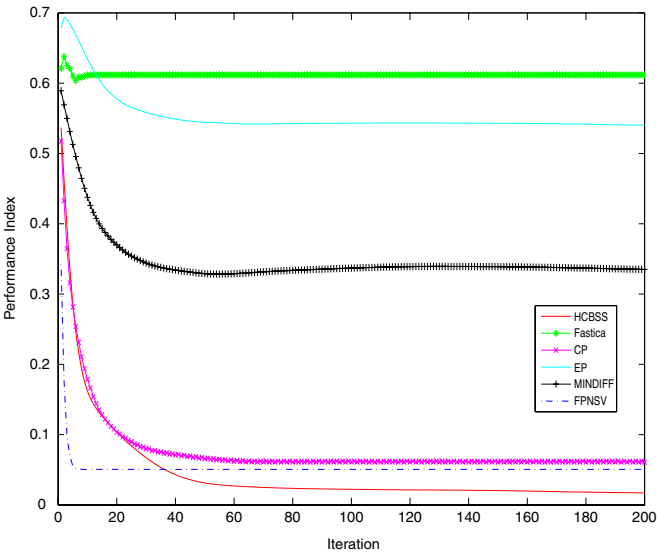


Fig. 10. Average PI values over 100 independent runs for six algorithms.

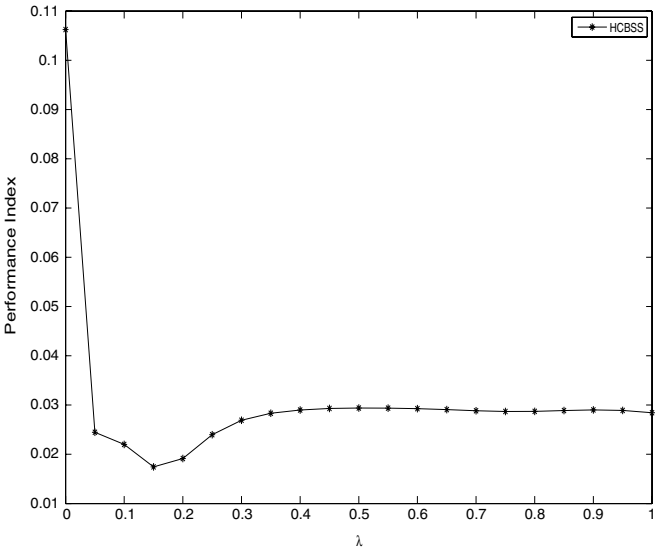


Fig. 11. The comparison of PI with  $\lambda$  for artificial ECG data.

**Table 2**  
The average SNR indexes over 100 independent trials for the BSS algorithms.

	HCBSS	FastICA	CP	EP	MINDIFF	FPNSV	JADE	RADICAL
SNR <sub>1</sub>	50.44	18.55	21.64	31.05	29.28	18.22	19.64	43.33
SNR <sub>2</sub>	55.50	42.48	23.16	56.06	58.68	31.41	50.73	51.90
SNR <sub>3</sub>	32.17	19.20	33.50	40.26	24.51	18.79	20.57	44.90

In this experiment  $\lambda$  equals 0.15. Fig. 11 shows the comparison of PI with  $\lambda$  for artificial ECG data. Note that this experiment is the same as two above-mentioned experiments. From Fig. 11, we can see that the performance of HCBSS is almost stationary even if the parameter value is changed between 0 and 1, and the minimum of PI is obtained at  $\lambda = 0.15$ .

4.3.2. Experiments on real-world ECG data

To check the validity of the proposed algorithm, we have performed experiments on real-world ECG data, distributed by De Moor [29]. This data set consists of a famous ECG measured from a pregnant woman (shown in Fig. 12). One can see the heart beating of both the mother (stronger and slower) and the fetus (weaker and faster). Note that the fetal influence

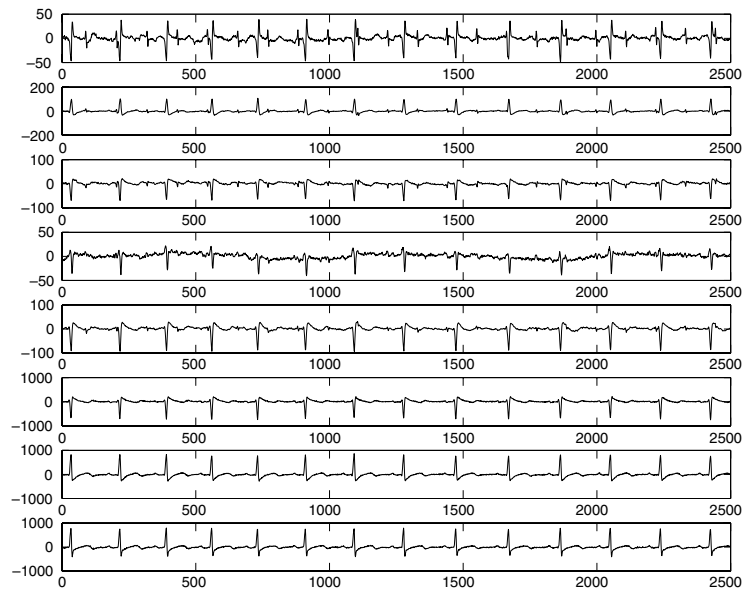


Fig. 12. The eight-channel ECG recording obtained from a pregnant woman.

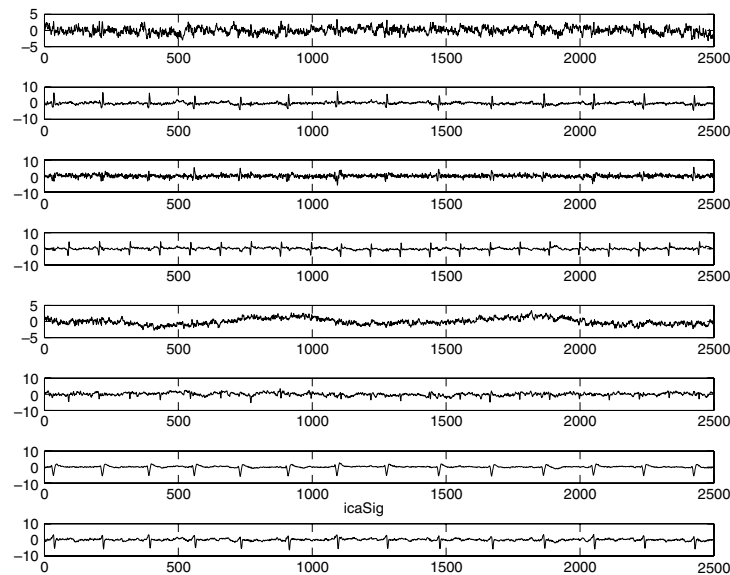


Fig. 13. The separations for real ECG signals obtained by the HCBSS algorithm.

Table 3

The average SNR indexes over 100 independent trials for the BSS algorithms.

	HCBSS	FastICA	CP	EP	MINDIFF	FPNSV	JADE	RADICAL
<b>SNR<sub>1</sub></b>	37.67	10.65	29.42	23.30	30.46	30.13	25.14	9.93
<b>SNR<sub>2</sub></b>	44.84	10.46	30.11	10.90	11.85	33.08	24.23	9.74
<b>SNR<sub>3</sub></b>	34.82	31.04	32.58	12.00	11.61	31.43	31.68	30.27
<b>SNR<sub>4</sub></b>	44.30	32.51	23.95	12.53	15.92	30.80	30.37	30.53

is stronger in the first channel of Fig. 12. The ECG measurements were recorded over 10 s and sampled at 250 Hz. Fig. 13 provides the separations obtained by the HCBSS algorithm. Since the mixing matrix and the source signals are not available, the performance indexes, such as **SNR** and **PI**, cannot be computed as above. But we can perceive distinctly the high quality of the separations through experience.

## 5. Conclusions

We have proposed a hybrid technique based on linear and nonlinear complexity pursuit for BSS (hybrid complexity BSS: HCBSS). We demonstrate the efficient implementation of the method. Also, many optimization contrast functions for BSS can be seen as special cases of the contrast function presented in this paper. The proposed method is also available for on-line learning, which is important for BSS in practice, and we will study this case in the future.

## Acknowledgments

The authors wish to gratefully thank the anonymous reviewers, who provided insightful and helpful comments. Also, the authors wish to gratefully thank Beijing Key Laboratory of Digital Media, Beihang University.

## References

- [1] A. Cichocki, S.-I. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, Wiley, New York, 2002.
- [2] A. Hyvärinen, J. Karhunen, E. Oja, *Independent Component Analysis*, Wiley, New York, 2001.
- [3] S.I. Amari, A. Cichocki, H. Yang, A new learning algorithm for blind source separation, in: *Advances in Neural Information Processing Systems*, vol. 8, 1996, pp. 757–763.
- [4] A. Bell, T. Sejnowski, An information-maximization approach to blind separation and blind deconvolution, *Neural Computation* 7 (6) (1995) 1129–1159.
- [5] J.-F. Cardoso, B.H. Laheld, Equivariant adaptive source separation, *IEEE Transactions on Signal Processing* 44 (12) (1996) 3017–3030.
- [6] P. Comon, Independent component analysis—a new concept? *Signal Processing* 36 (1994) 287–314.
- [7] A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis, *IEEE Transactions on Neural Networks* 10 (3) (1999) 626–634.
- [8] C. Jutten, J. Herault, Blind separation of sources, part I: an adaptive algorithm based on neuromimetic architecture, *Signal Processing* 24 (1991) 1–10.
- [9] T.-W. Lee, M. Girolami, T. Sejnowski, Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and super-Gaussian sources, *Neural Computation* 11 (2) (1999) 417–441.
- [10] H. Zhang, C. Guo, Z. Shi, E. Feng, A new constrained fixed-point algorithm for ordering independent components, *Journal of Computational and Applied Mathematics* 220 (1–2) (2008) 548–558.
- [11] A.K. Barros, A. Cichocki, Extraction of specific signals with temporal structure, *Neural Computation* 13 (9) (2001) 1995–2003.
- [12] A. Belouchrani, K.A. Meraim, J.-F. Cardoso, E. Moulines, A blind source separation technique based on second order statistics, *IEEE Transactions on Signal Processing* 45 (2) (1997) 434–444.
- [13] L. Molgedey, H.G. Schuster, Separation of a mixture of independent signals using time delayed correlations, *Physical Review Letters* 72 (23) (1994) 3634–3637.
- [14] L. Tong, R.-W. Liu, V. Soon, Y.-F. Huang, Indeterminacy and identifiability of blind identification, *IEEE Transactions on Circuits and Systems* 38 (5) (1991) 499–509.
- [15] A. Hyvärinen, Complexity pursuit: separating interesting components from time-series, *Neural Computation* 13 (4) (2001) 883–898.
- [16] Z. Shi, H. Tang, Y. Tang, A fast fixed-point algorithm for complexity pursuit, *Neurocomputing* 64 (2005) 529–536.
- [17] Z. Shi, C. Zhang, Gaussian moments for noisy complexity pursuit, *Neurocomputing* 69 (7–9) (2006) 917–921.
- [18] Z. Shi, C. Zhang, Semi-blind source extraction for fetal electrocardiogram extraction by combining non-Gaussianity and time-correlation, *Neurocomputing* 70 (2007) 1574–1581.
- [19] H. Zhang, Z. Shi, C. Guo, E. Feng, Semi-blind source extraction algorithm for fetal electrocardiogram based on generalized autocorrelations and reference signals, *Journal of Computational and Applied Mathematics* 223 (1) (2009) 409–420.
- [20] J.V. Stone, Blind source separation using temporal predictability, *Neural Computation* 13 (2001) 1559–1574.
- [21] A. Hyvärinen, Blind source separation by nonstationarity of variance: a cumulant-based approach, *IEEE Transactions on Neural Networks* 12 (6) (2001) 1471–1474.
- [22] K. Matsuoka, M. Ohya, M. Kawamoto, A neural net for blind separation of nonstationary signals, *Neural Networks* 8 (3) (1995) 411–419.
- [23] D.-T. Pham, J.-F. Cardoso, Blind separation of instantaneous mixtures of non stationary sources, *IEEE Transactions on Signal Processing* 49 (9) (2001) 1837–1848.
- [24] Z. Shi, C. Zhang, Energy predictability to blind source separation, *Electronics Letters* 42 (17) (2006) 1006–1008.
- [25] Z. Shi, C. Zhang, Nonlinear innovation to blind source separation, *Neurocomputing* 71 (2007) 406–410.
- [26] Z. Shi, Z. Jiang, F. Zhou, A fixed-point algorithm for blind source separation with nonlinear autocorrelation, *Journal of Computational and Applied Mathematics* 223 (2) (2009) 908–915.
- [27] Z. Shi, Z. Jiang, F. Zhou, J. Yin, Blind source separation with nonlinear autocorrelation and non-Gaussianity, *Journal of Computational and Applied Mathematics* 229 (1) (2009) 240–247.
- [28] Z. Shi, C. Zhang, Fast nonlinear autocorrelation algorithm for source separation, *Pattern Recognition* 42 (9) (2009) 1732–1741.
- [29] D. De Moor (Ed.) *Daisy: database for the identification of systems*. Available online at: <http://www.esat.kuleuven.ac.be/sista/daisy>.
- [30] A. Hyvärinen, A unifying model for blind separation of independent sources, *Signal Processing* 85 (7) (2005) 1419–1427.
- [31] J.-F. Cardoso, A. Souloumiac, Blind beamforming for non-Gaussian signals, *IEEE Proceedings F (Communications, Radar and Signal Processing)* 140 (1993) 362–370.
- [32] E.G. Learned-Miller, J.W. Fisher III, ICA using spacings estimates of entropy, *Journal of Machine Learning Research* 4 (2003) 1271–1295.