# A Novel FastICA Method for the Reference-based Contrast Functions

*Wei ZHAO*[1], *Yimin WEI*[2], *Yuehong SHEN*[2], *Zhigang YUAN*[2], *Pengcheng XU*[1], *Wei JIAN*[2]

[1]College of Communications Engineering, PLA University of Science and Technology, Yudao Street, 210007 Nanjing, China
[2]Wireless Communications Lab of College of Communications Engineering, PLA University of Science and Technology, Yudao Street, 210007 Nanjing, China

gmajyie@126.com

**Abstract.** *This paper deals with the efficient optimization problem of Cumulant-based contrast criteria in the Blind Source Separation (BSS) framework, in which sources are retrieved by maximizing the Kurtosis contrast function. Combined with the recently proposed reference-based contrast schemes, a new fast fixed-point (FastICA) algorithm is proposed for the case of linear and instantaneous mixture. Due to its quadratic dependence on the number of searched parameters, the main advantage of this new method consists in the significant decrement of computational speed, which is particularly striking with large number of samples. The method is essentially similar to the classical algorithm based on the Kurtosis contrast function, but differs in the fact that the reference-based idea is utilized. The validity of this new method was demonstrated by simulations.*

## Keywords

Blind source separation, FastICA, reference-based contrast functions, Kurtosis contrast function.

## 1. Introduction

For the latest decades, Blind Source Separation (BSS) has been applied in a wide variety of fields such as array processing, passive sonar, seismic exploration, speech processing, multi-user wireless communications, etc [1]. In the case of a linear multi-input/multi-output (MIMO) instantaneous system, BSS corresponds to Independent Component Analysis (ICA), which is now a well recognized concept. In this contribution, we mainly consider the efficient optimization issue of BSS in the FastICA framework, where statistically independent sources are linearly and instantaneously mixed.

In the linear MIMO systems, BSS has found interesting solutions through the optimization of so-called contrast functions [2], which are generally treated as separation criteria. Many separation criteria rely on higher-order statistics (e.g., the Kurtosis contrast function [3], [4]) or can be linked to higher-order statistics (e.g., the Constant Modulus contrast function [5]). These criteria are known to provide good results. Recently, some novel contrast schemes referred to as "reference-based" have been proposed in [6] and [7]. They are essentially the cross-statistics or cross-cumulants between the estimated outputs and reference signals [8]-[9]. These reference-based contrast functions have an appealing feature in common: the corresponding optimization algorithms are quadratic with respect to the searched parameters. First of all, we give a brief review of previous works on this subject.

- A maximization algorithm based on Singular Value Decomposition (SVD) has been proposed in [6] and [10], and it was shown to be significantly quicker than other maximization algorithms. However, the method often suffers from the need to have a good knowledge of the filter orders due to its sensitivity on a rank estimation [11].
- A gradient optimization method based on Kurtosis with reference signals introduced has been proposed in [12], which obtains an optimal step size and dose not require any rank estimation. Therefore, the drawback of the SVD-based method can be well overcome. But the reference signals involved in this method are fixed during the optimization process, which may lead to bad separation performance because of inappropriate initialization value of the corresponding reference signals.
- A similar gradient optimization algorithm based on Kurtosis maximization has been proposed in [13], in which the reference signals used in the separation process update after each one-dimensional optimization. So the separation quality of this method is better than that in [12].
- On the basis of the algorithms in [12] and [13], a new improvement method has been proposed in [11]. For this method, a tradeoff can be adjusted between performance and speed by introducing a new iterative update parameter. Moreover, the global convergence of this algorithm is proved in detail.

Inspired by [11] and [13], in this paper, we propose a new algorithm by considering the reference-based Kurtosis maximization in the FastICA optimization framework. The papers most directly linked to our approach are [11] on the one hand and [13] on the other hand. The former has provided a proof of stationary consistency of the reference-based contrast function, which is directly cited in our paper. The latter has proposed a gradient optimization algorithm with the reference signals updated, which contributes to the proposal of our method primarily. Besides the gradient optimization algorithms in [11]-[13], our proposed method provides another approach to an efficient optimization of reference-based contrast functions. To our knowledge, it has not been investigated yet despite its simplicity.

This paper is organized as follows. Section 2 describes the model and assumptions we consider in this paper. In Section 3, the separation criteria we use are presented. Our proposed algorithm can be found in Section 4. Simulation results are illustrated in Section 5 and Section 6 concludes this paper.

# 2. System Model and Assumptions

## 2.1 Instantaneous Mixture

We consider an observed M-dimensional ($M \geq 2$) discrete-time signal, the $n$th sample of which is denoted by the column vector $\mathbf{x}(n)$ (where $n \geq 1$ holds implicitly in this whole paper). The observed signals result from a noise-free linear MIMO system, for which the input and output relationship is described as follows:

$$\mathbf{x}(\mathrm{n}) = \mathbf{A}\mathbf{s}(\mathrm{n}) \tag{1}$$

where A is a linear mixture matrix of $M \times N$, the elements of which are unknown constant. The $N$-dimensional ($N \geq 2$) source vector $\mathbf{s}(\mathrm{n})$ is unknown and unobserved.

The objective of BSS is to recover source signals blindly only by using the observations. Similarly, we consider a linear separator, the output of which is described as:

$$\mathbf{y}(\mathrm{n}) = \mathbf{W}\mathbf{x}(\mathrm{n}) \tag{2}$$

where $\mathbf{W}$ is the separation matrix of $N \times M$. $\mathbf{y}(\mathrm{n})$ is the approximate estimation of $\mathbf{s}(\mathrm{n})$. We mention above that our proposed method presents a quadratic dependence of the searched parameters, where the "searched parameters" are the row vectors of $\mathbf{W}$.

## 2.2 Assumptions on the Sources

To be able to carry out the estimation blindly and successfully, we make some assumptions on the sources [7], which are shown as follows:

A1: For all $i$, the source sequence $s_i(n)$ is stationary, zero-mean and with unit variance.

A2: The source vectors $s_i(n), i \in \{1, \cdots, N\}$ are statistically mutually independent.

# 3. Separation Criteria

## 3.1 Notations

In order to describe the reference-based Kurtosis contrast function we utilize in this paper, we first introduce some notations. The Cumulant of a set of random variables is denoted by $Cum\{\bullet\}$. Note that we only consider the Cumulant of real-valued signals in this paper, even though the signals can be complex- or real-valued. The complex-valued signals case will be considered in our work later.

For any jointly stationary signals $y(n)$ and $z(n)$, we set

$$C\{y\} \stackrel{\Delta}{=} Cum\{y(n), y(n), y(n), y(n)\} = E\{y(n)^4\} - 3E\{y(n)^2\}^2 \tag{3}$$

$$\begin{aligned} C_z\{y\} &\stackrel{\Delta}{=} Cum(y(n), y(n), z(n), z(n)) \\ &= E\{y(n)^2 z(n)^2\} - E\{y(n)^2\}E\{z(n)^2\} - 2E^2\{y(n)z(n)\} \end{aligned} \tag{4}$$

where $E\{\bullet\}$ denotes the expectation value.

## 3.2 Reference Signals

Before introducing the reference-based contrast functions, we first introduce the corresponding "reference signals" we have mentioned above. Similarly to (2), we consider a separation matrix of $N \times M$ denoted by $\mathbf{V}$. The corresponding output can be denoted by

$$\mathbf{z}(\mathrm{n}) = \mathbf{V}\mathbf{x}(\mathrm{n}) \tag{5}$$

where the components of $\mathbf{z}(\mathrm{n})$ are the reference signals. Note that the reference signals have direct influence on the reference-based contrast function and their values do impact the optimization results, especially the initialization value.

As described in [11], the reference signals are artificially introduced in the algorithm for the purpose of facilitating the maximization of the contrast function. In [12], the reference signals are initialized arbitrarily and kept the same during whole optimization process. In [13], the reference signals are indirectly involved in the iterative optimization process. In other words, the reference signals update following the objective signals. More precisely, $\mathbf{V}$ updates following $\mathbf{W}$ in each loop iteration step. Then the separation quality of the algorithm in [13] is better than that in [12]. In [14] and [15], we have done some corresponding work to investigate the impact of reference signals, which is similar to [12]. Inspired by [13], we consider the case that the reference signals update circularly in this paper. Therefore, the performance of our algorithm in this paper is much better than those in [14] and [15].

## 3.3 Contrast Functions

Let us introduce the following criteria:

$$J(w) = \left| \frac{C\{y(n)\}}{E\left\{(y(n))^2\right\}^2} \right|^2, \tag{6}$$

$$I(w,v) = \left| \frac{C_z\{y(n)\}}{E\left\{(y(n))^2\right\} E\left\{(z(n))^2\right\}} \right|^2 \tag{7}$$

where $J$ is the well-known Kurtosis contrast function, which has been proved to be a contrast function in [3] and [4]. $I$ is the reference-based contrast function used in this paper, which has been proposed in [6] and the consistency of which to a stationary point has been proved in [11]. We mainly focus our attention on the efficient optimization of $I$ in the FastICA framework, which leads to the proposal of our new algorithm.

Furthermore, besides the gradient algorithms in [11]-[13], our method provides another new approach to the efficient optimization of reference-based contrast functions. Recently, we have also done some work by introducing the reference signals to the Negentropy contrast criterion, based on which a family of more efficient and robust algorithms have been proposed such as the algorithm in [16].

# 4. Optimization Method

## 4.1 New Algorithm

We now introduce our new FastICA algorithm based on the reference-based Kurtosis contrast function. First, we give some definitions used in our method. As described in [11], $\nabla$ denotes a gradient operator. $\nabla_1$ and $\nabla_2$ denote partial gradient operators with respect to the first and second parameters, respectively. More precisely, $\nabla J(w)$ is the vector composed of all partial derivatives of $J(w)$, whereas $\nabla_1 I(w,v)$ and $\nabla_2 I(w,v)$ are the vectors of partial derivatives of $I(w,v)$ with respect to $w$ and $v$. We denote $\mathbf{W}$ and $\mathbf{V}$ by $\left(w^1,\ldots,w^N\right)^T$ and $\left(v^1,\ldots,v^N\right)^T$, where $T$ means the transpose.

Combined with (4) and (7), we can get

$$\begin{aligned} I(w,v) &= \left| \frac{C_z\{y(n)\}}{E\left\{(y(n))^2\right\}E\left\{(z(n))^2\right\}} \right|^2 \\ &= \left| \frac{E[(w\mathbf{x}(\mathrm{n}))^2(v\mathbf{x}(\mathrm{n}))^2]-E[(w\mathbf{x}(\mathrm{n}))^2]E[(v\mathbf{x}(\mathrm{n}))^2]-2\{E[(w\mathbf{x}(\mathrm{n}))(v\mathbf{x}(\mathrm{n}))]\}^2}{E\left\{(w\mathbf{x}(\mathrm{n}))^2\right\}E\left\{(v\mathbf{x}(\mathrm{n}))^2\right\}} \right|^2. \end{aligned} \tag{8}$$

Because $\mathbf{x}(\mathrm{n})$ is prewhitened and $\mathbf{W}$ and $\mathbf{V}$ are normalized in our method, we can get $E\left\{(w\mathbf{x}(n))^2\right\} = E\left\{(v\mathbf{x}(n))^2\right\} = 1$ and $ww^T = vv^T = 1$. Then (8) can be reduced to

$$I(w,v) = \left| E[(w\mathbf{x}(\mathrm{n}))^2(v\mathbf{x}(\mathrm{n}))^2] - 3w^2 \right|^2. \tag{9}$$

Hence, the partial derivative of $I(w,v)$ with respect to $w$ can be expressed as

$$\begin{aligned} \nabla_1 I(w,v) = \tfrac{\partial I(w,v)}{\partial w} = 4\left| E[(w\mathbf{x}(\mathrm{n}))^2(v\mathbf{x}(\mathrm{n}))^2] - 3w^2 \right| \\ (E[\mathbf{x}(\mathrm{n})\,(w\mathbf{x}(\mathrm{n}))(v\mathbf{x}(\mathrm{n}))^2] - 3w). \end{aligned} \tag{10}$$

Our new proposed algorithm is shown as follows:

- Eliminate the mean value of $\mathbf{x}$ and prewhiten it.
- Initialize $\mathbf{W}$ and normalize it.
- $(M0)$ For $i = 1,2,\cdots,N$ repeat $(M0)$

  Set $\mathbf{w}_0^i = \mathbf{w}_i$, $\mathbf{v}_0^i = \mathbf{w}_0^i$

  $(M0')$ For $k = 0,1,\cdots,k_{\max}-1$ repeat $(M0')$

  - Set $\mathbf{d}_k = \nabla_1 I(\mathbf{w}_k^i, \mathbf{v}_k^i)$
  - $\mathbf{w}_{k+1}^i = \mathbf{d}_k$
  - Normalize $\mathbf{w}_{k+1}^i$
  - $\mathbf{w}_{k+1}^i = \mathbf{w}_{k+1}^i - \sum_{j=1}^{i-1} \mathbf{w}_j \mathbf{w}_j^T \mathbf{w}_i$
  - Renormalize $\mathbf{w}_{k+1}^i$
  - Set $\mathbf{v}_{k+1}^i = \mathbf{w}_{k+1}^i$
  - Renormalize $\mathbf{w}_{k+1}^i$
  - $\mathbf{w}_i = \mathbf{w}_{k_{\max}-1}^i$
- $\mathbf{y} = \mathbf{W}\mathbf{x}$

Here, $\mathbf{W}_0$ is the initializations of $\mathbf{W}$, which is chosen randomly. From the whole process, we can see the source signals are recovered one by one through each one-dimensional optimization in a deflationary manner. To prevent different one-dimensional optimization converging to the same maxima, a Gram-Schmidt-like decorrelation scheme is adopted as shown above.

## 4.2 Convergent Results

Because of the symmetry of $I(w,v)$, we can get the following relationship.

$$\begin{aligned} &I(w,v) = I(v,w), \\ &\nabla_1 I(w,v) = \nabla_2 I(v,w), \\ &\nabla J(w) = 2\nabla_1 I(w,w) = 2\nabla_2 I(w,w). \end{aligned} \tag{11}$$

From (11), we can see that, during one-dimensional optimization process, the reference-based contrast function is $I(w,v)$ with $v$ fixed instead of $J(w)$. So the optimization algorithm is quadratic dependence on the searched parameters. To justify the convergence of the algorithm, besides the assumptions on source signals mentioned above, another assumption [11] is needed.

A3: The sources sequences $s_i(n), i \in \{1,\cdots,N\}$ are temporally independent and identically distributed (i.i.d.). Moreover, they have fourth-order Cumulants which are all of the same sign.

Now, we can give the following proposition [11] shown below:

*Proposition 1:* Assume that the sequences $w_k^i, i = 1,\ldots,N$ are obtained according to the above algorithm with $k_{\max}$ infinite and all $w_k^i, i = 1,\ldots,N$ are contained in a compat set. Then, under assumptions A1-A3, any convergent subsequence of $w_k^i, i = 1,\ldots,N$ converges to points $(w^i)^*, i = 1,\ldots,N$ respectively such that $\nabla J((w^i)^*) = 0, i = 1,\ldots,N$.

| Separation Method | | Average MSE | | | | Median MSE | | | | Average Execution Time (s) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Number of Samples | | | | Number of Samples | | | | Number of Samples | | | |
| | | 1000 | 5000 | 10000 | 30000 | 1000 | 5000 | 10000 | 30000 | 1000 | 5000 | 10000 | 30000 |
| | 1st | 0.0338 | 0.0123 | 0.0088 | 0.0052 | 0.0314 | 0.0125 | 0.0089 | 0.0051 | | | | |
| G-1 | 2nd | 0.0365 | 0.0124 | 0.0088 | 0.0051 | 0.0346 | 0.0125 | 0.0089 | 0.0051 | 3.793 | 10.743 | 18.934 | 53.401 |
| | 3rd | 0.0367 | 0.0124 | 0.0088 | 0.0051 | 0.0356 | 0.0125 | 0.0089 | 0.0051 | | | | |
| | 1st | 0.0364 | 0.0124 | 0.0088 | 0.0051 | 0.0338 | 0.0125 | 0.0089 | 0.0051 | | | | |
| G-2 | 2nd | 0.0348 | 0.0124 | 0.0088 | 0.0051 | 0.0318 | 0.0125 | 0.0089 | 0.0051 | 3.408 | 7.927 | 13.921 | 42.073 |
| | 3rd | 0.0354 | 0.0123 | 0.0088 | 0.0051 | 0.0327 | 0.0124 | 0.0089 | 0.0051 | | | | |
| | 1st | 0.0375 | 0.0123 | 0.0088 | 0.0051 | 0.0383 | 0.0124 | 0.0089 | 0.0051 | | | | |
| F-1 | 2nd | 0.0352 | 0.0123 | 0.0087 | 0.0051 | 0.0355 | 0.0124 | 0.0089 | 0.0051 | 0.687 | 3.095 | 6.069 | 16.389 |
| | 3rd | 0.0361 | 0.0122 | 0.0086 | 0.0051 | 0.0370 | 0.0124 | 0.0084 | 0.0051 | | | | |
| | 1st | 0.0370 | 0.0124 | 0.0088 | 0.0052 | 0.0361 | 0.0125 | 0.0089 | 0.0052 | | | | |
| F-2 | 2nd | 0.0357 | 0.0124 | 0.0088 | 0.0051 | 0.0344 | 0.0125 | 0.0089 | 0.0051 | 0.150 | 0.366 | 0.641 | 2.564 |
| | 3rd | 0.0352 | 0.0124 | 0.0088 | 0.0051 | 0.0338 | 0.0125 | 0.0089 | 0.0051 | | | | |

**Tab. 1.** MSE and execution time for different separation methods (100 Monte-Carlo runs).

Based on the assumptions A1-A3, the consistent convergence of *Proposition 1* can be proved by referring to *Proposition 1* in [11] and Zangwill's convergence theorem in [17].

# 5. Simulation Results

## 5.1 Experimental Data

In this section, we choose three speech signals as sources. Without loss of generality, we assume that the number of observations and sources are equal, i.e., $N = M = 3$. The iterative parameter $k_{\max} = 1000$. The initialization value of mixture matrix $\mathbf{W}_0$ is randomly chosen. And we implement each algorithm 1000 times independently and obtain the average value. Corresponding notations are described as follows:

G-1 denotes the classical gradient optimization algorithm based on the Kurtosis contrast function. However, the optimization step size is maximized to be optimal as presented in [11]-[13]. G-2 denotes the gradient optimization algorithm based on the reference-based Kurtosis contrast function proposed in [13], but differs in the fact that we apply it in the case of linear and instantaneous mixture in this paper. F-1 denotes the classical FastICA algorithm based on the Kurtosis. F-2 denotes our new FastICA algorithm in this paper.

Average MSE denotes the mean square estimation errors between sources and observations averaged over 1000 Monte-Carlo runs. Median MSE denotes the median mean square estimation errors between sources and observations averaged over 1000 Monte-Carlo runs. Average execution time denotes the execution time of recovering all sources averaged over 1000 Monte-Carlo runs.

1st, 2nd and 3rd denote the first, second and third recovered source signals, respectively. The experimental data in detail is shown in Tab. 1 on the top of this page.

## 5.2 Comments and Analysis

To compare the performance of above four algorithms clearly, the average execution time for samples from 1000 to 30000 is illustrated in Fig. 1.

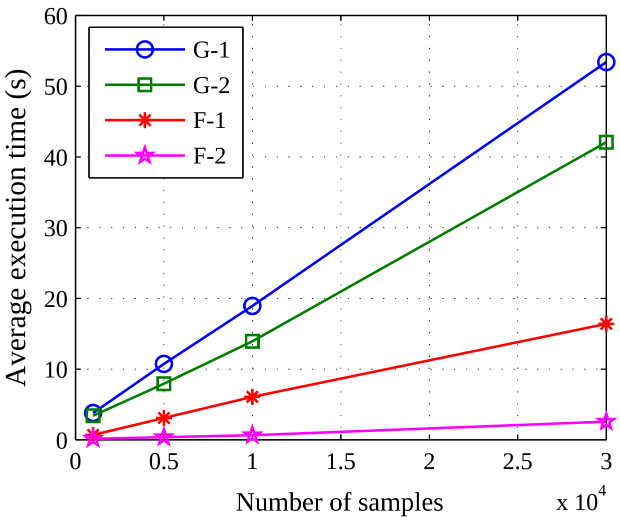


**Fig. 1.** Average execution time for samples from 1000 to 30000 (1000 Monte-Carlo runs).

Firstly, in the gradient optimization framework where G-1 and G-2 are considered, we can see that G-2 shows better performance than G-1 from Tab. 1 and Fig. 1. This has been confirmed and validated in [11] and [13], so detailed comparison and analysis are omitted here.

Secondly, in the FastICA optimization framework where F-1 and F-2 are considered, it can be clearly seen that

F-2 yields similar MSE value to F-1 from the corresponding rows in Tab. 1. This means our new algorithm F-2 performs well and can converge to the same stationary point. Moreover, the average MSE and median MSE value are very close to each other, which means that our method F-2 is very stable. However, we can obviously see that the computational time of F-2 is much less than that of F-1 with varying number of samples, which means our method is much quicker in terms of convergence speed. Furthermore, with an increasing number of samples, the advantage of our new method over F-1 is significantly apparent.

Finally, when the gradient and FastICA optimization schemes are considered together, the similar MSE value of G-1, G-2, F-1 and F-2 in Tab. 1 confirms the convergence performance for all of them. However, it can be observed clearly from Fig. 1 that our algorithm F-2 needs the least execution time among them. In other words, our proposed method F-2 is much quicker than F-1, and also much quicker than G-1 and G-2. It means that our method provides better performance than the corresponding classical algorithm on the one hand and than some other gradient optimization ones under same circumstances on the other hand.

## 6. Conclusion

A novel FastICA optimization algorithm based on the recently proposed reference-based Kurtosis contrast function is proposed in this paper. It is much more efficient than corresponding classical one in terms of computational speed. The performance of this new method is validated through simulations. Our future work includes the extension of our method to complex-valued signals and the application of our method to more complicated channel such as convolution mixture.

## Acknowledgements

This work is supported by the NSF of Jiangsu Province of China under Grant No. BK2011117 and BK2012057, and by the National Natural Science Foundation of China under Grant No. 61172061 and 61201242.

## References

[1] LAHAT, D., CARDOSO, J. F., MESSER, H. Second-order multidimensional ICA: Performance analysis. *IEEE Transactions on Signal Processing*, 2012, vol. 60, no. 9, p. 4598 - 4610.

[2] COMON, P., JUTTEN, C. (Eds.) *Handbook of Blind Source Separation, Independent Component Analysis and Applications.* New York, 2010.

[3] SIMON, C., LOUBATON, P., JUTTEN, C. Separation of a class of convolutive mixtures: A contrast function approach. *Signal Processing*, 2001, vol. 4, no. 81, p. 883 - 887.

[4] TUGNAIT, J. K. Identification and deconvolution of multichannel linear non-Gaussian processes using higher order statistics and inverse filter criteria. *IEEE Transactions on Signal Processing*, 1997, vol. 45, no. 3, p. 658 - 672.

[5] GODARD, D. N. Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *IEEE Transactions on Communications*, 1980, vol. 28, no. 11, p. 1867 - 1875.

[6] CASTELLA, M., RHIOUI, S., MOREAU, E., PESQUET, J. C. Quadratic higher-order criteria for iterative blind separation of a MIMO convolutive mixture of sources. *IEEE Transactions on Signal Processing*, 2007, vol. 55, no. 1, p. 218 - 232.

[7] DUBROCA, R., DELUIGI, C., CASTELLA, M., MOREAU, E. A general algebraic algorithm for blind extraction of one source in a MIMO convolutive mixture. *IEEE Transactions on Signal Processing*, 2010, vol. 58, no. 5, p. 2484 - 2493.

[8] ADIB, A., MOREAU, E., ABOUTAJDINE, D. Source separation contrasts using a reference signal. *IEEE Signal Processing Letters*, 2004, vol. 11, no. 3, p. 312 - 315.

[9] CASTELLA, M., RHIOUI, S., MOREAU, E., PESQUET, J. C. Source separation by quadratic contrast functions: A blind approach based on any higher-order statistics. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing.* Philadelphia (USA), 2005, vol. 3, p. III/569-III/572.

[10] KAWAMOTO, M., KOHNO, K., INOUYE, Y. Eigenvector algorithms incorporated with reference systems for solving blind deconvolution of MIMO-IIR linear systems. *IEEE Signal Processing Letters*, 2007, vol. 14, no. 12, p. 996 - 999.

[11] CASTELLA, M., MOREAU, E. New kurtosis optimization schemes for MISO equalization. *IEEE Transactions on Signal Processing*, 2012, vol. 60, no. 3, p. 1319 - 1330.

[12] CASTELLA, M., MOREAU, E. A new optimization method for reference-based quadratic contrast functions in a deflation scenario. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing.* Taipei (Taiwan), 2009, p. 3161 - 3164.

[13] CASTELLA, M., MOREAU, E. A new method for kurtosis maximization and source separation. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing.* Dallas (TX, USA), 2010, p. 2670 - 2673.

[14] ZHAO, W., YUE, H. S., WANG, J. G., YUAN, Z. G., JIAN, W. New methods for the efficient optimization of cumulant-based contrast functions. In *Proceedings of the $5^{th}$ IET International Conference on Wireless, Mobile and Multimedia Networks.* Beijing (China), 2013, p. 345 - 349.

[15] ZHAO, W., YUE, H. S., WANG, J. G., YUAN, Z. G., JIAN, W. New kurtosis optimization algorithms for independent component analysis. In *Proceedings of the $4^{th}$ IEEE International Conference on Information Science and Technology.* Shenzhen (China), 2014, p. 23 - 28.

[16] ZHAO, W., YUE, H. S., XU, P. C., WANG, J. G., YUAN, Z. G., WEI, Y. M., JIAN, W., LI, H. A new efficient reference-based negentropic algorithm for instantaneous ICA. In *Proceedings of the $3^{th}$ International Conference on Convergence and its Application.* Seoul (Korea), 2014, p. 3 - 4.

[17] ZANGWILL, W. I. *Nonlinear Programming: A Unified Approach.* Englewood Cliffs (NJ, USA): Prentice-Hall, 1969.