Blind Identification of Underdetermined Mixtures Based on Charrelation Matrix

Zhongqiang Luo, Lidong Zhu
National Key Lab. of Science and Technology on Communications
University of Electronic Science and Technology of China
CHINA
No.2006, Xiyuan Ave, West Hi-Tech Zone, 611731,Chengdu, Sichuan, P.R.China
Email: luozhongqiang@126.com

Abstract: In this paper, we propose a novel algorithm for underdetermined blind identification problems in blind signal separation. The proposed algorithm is based on the charrelation matrix of observations. The charrelation matrix can not only be considered as a generalized covariance matrix, but also incorporates higher-order information. It is significant for blind separation problem based on statistic characteristics to extract statistical information. The problem of underdetermined blind identification is converted as a tensor decomposition model. The mixing matrix is estimated from the rank-1 terms of the tensor decomposition. Theoretical analysis and simulation results illustrate that the proposed algorithm performs better estimated performance than the underdetermined blind identification algorithm based on second-order covariance and four-order cumulant respectively.

Key-words: blind identification; blind source separation; tensor decomposition

1 Introduction

Recently, blind signal separation (BSS) technique has become increasingly active and attractive applied in wireless communication systems [1-9]. In our opinion, some reasons can account for it. With the help of BSS technique, on the one hand, the frequently used pilot sequence can be spared for improving spectral efficiency. On the other hand, it can enhance the capacity of source recovery and resist unpredictable interference in communications in spite of fewer prior information [4-7]. Aided by BSS technique, the blind communication technique can be realized. This is promising for some communication environments for anti-interference, whose case are difficult to acquire prior information in communication link.

The aim of the BSS is to recover the unknown source signal or the mixing matrix from only observed signals. Nowadays, underdetermined BSS is a hot topic and difficult problem in the this subject, which has brought about much attention [10-16]. Underdetermined case is ubiquitous in the communication systems. It is a representative case when the number of received sensors is less than that of sources, namely underdetermined multiple input multiple output (MIMO) systems. In the overdetermined or determined case, the unknown source signals can be obtained when the mixed signal is multiplied by the inverse or pseudo-inverse of mixing matrix. Therefore, the estimation of mixing matrix is enough to implement the separation work. Although the underdetermined case does not like the previous operation, the estimation of mixing matrix is a improtant step for further source recovery. The problem of blind identification of underdetermined mixtures will be focused on in this paper.

As far as we are concerned, the algorithm of underdetermined blind identification can be divided into two types. The first type of algorithms are from the assumption that the sources are sparse or the sources may become sparse by executed some preprocessing linear transform (e.g. Short-Time Fourier transform) [10-13]. These algorithms need clustering techniques to implement an exhaustive search for the mixing vector space, and are therefore very expensive, especially when there are more than two observation sensors. The second type of algorithms come from the algebraic structure of statistical characteristic [14-16]. In such case, the sparsity of the source signal can be avoided, these algorithms may be more in common use. There are two typical representative algorithmic, second-order blind identification of underdetermined mixtures (SOBIUM) [14] and fourth-order blind identification of underdetermined mixtures (FOBIUM) [16]. SOBIUM algorithm is based on the algebraic structure of second-order covariance, and FOBIUM algorithm is based on the algebraic structure of fourth-order cumulant. From the complexity perspective, the FOBIUM exceed the SOBIUM.

From the statistical property perspective, the performance of FOBUIM is sensitive to the length of sample due to the estimation problem of higher-order cumulant. However, the SOBIUM can not resist Gaussian noise due to its second-order property compared to the fourth-order cumulant.

In order to coordinate the previous shortcomings, we use a new generic statistics charrelation matrix for blind identification problem. The charrelation matrix is a generalization of covariance matrix, encompassing statistical information beyond second order while maintaining a convenient 2-dimensional structure [17]. As far as we know, there is little literature to report charrelation matrix based blind identification of underdetermined mixtures. Therefore, the main work of this paper is that firstly, we extend this new statistical tool to the more attractive case of complex random signal; then the charrelation matrix based blind identification of underdetermined mixture (CMBIUM) is proposed. Theoretical analysis and simulation results illustrate that the proposed algorithm has nearly similar complexity to SOBIUM algorithm, but better blind identification performance. Moreover, the blind identification performance of the proposed algorithm is superior to the FOBIUM in short data symbols (few sample numbers).

The organization structure of the remainder paper is as follows. The charrelation matrix is defined and its relevant properties is illustrated in section 2.The underdetermined blind identification problem and the CMBIUM algorithm are presented in section 3. Simulation results of the CMBIUM for different signal to noise ratio (SNR) and different length of data sample compared with SOBIUM and FOBIUM are presented in section 4. Concluding remarks are summarized in section 5.

Notation: Scalars are denoted by lower-case italic letters $(a,b,\ldots)$, vectors by lower-case boldface letters $(\mathbf{a},\mathbf{b},\ldots)$, matrices by boldface capitals $(\mathbf{A},\mathbf{B},\ldots)$ and tensor by caligraphic letters $(\mathcal{A},\mathcal{B},\ldots,)$. Italic captial are used to denote index upper bounds $(j=1,2\ldots,J)$. The entry with row index i and column index j in a matrix $\mathbf{A}$, i.e., $(\mathbf{A})_{ij}$,is symbolized by a_{ij}. Likewise, we have $(\mathcal{A})_{ijk}=a_{ijk}$. The columns of $\mathbf{A}$ are denoted by $\mathbf{a}_1,\mathbf{a}_2,\ldots$. Conversely, the matrix with columns $\mathbf{a}_1,\mathbf{a}_2,\ldots$ is denote by $\mathbf{A}$. The superscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote the transpose, the complex conjugate, and the complex conjugated transpose, respectively. $\mathbb{R}$ and $\mathbb{C}$ denote real number field and complex number field, respectively.

2 Charrelation Matrix

In this section, the definition and relevant properties of the charrelation matrix for complex random signal are presented. First, the charmean is illustrated, which can be considered as a generalization of expectation (mean) operator [17]. Then the charrelation matrix is illustrated.

Definition 1. (Charmean)

Given a random vector $\mathbf{x}\in\mathbb{C}^K$, and a function $\mathbf{g}(\cdot):\mathbb{C}^K\to\mathbb{C}^L$, the charmean of $\mathbf{g}(\mathbf{x})\in\mathbb{C}^L$ with respect to (w.r.t.) $\mathbf{x}$ at an arbitrary processing-point $\boldsymbol{\tau}\in\mathbb{C}^K$ is defined as:

$$\boldsymbol{\eta}_x\left[\mathbf{g}(\mathbf{x});\boldsymbol{\tau}\right]\triangleq\frac{E\left[\mathbf{g}(\mathbf{x})\exp\left\{\mathbf{x}^H\boldsymbol{\tau}\right\}\right]}{E\left[\exp\left\{\mathbf{x}^H\boldsymbol{\tau}\right\}\right]}\in\mathbb{C}^L \qquad (1)$$

Whenever both means (taken w.r.t. $\mathbf{x}$) exist.

The charmean shares many properties with the conventional expectation operator (e.g., linearity in $\mathbf{g}(\mathbf{x})$,separability in the case of the statistical independence), and for $\boldsymbol{\tau}=\mathbf{0}$ both operator coincide.

Definition 2. (Cross-charrelation and charrelation matrices)

Given a random vector $\mathbf{x}\in\mathbb{C}^K$ and function $\mathbf{g}_1(\cdot):\mathbb{C}^K\to\mathbb{C}^{L_1}$ and $\mathbf{g}_2(\cdot):\mathbb{C}^K\to\mathbb{C}^{L_2}$, the cross-charrelation matrix between $\mathbf{g}_1(\mathbf{x})$ and $\mathbf{g}_2(\mathbf{x})$ w.r.t. $\mathbf{x}$ at an arbitrary processing-point $\boldsymbol{\tau}\in\mathbb{C}^K$ is defined as:

$$\boldsymbol{\Psi}_{\mathbf{x}}\left[\mathbf{g}_1(\mathbf{x}),\mathbf{g}_2(\mathbf{x});\boldsymbol{\tau}\right]\triangleq\boldsymbol{\eta}_{\mathbf{x}}\left[\mathbf{g}_1(\mathbf{x})\mathbf{g}_2^H(\mathbf{x});\boldsymbol{\tau}\right]-\boldsymbol{\eta}_{\mathbf{x}}\left[\mathbf{g}_1(\mathbf{x});\boldsymbol{\tau}\right]\boldsymbol{\eta}_{\mathbf{x}}^H\left[\mathbf{g}_2(\mathbf{x});\boldsymbol{\tau}\right]\in\mathbb{C}^{L_1\times L_2} \qquad (2)$$

Whenever all the charmeans involved exist. Similary, for $\mathbf{g}(\cdot):\mathbb{C}^K\to\mathbb{C}^L$, the charrelation matrix of $\mathbf{g}(\mathbf{x})$ (w.r.t. $\mathbf{x}$, at $\boldsymbol{\tau}$) is simply defined as the cross-charrelation between $\mathbf{g}(\mathbf{x})$ and itself, namely $\boldsymbol{\Psi}_{\mathbf{x}}\left[\mathbf{g}(\mathbf{x});\boldsymbol{\tau}\right]\triangleq\boldsymbol{\Psi}_{\mathbf{x}}\left[\mathbf{g}(\mathbf{x}),\mathbf{g}(\mathbf{x});\boldsymbol{\tau}\right]$.

The charrelation matrix is a symmetric, positive semi-definite matrix, sharing many properties with the conventional covariance matrix. Both the cross-charrelation and charrelation matrices coincide

with the cross-covariance and covariance matrices (resp.) for $\boldsymbol{\tau}=0$. For $\mathbf{g}(\mathbf{x})=\mathbf{x}$, the charrelation matrix coincides with the Hessian (at $\boldsymbol{\tau}$) of the second generalized characteristic funtion of $\mathbf{x}$, namely, $\boldsymbol{\Psi}_{\mathbf{x}}[\mathbf{x};\boldsymbol{\tau}]=\frac{\partial^2}{\partial\boldsymbol{\tau}^*\partial\boldsymbol{\tau}^T}\log E\left[\exp\left\{\mathbf{x}^H\boldsymbol{\tau}\right\}\right]$.

The following addition properties are relatively straightforward to derive [17], and would be useful in our subsequent derivations:

Properties (Charrelation matrix)

(1) Linear transformations: If $\mathbf{C}\in\mathbb{C}^{L\times K}$ and $\mathbf{c}\in\mathbb{C}^{L}$ are some constant matrix and vector, and $\mathbf{y}=\mathbf{C}\mathbf{x}+\mathbf{c}\in\mathbb{C}^{L}$,then

$$\begin{aligned}\boldsymbol{\Psi}_{\mathbf{x}}[\mathbf{y};\boldsymbol{\tau}]&=\mathbf{C}\boldsymbol{\Psi}_{\mathbf{x}}[\mathbf{x};\boldsymbol{\tau}]\mathbf{C}^H\in\mathbb{C}^{L\times L}\\ \boldsymbol{\Psi}_{\mathbf{y}}[\mathbf{y};\boldsymbol{\tau}]&=\mathbf{C}\boldsymbol{\Psi}_{\mathbf{x}}\left[\mathbf{x};\mathbf{C}^H\boldsymbol{\tau}\right]\mathbf{C}^H\in\mathbb{C}^{L\times L}\end{aligned}\tag{3}$$

(2) Independence: If $\mathbf{x}\in\mathbb{C}^{K}$ can be partitioned into two statistically independent groups $\mathbf{x}_1\in\mathbb{C}^{K_1}$, $\mathbf{x}_2\in\mathbb{C}^{K_2}$ with $K_1+K_2=K$, then $\boldsymbol{\Psi}_{\mathbf{x}}[\mathbf{x};\boldsymbol{\tau}]\in\mathbb{C}^{K\times K}$ is block-diagonal (with the respective partition) for all $\boldsymbol{\tau}\in\mathbb{C}^{K}$ at which it exist.

Convenient estimates of the charmean and the charrelation matrix are obtained from the sample-charmean and the sample-charrelation (resp.):

$$\hat{\boldsymbol{\eta}}_{\mathbf{x}}[\mathbf{g}(\mathbf{x});\boldsymbol{\tau}]=\frac{\sum_{n=1}^{N}\mathbf{g}(\mathbf{x}[n])\exp\left\{\mathbf{x}^H[n]\boldsymbol{\tau}\right\}}{\sum_{n=1}^{N}\exp\left\{\mathbf{x}^H[n]\boldsymbol{\tau}\right\}}\tag{4}$$

$$\begin{aligned}\hat{\boldsymbol{\Psi}}_{\mathbf{x}}[\mathbf{g}(\mathbf{x});\boldsymbol{\tau}]=\hat{\boldsymbol{\eta}}_{\mathbf{x}}[\mathbf{g}(\mathbf{x})\mathbf{g}^H(\mathbf{x});\boldsymbol{\tau}]-\\ \hat{\boldsymbol{\eta}}_{\mathbf{x}}[\mathbf{g}(x);\boldsymbol{\tau}]\hat{\boldsymbol{\eta}}_{\mathbf{x}}^H[\mathbf{g}(x);\boldsymbol{\tau}]\end{aligned}\tag{5}$$

Though biased in general, both estimates are asymptotically unbiased and consistent. To simplify the exposition, we shall from now on use $\boldsymbol{\Psi}_{\mathbf{x}}(\boldsymbol{\tau})$ as shorthand for $\boldsymbol{\Psi}_{\mathbf{x}}(\mathbf{x};\boldsymbol{\tau})$ in instructions below.

3 Charrealtion Matrix Based Underdetermined Blind Identification

3.1 Problem formulation

Consider the following basic linear mixture model:

$$\mathbf{x}(t)=\mathbf{A}\mathbf{s}(t)+\mathbf{n}(t)\tag{6}$$

The stochastic vector $\mathbf{x}(t)\in\mathbb{C}^{J}$ represents the observation signals, the components of the stochastic vector $\mathbf{s}(t)\in\mathbb{C}^{Q}$ correspond to unobserved source signals, and $\mathbf{n}(t)\in\mathbb{C}^{J}$ denotes additive Gaussian noise. The unknown mixing matrix $\mathbf{A}\in\mathbb{C}^{J\times Q}$ characterizes the way that the sources are acquired by the sensors. The aim of blind identification is to estimate the mixing matrix from the observations based on the assumption that the source signals are statistically independent. The mixing matrix obtained may in turn be used to estimate the original source signals from the observations. The overdetermined case has been addressed in most of literature about BSS, where $J\geq Q$. In this paper, the underdetermined case is considered, where $J<Q$.

3.2 The proposed algorithm: CMBIUM

According to the properties of charrelation matrix in the previous section Ⅱ, the model funtion is derived. The model function is a function relationship between the mixing matrix and the observation' charrelation matrices, evaluated at K arbitrary processing-points $\boldsymbol{\tau}_1,\ldots,\boldsymbol{\tau}_K$. From property 1 above of Charrelation matrix, and from the BSS model $\mathbf{x}=\mathbf{A}\mathbf{s}$ (For convenience, without loss of generality, the noise $\mathbf{n}$ is ignored except when running simulation experiments), we have

$$\boldsymbol{\Psi}_{\mathbf{x}}(\boldsymbol{\tau})=\mathbf{A}\boldsymbol{\Psi}_{\mathbf{s}}\left(\mathbf{A}^H\boldsymbol{\tau}\right)\mathbf{A}^H\tag{7}$$

Moreover, from property 2, the charrelation matrix $\boldsymbol{\Psi}_{\mathbf{s}}\left(\mathbf{A}^H\boldsymbol{\tau}\right)$ of the random vector $\mathbf{s}$ with mutually independent elements is strictly diagonal (more detail can be found in Appendix A). Consequently, the charrelation matrices of the observations satisfy when the K arbitrary processing-points is considered.

$$\begin{aligned}\boldsymbol{\Psi}_{\mathbf{x}}^1&=\boldsymbol{\Psi}_{\mathbf{x}}(\boldsymbol{\tau}_1)=\mathbf{A}\boldsymbol{\Psi}_{\mathbf{s}}\left(\mathbf{A}^H\boldsymbol{\tau}_1\right)\mathbf{A}^H\\ &\vdots\\ \boldsymbol{\Psi}_{\mathbf{x}}^K&=\boldsymbol{\Psi}_{\mathbf{x}}(\boldsymbol{\tau}_K)=\mathbf{A}\boldsymbol{\Psi}_{\mathbf{s}}\left(\mathbf{A}^H\boldsymbol{\tau}_K\right)\mathbf{A}^H\end{aligned}\tag{8}$$

Let us stack the matrices $\boldsymbol{\Psi}_{\mathbf{x}}^1,\ldots,\boldsymbol{\Psi}_{\mathbf{x}}^K$ in model function (8) in a tensor $\mathcal{M}\in\mathbb{C}^{J\times J\times K}$, where $(\mathcal{M})_{ijk}\triangleq\left(\boldsymbol{\Psi}_{\mathbf{x}}^k\right)_{ij}$, $i=1,\ldots,J$, $j=1,\ldots,J$, $k=1,\ldots,K$. Define a matrix $\mathbf{D}\in\mathbb{C}^{K\times Q}$ by $(\mathbf{D})_{kq}\triangleq\left(\boldsymbol{\Psi}_{\mathbf{s}}\left(\mathbf{A}^H\boldsymbol{\tau}_k\right)\right)_{qq}$, $k=1,\ldots,K$,

$q = 1, \ldots, Q$. Then we have

$$m_{ijk} = \sum_{q=1}^{Q} a_{iq} a_{jq}^{*} d_{kq} \tag{9}$$

which we can write as

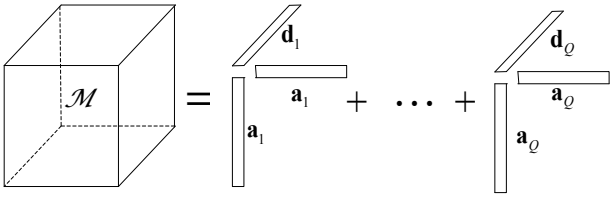

Fig 1. Rank-Q Tensor decomposition

$$\mathcal{M} = \sum_{q=1}^{Q} \mathbf{a}_q \circ \mathbf{a}_q^{*} \circ \mathbf{d}_q \tag{10}$$

in which " $\circ$ " denotes the outer product and in which $\{\mathbf{a}_q\}$ and $\{\mathbf{d}_q\}$ are the columns of $\mathbf{A}$ and $\mathbf{D}$, respectively. Equation (10) is decomposition of tensor $\mathcal{M}$ in a sum Q rank-1 terms, as shown in fig. 1. In the literature, this is called a "Canonical Decomposition" (CANDECOMP) or "Parallel Factors Model" (PARAFAC)[18, 19]. The minimal number of rank-1 tensors in which a higher-order tensor can be decomposed, is called its rank. Note that each rank-1 term in (10) is consisted of the contribution of one distinct source to $\mathcal{M}$. Therefore, in terms of this tensor, "source separation" corresponds to the computation of decomposition (10), provided it is unique. Tensor rank-1 decomposition can be unique, which allows for the determination of the mixing matrix (up to a scaling and permutation of its columns) in the underdetermined case.

A important uniqueness condition needs the notion of Kruskal-rank or *k*-rank $k(\mathbf{A})$ of a matrix $\mathbf{A}$. It is defined as the maximal number of k such that any set of k columns of $\mathbf{A}$ is linearly independent [18, 19]. We can acquire the condition that decomposition (10) is essentially unique when

$$2k(\mathbf{A}) + k(\mathbf{D}) \geq 2(Q+1) \tag{11}$$

The standard way to compute tensor rank-1 decomposition, is by means of an "Altenating Least Squares (ALS)" algorithm [20, 21]. More specifically, one optimizes the cost function as follows.

$$\min_{\mathbf{A},\mathbf{D}} \left\| \mathcal{M} - \sum_{q=1}^{Q} \mathbf{a}_q \circ \mathbf{a}_q^{*} \circ \mathbf{d}_q \right\|_F^2 \tag{12}$$

Due to multiliearity of the model, the estimation of one of the arguments, given the other two, is a classical linear least squuares problem.The aim is to minimize the (squared) Frobenius norm of the difference between $\mathcal{M}$ and its estimated decomposition in rank-1 terms by means of an iteration in which each step consists of fixing a subset of unknown parameters to their current estimates, and optimizing w.r.t. the remaining unknowns, followed by fixing another subset of parameters, and optimizing w.r.t. the complimentary set, etc. Although the ALS algorithm of literature [21, 21] can direct be applied in previous optimization problem (12), the global optimum may not be found (especially, the stacked tensor $\mathcal{M}$ is ill-conditioned). Therefore, we consider the way of matrix decomposition to solve the problem (12).

In such way, we consider the pervious $(J \times J \times K)$ —tensor $\mathcal{M}$ is given by (9) or (10), in which $\mathbf{A} \in \mathbb{C}^{J \times Q}$, $\mathbf{D} \in \mathbb{C}^{K \times Q}$. We assume that $\min(J^2, K) \geq Q$. Consider a matrix $\mathbf{M} \in \mathbb{C}^{J^2 \times K}$ in which the entrier of $\mathcal{M}$ are stacked as follows:

$$(\mathbf{M})_{(i-1)J+j,k} = m_{ijk} \tag{13}$$

We have

$$\mathbf{M} = (\mathbf{A} \odot \mathbf{A}^{*}) \mathbf{D}^{T} \tag{14}$$

We assume that both $\mathbf{A} \odot \mathbf{A}^{*}$ and $\mathbf{D}$ are full column rank. Both conditions are generically satisfied if $\min(J^2, K) \geq Q$. Note that, in this case, the rank of the tensor $\mathcal{M}$ is equal to the rank of its matrix representation $\mathbf{M}$. Consider a factorization of $\mathbf{M}$ the form

$$\mathbf{M} = \mathbf{E}\mathbf{F}^{H} \tag{15}$$

with $\mathbf{E} \in \mathbb{C}^{IJ \times Q}$ and $\mathbf{F} \in \mathbb{C}^{K \times Q}$ full column rank. Because of (13) and (14), we have

$$\mathbf{A} \odot \mathbf{A}^{*} = \mathbf{E}\mathbf{W} \tag{16}$$

For some nonsingular $\mathbf{W} \in \mathbb{C}^{Q \times Q}$. The task is now to seek for $\mathbf{W}$ such that the columns of $\mathbf{E}\mathbf{W}$ are Kronecker products [15]. A vector that is equal to the Kronecker product of a vector $\mathbf{a} \in \mathbb{C}^{I}$ and a vector $\mathbf{b} \in \mathbb{C}^{J}$ can be represented as an $(I \times J)$ rank-1 matrix. Matrices with rank at most 1 and matrices of which the introduced in the following theorem .

Theorem 1 [15]. Consider the mapping $\Phi: (\mathbf{X}, \mathbf{Y}) \in \mathbb{C}^{I \times J} \times \mathbb{C}^{I \times J} \to \Phi(\mathbf{X}, \mathbf{Y}) \in \mathbb{C}^{I \times I \times J \times J}$ defined by

$$\left(\Phi(\mathbf{X}, \mathbf{Y})\right)_{ijkl} = x_{ik} y_{jl} + y_{ik} x_{jl} - x_{il} y_{jk} - y_{il} x_{jk} \tag{17}$$

Then we have $\Phi(\mathbf{X}, \mathbf{X}) = O$ if and only if $\mathbf{X}$ is at most rank 1.

Define matrices $\mathbf{E}_1, \ldots, \mathbf{E}_Q \in \mathbb{C}^{J \times J}$ corresponding to each column of $\mathbf{E}$ in (14) so that

$$\left(\mathbf{E}_q\right)_{ij} = e_{(i-1)+j,\mathrm{q}} \quad \forall i,j,\mathrm{q} \tag{18}$$

And let $\mathcal{P}_{qs} = \Phi\left(\mathbf{E}_q, \mathbf{E}_s\right)$. Note that Φ is symmetric in its arguments; hence

$$\mathcal{P}_{qs} = \mathcal{P}_{sq} \quad \forall q,s \tag{19}$$

Since Φ is bilinear, we have from (15)

$$\mathcal{P}_{qs} = \sum_{v,u=1}^{R} \left(\mathbf{W}^{-1}\right)_{vq} \left(\mathbf{W}^{-1}\right)_{us} \Phi\left(\mathbf{a}_v\mathbf{a}_v^H, \mathbf{a}_u\mathbf{a}_u^H\right) \tag{20}$$

Assume at this point that there exists a symmetric matrix $\mathbf{G}$ of which the entries satisfy the following set of homogeneous linear equations[15].

$$\sum_{q,s=1}^{R} g_{qs}\mathcal{P}_{qs} = O \tag{21}$$

Substitution of (19) in (20) yields

$$\sum_{\mathrm{q},s=1}^{R}\sum_{\mathrm{v},u}^{R}\left(\mathbf{W}^{-1}\right)_{vq}\left(\mathbf{W}^{-1}\right)_{us} g_{qs}\Phi\left(\mathbf{a}_v\mathbf{a}_v^H, \mathbf{a}_u\mathbf{a}_u^H\right) = O \tag{22}$$

According to Theorem 1, we have $\Phi\left(\mathbf{a}_v\mathbf{a}_v^H, \mathbf{a}_v\mathbf{a}_v^H\right) = O$, $1 \le v \le Q$. Hence

$$\sum_{q,s=1}^{R}\sum_{\substack{\mathrm{v},u=1 \\ v\neq u}}^{R}\left(\mathbf{W}^{-1}\right)_{vq}\left(\mathbf{W}^{-1}\right)_{us} g_{qs}\Phi\left(\mathbf{a}_v\mathbf{a}_v^H, \mathbf{a}_u\mathbf{a}_u^H\right) = O \tag{23}$$

Furthermore, due to (18)and the symmetry of $\mathbf{G}$ we have

$$\sum_{q,s=1}^{R}\sum_{\substack{\mathrm{v},u=1 \\ v<u}}^{R}\left(\mathbf{W}^{-1}\right)_{vq}\left(\mathbf{W}^{-1}\right)_{us} g_{qs}\Phi\left(\mathbf{a}_v\mathbf{a}_v^H, \mathbf{a}_u\mathbf{a}_u^H\right) = O \tag{24}$$

Denote

$$\lambda_{vu} = \sum_{q,s=1}^{R}\left(\mathbf{W}^{-1}\right)_{vq}\left(\mathbf{W}^{-1}\right)_{us} g_{qs} \tag{25}$$

Let us now make the crucial assumption that the tensors $\Phi\left(\mathbf{a}_v\mathbf{a}_v^H, \mathbf{a}_u\mathbf{a}_u^H\right), 1 \le v < u \le R$, are linearly independent. Then (23) implies that $\lambda_{vu} = 0$ when $v \neq u$. As a consequence, (24) can be written in a matrix format as

$$\mathbf{G} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^H \tag{26}$$

in which $\mathbf{\Lambda}$ is diagonal. Actually, one can see that any diagonal matrix $\mathbf{\Lambda}$ generates a matrix $\mathbf{G}$ that satisfies (20). Hence, if the tensors $\left\{\Phi\left(\mathbf{a}_v\mathbf{a}_v^H, \mathbf{a}_u\mathbf{a}_u^H\right)\right\}_{v<u}$ are linearly independent, these matrices form an Q-dimensional subspace of the symmetric $Q\times Q$ matirces. Let $\left\{\mathbf{G}_q\right\}$ represent a basis of this subspace. We have

$$\begin{aligned} \mathbf{G}_1 &= \mathbf{W}\mathbf{\Lambda}_1\mathbf{W}^H \\ &\vdots \\ \mathbf{G}_Q &= \mathbf{W}\mathbf{\Lambda}_Q\mathbf{W}^H \end{aligned} \tag{27}$$

in which $\mathbf{\Lambda}_1,\ldots,\mathbf{\Lambda}_Q$ are diagonal. The matrix $\mathbf{W}$ can be determined from this simultaneous matrix decomposition by means of the algorithms presented in [22- 25].

Once $\mathbf{W}$ is known, $\mathbf{A}\odot\mathbf{A}^*$ can be obtained from (15). Let the columns of $\mathbf{A}\odot\mathbf{A}^*$ be mapped to $\left(J\times J\right)$ matrices $\mathbf{H}_q$ as follows:

$$\left(\mathbf{H}_q\right)_{ij} = \left(\mathbf{A}\odot\mathbf{A}^*\right)_{(i-1)J+j,\mathrm{q}}, q = 1,\ldots,Q \tag{28}$$

Then we have

$$\mathbf{H}_q = \mathbf{a}_q\mathbf{a}_q^H, q = 1,\ldots,Q \tag{29}$$

from which the mixing matrix $\mathbf{A}$ can be obtained.

The proposed agorithm can be summarized as follows:

(1) According to the defintion and properties, the model function $\mathbf{\Psi}_{\mathbf{x}}\left(\boldsymbol{\tau}_i\right) = \mathbf{A}\mathbf{\Psi}_{\mathbf{s}}\left(\mathbf{A}^H\boldsymbol{\tau}_i\right)\mathbf{A}^H \quad , i = 1,\ldots,K$ is established at K processing-points $\boldsymbol{\tau}_1,\ldots,\boldsymbol{\tau}_K$.

(2) Stack K charrelation matrices of the observations as $\mathcal{M} \in \mathbb{C}^{J\times J\times K}$ satisfying $\mathcal{M} = \sum_{q=1}^{Q}\mathbf{a}_q \circ \mathbf{a}_q^* \circ \mathbf{d}_q$, with both $\left(\mathbf{d}_q\right)_{1\le q\le Q}$ and $\Phi\left(\mathbf{a}_v\mathbf{a}_v^H, \mathbf{a}_u\mathbf{a}_u^H\right)_{1\le v<u\le R}$ linearly independent.

(3) Stack $\mathcal{M}$ in $\mathbf{M} \in \mathbb{C}^{J^2\times K}$ as follows: $\left(\mathbf{M}\right)_{(i-1)J+j,k} = m_{ijk} \quad , \forall i,j,k \quad . Q = rank\left(\mathbf{M}\right)$;

(4) Compute factorization $\mathbf{M} = \mathbf{E}\mathbf{F}^H$ with $\mathbf{E} \in \mathbb{C}^{J^2\times Q}$ and $\mathbf{F} \in \mathbb{C}^{K\times Q}$ full column rank.

(5) Stack $\mathbf{E}$ in $\mathcal{E} \in \mathbb{C}^{J\times J\times Q}$ as follows: $\left(\mathcal{E}\right)_{ijq} = \left(\mathbf{E}\right)_{(i-1)J+j,\mathrm{q}} \quad \forall i,j,\mathrm{q}$

(6) Compute $\mathcal{P}_{qs} \in \mathbb{C}^{J\times J\times J\times J}, 1 \le q,s \le Q$, as follows: $\left(\mathcal{P}_{qs}\right)_{ijkl} = e_{ikq}e_{jls} + e_{iks}e_{jlq} - e_{ilq}e_{jks} - e_{ils}e_{jkq}$, $\forall i,j,k,l$

(7) Compute the kernel of $\sum_{s,\mathrm{v}=1}^{R} g_{sv}\mathcal{P}_{sv} = O$, under the constraint $g_{sv} = g_{vs} \ \forall s,v$. Stack Q linearly independent solutions in symmetric matrices

$\mathbf{G}_1,\ldots,\mathbf{G}_Q \in \mathbb{C}^{Q\times Q}$.

(8) Determined $\mathbf{W}\in\mathbb{C}^{Q\times Q}$ that simultaneously diagonalizes

$$\mathbf{G}_1 = \mathbf{W}\mathbf{\Lambda}_1\mathbf{W}^H$$
$$\vdots$$
$$\mathbf{G}_Q = \mathbf{W}\mathbf{\Lambda}_Q\mathbf{W}^H$$

(9) $\mathbf{A}\odot\mathbf{A}^* = \mathbf{E}\mathbf{W}$, $\mathbf{D} = \mathbf{F}\mathbf{W}^{-H}$. Stack $\mathbf{A}\odot\mathbf{A}^*$ in $\mathbf{H}_1,\ldots,\mathbf{H}_Q \in \mathbb{C}^{J\times J}$ as follows:

$$\left(\mathbf{H}_q\right)_{ij} = \left(\mathbf{A}\odot\mathbf{A}^*\right)_{(i-1)J+j,q}$$

(10) Obtain $\mathbf{a}_q$ from $\mathbf{H}_q = \mathbf{a}_q\mathbf{a}_q^H, q=1,\ldots,Q$, then the mixing matrix is $\mathbf{A} = \left[\mathbf{a}_1,\ldots,\mathbf{a}_Q\right]$.

3.3 Performance Evaluation

In order to demonstrate the superiority of the proposed algorithm, we consider aspects of complexity and statistical characteristic to evaluate the performance of the proposed CMBIUM algorithm compared to SOBIUM and FOBIUM. On the one hand, the proposed algorithm has relatively equivalent complexity to the SOBIUM from the complexity point. Becasue both of them are from the two-dimension algebraic structure in covariance matrix and charrelation matrix respectively. If the multiplication operation in matrix is mainly considered, both of them are nearly $O\left(J^2Q^3K\right)$. Likewise, the case in FOBIUM is $O\left(J^4Q^4K\right)$ due to the four-dimension algebraic structure in cumulant tensor.

On the other hand, FOBIUM can suppress the influence of additive Gaussian noise, but that of case is failure to SOBIUM from the statistical characteristic point. However, there is a negative factor for FOBIUM that the length of samples influence the its performance due to higer-order statistical estimation problem. The proposed algorithm can improve the pervious negative factor with help of charrelation matrix. Because the charrelation matrix offer the structural simplicity and controllable statistical stability of second order statistics on the one hand, while retaining higher-order statistical information on the other hand. In next section, the simulation experiments are carried out to verify the pervious exposure.

4 Simulation Results

To demonstrate the effectiveness of the proposed algorithm CMBIUM, we conduct simulation experiments to evaluate the performance of CMBIUM. For making a comparison, the SOBIUM and FOBIUM are also illustrated. The peroformance of the tested algorithms is evaluated and compared in terms of the relative error performance index (PI) as a function of the sample size and the signal-to-noise ratio (SNR) of the observations. The relative error PI is given by as [14]

$$PI = E\left\{\left\|\mathbf{A}-\hat{\mathbf{A}}\right\|\Big/\left\|\mathbf{A}\right\|\right\} \tag{30}$$

In which the norm is the Frobenius norm and $\hat{\mathbf{A}}$ denotes the optimally ordered and scaled estimate of the mixing matrix $\mathbf{A}$.

We consider $Q=5$ narrow sources, received by a Uniform Circular Array (UCA) of $J=4$ identical sensors of radius R_a. Considering a free space propagation model, the entries of the mixing matrix before normalization are given by

$$a_{jq} = \exp\left(2\pi i\left(\alpha_j\cos\left(\theta_q\right)\cos\left(\phi_q\right)+\beta_j\cos\left(\theta_q\right)\sin\left(\phi_q\right)\right)\right)$$

where $\alpha_j = \left(R_a/\lambda\right)\cos\left(2\pi\left(j-1\right)/J\right)$, $\beta_j = \left(R_a/\lambda\right)\sin\left(2\pi\left(j-1\right)/J\right)$, and $i=\sqrt{-1}$. We have $R_a/\lambda = 0.55$. The mixing matrix $\mathbf{A}$ is obtained by dividing the columns of $\mathbf{A}$ by their Frobenius norm. The sources are unit-variance quadrature phase-shift keying (QPSK) with a uniform distribution, shaped by a raise cosine pulse shape filter with roll-off $\rho=0.3$. All sources have the same symbol duration $T=4T_e$, where T_e is sample period. The directions-of-arrival (DOAs) of the different sources are given by $\theta_1=3\pi/10$, $\theta_2=3\pi/10$, $\theta_3=2\pi/5$, $\theta_4=0$, $\theta_5=\pi/10$ and $\phi_1=7\pi/10$, $\phi_2=9\pi/10$, $\phi_3=3\pi/5$, $\phi_4=4\pi/5$, $\phi_5=3\pi/5$. The observations are contaminated by additive zeros-mean complex Gaussian noise. In such case, an equivalent underdetermined MIMO system is attained. The charrelation matrices of the observations are computed for $K=12\left(K>Q\right)$ different processing-point. The prossing-points $\boldsymbol{\tau}$ are randomly drawn from in the range [-1;1]. N denotes the number of data symbol.

Fig. 2 shows the relative error PI as a function of SNR when the number of symbols is $N=1000$. From the Fig. 2, we can see that the proposed algorithm is superior to SOBIUM and FOBIUM as

analysed result. We know that the length of the data symbols influences the performance due to the problem estimation of statistical information sensitive to the length of samples. The proposed algorithm and SOBIUM is better than FOBIUM when the length of data symbols is short.

Fig. 3 shows the relative error PI as a function of the number of data symbols, when the SNR is 10dB. According the Fig. 3, we obtain the result that the proposed algorithm CMBIUM has better performance than the SOBIUM and FOBIUM algorithm.

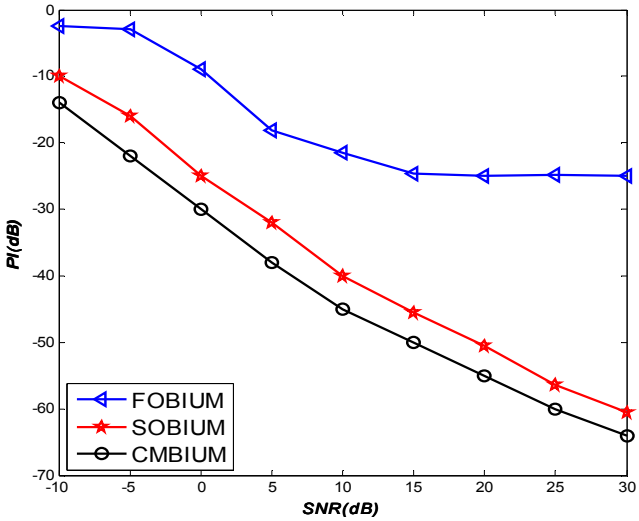

Fig.2 The relative error PI versus SNR

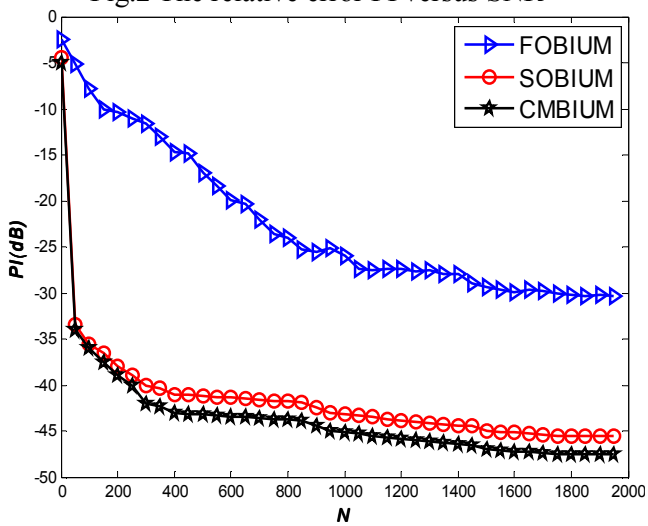

Fig. 3 The relative error PI versus length of the data symbols

Taking into the previous simulation results account, we can summarise some reasons for those results. For one thing, for the estimation of the fourth-order cumulants in FOBIUM more samples are required than for the estimation of the charrelation matrices in CMBIUM and covariance matrices in SOBIUM. For another, the charrelation matrix is generalized covariance matrix and also contain higher-order information to improve performance.

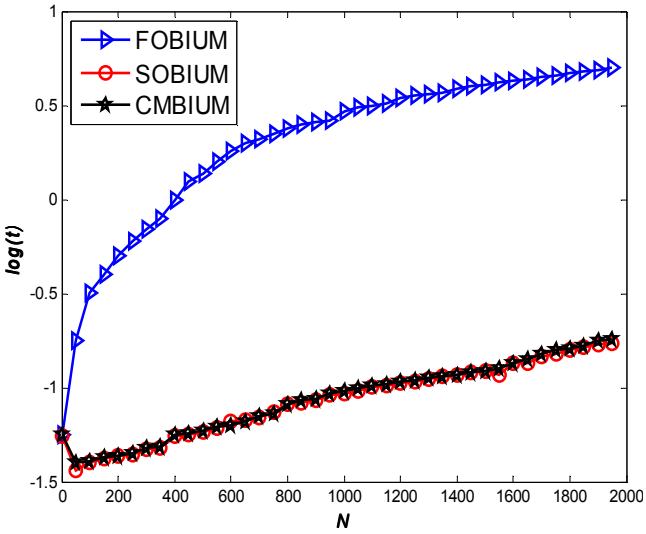

Fig. 4 Computation time versus length of the data symbols

Fig. 4 shows the compution time as a function of the number of data symbols, when the SNR is 10dB. We can see that the proposed algorithm has nearly similar computation time as SOUBIUM and the FOBIUM is higher than them in Fig. 4. According to Fig. 4, we can know that the complexity of computation becomes worse with the increasing of the symbol numbers.

5 Conclusions

The new statistical tool (charrelation matrix) is used to ameliorate the performance of underdetermined blind identification method. Due to the superior properties of charrelation matrix, we develop a undetermined blind identification algorithm based on this new tool. The underdetermined blind identification problem is described in terms of rank-1 decomposition of a three-order tensor. The proposed algorithm has better performance of blind identification than SOBIUM algorithm, with essentially similar complexity. Moreover, the performance of blind identification of the proposed algorithm is superior to the FOBIUM algorithm in fewer sample numbers. The proposed algorithm would further be extended in undetermined MIMO systems in future work.

Appendix A Diagonalization Verification

Let $\varphi_s(\tilde{\boldsymbol{\tau}})$ denote the "generalized characteristic function" of source signal $\mathbf{s}(t)$. Due to the statistical independence of elements of $\mathbf{s}(t)=\left[s_1(t),\cdots,s_Q(t)\right]^T$, we get

$$\varphi_{\mathbf{s}}(\tilde{\boldsymbol{\tau}}) = \varphi_{s_1}(\tilde{\tau}_1) \cdot \varphi_{s_2}(\tilde{\tau}_2) \cdots \varphi_{s_Q}(\tilde{\tau}_Q) \quad (31)$$

Where $\varphi_{s_i}(\tilde{\tau}_i) = E\left[\exp\left(\tilde{\tau}_i^* s_i(t)\right)\right], i = 1, \cdots, Q$ .Defining $\delta_{\mathbf{s}}(\tilde{\boldsymbol{\tau}}) = \log \varphi_{\mathbf{s}}(\tilde{\boldsymbol{\tau}})$ is called "second generalized characteristic function" of source signal $\mathbf{s}(t)$. Hence, we obtain

$$\delta_{\mathbf{s}}(\tilde{\boldsymbol{\tau}}) = \delta_{\mathbf{s}_1}(\tilde{\tau}_1) + \delta_{\mathbf{s}_2}(\tilde{\tau}_2) + \cdots + \delta_{\mathbf{s}_Q}(\tilde{\tau}_Q) \quad (32)$$

Consequently, the charrelation matrix (Hessian matrix) $\boldsymbol{\Psi}_{\mathbf{s}}(\tilde{\boldsymbol{\tau}})$ can be easily gained

$$\boldsymbol{\Psi}_{\mathbf{s}}(\boldsymbol{\tau}) = \nabla_{\boldsymbol{\tau}^T}\left[\nabla_{\boldsymbol{\tau}^*}\delta_{\mathbf{s}}(\boldsymbol{\tau})\right] = \frac{\partial}{\partial \boldsymbol{\tau}^T}\left[\frac{\partial \delta_{\mathbf{s}}(\boldsymbol{\tau})}{\partial \boldsymbol{\tau}^*}\right]$$

$$= diag\left(\frac{\partial \delta_{s_1}(\tilde{\tau}_1)}{\partial \tilde{\tau}_1 \partial \tilde{\tau}_1^*}, \frac{\partial \delta_{s_2}(\tilde{\tau}_2)}{\partial \tilde{\tau}_2 \partial \tilde{\tau}_2}, \cdots, \frac{\partial \delta_{s_Q}(\tilde{\tau}_Q)}{\partial \tilde{\tau}_Q \partial \tilde{\tau}_Q}\right) \quad (33)$$

Where $diag(\cdot)$ indicates a diagonal matrix with elements on the main diagonal.

Acknowledgments

This work is fully supported by a grant from the national High Technology Research and development Program of China (863 Program) (No. 2012AA01A502), and National Natural Science Foundation of China (No. 61179006), and Science and Technology Support Program of Sichuan Province (No. 2014GZX0004).

References:

[1] M.G.S. Sriyananda, Jyrki Joutsensalo, Timo Hämäläinen, "Blind Source Separation for OFDM with Filtering Colored Noise and Jamming Signal", *Journal of Communications and Networks*, Vol. 14, No. 4, 2012, pp. 410-417.

[2] Jiang Zhang, Hang Zhang and Zhifu Cui, "Dual-antenna-based blind joint hostile jamming cancellation and multi-user detection for uplink of asynchronous direct-sequence code-division multiple access systems", *IET Commun.*, Vol. 7, No. 10, 2013, pp. 911-921.

[3] M.G.S. Sriyananda, Jyrki Joutsensalo, Timo Hämäläinen, "Interference Cancellation Schemes for Spread Spectrum Systems with Blind Principles", *in Proc. 27th International Conference on Advanced Information Networking and Applications*, Barcelona, 25-28 March, 2013, pp. 1078–1082.

[4] Jingbo Gao, Xu Zhu, and Asoke K. Nandi, "PAPR Reduction in Blind MIMO OFDM System Based on Independent Component Analysis", *in Proc. 17th Eutopean Signal Processing Conference, Glasgow*, Scotland, Aug. 24-28, 2009, pp. 318-322.

[5] FU Wei-hong, YANG Xiao-niu, LIU Nai-an, "The Multi-User Detection and Chip Sequence Estimation for CDMA System Based on the Blind Source Separation", *ACTA ELECTRONICA SINICA*, Vol. 36, No. 7, 2008, pp. 1319-1323.

[6] K. Raju, T. Ristaniemi, J. Karhunen, and E. Oja, "Jammer Suppression in DS-CDMA Arrays Using Independent Component Analysis", *IEEE Trans. on Wirless Commun.*, Vol. 5, No. 1, 2006, pp. 77-82.

[7] E. Gopalakrishna Sarma, Sakuntala S. Pillai, "A Robust Technique for Blind Multiuser detection in Fading Channels", *International Journal of Hybrid Information Technology*, Vol. 4, No. 2, Apr. 2011, pp. 13-22.

[8] NIU Dezhi, CHEN Changxing, WANG Shuzhao, et al., "Research on channel estimation of MIMO system based on FastICA", *Computer Engineering and Applications*, Vol. 48, No. 6, 2012, pp.106-109.

[9] Chih-Wei Jen, Shyh-Jye Jou, "Blind ICA detection based on second-order cone programming for MC-CDMA system", *EURASIP Journal on Advances in Signal Processing*, Vol.151, 2014, pp. 1-14.

[10] F. Abrard, Y. Deville, "A time-frequency blind signal separation method applicable to underdetermined mixtures of dependent sources", *Signal Process.*, Vol. 85, No. 7, Jul. 2005, pp. 1389-1403.

[11] P. Bofill, M. Zibulevsky, "Underdetermined blind source separation using sparse representations", *Signal Process.*, Vol. 81, No. 11, Nov. 2001, pp. 2353-2362.

[12] SangGyun Kim, Chang D.Yoo, "Underdetermined Blind Source Separation Based on Subspace Representation", *IEEE Trans. Signal Process.* , Vol. 57, No. 7, Jul. 2009, pp. 2604-2614.

[13] Shengli Xie, Liu Yang, Jun-Mei Yang, etc., "Time-Frequency Approach to Underdetermined Blind Source Separation", *IEEE Trans. on Neural Networks and Learning Systems*, Vol. 23, No. 2, Feb. 2012, pp. 306-316.

[14] Lieven De Lathauwer, Joséphine, "Second-Order Blind Identification of Underdetermined Mixtures", *in Proc. Independent Component Analysis and Blind Signal Separation (ICA)*, J.Rosca, Ed.,

Charleston, SC, Vol. 3889, Mar. 5-8, 2006, pp. 40-47.

[15] Lieven De Lathauwer, Joséphine, "Blind Identification of Underdetermined Mixtures by Simultaneous Matrix Diagonalization", *IEEE Trans. on Signal Processing*, Vol. 56, No. 3, 2008, pp. 1096-1105.

[16] A. Ferréol, L. Albera, and P. Chevalier, "Fourth order blind identification of underdetermined mixtures of sources", *IEEE Trans. on Signal Process.* , Vol. 53, No. 5, 2005, pp. 1640-1653.

[17] Alon Slapak and Arie Yeredor, "Charrelation and Charm: Generic Statistics Incorporating Higher-Order Information", *IEEE Trans. on Signal Processing*, Vol. 60, No. 10, 2012, pp. 5089-5106.

[18] A. Stegeman, N. D. Sidiropoulos, "On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition", *Linear Algebr. Applicat.*, Vol. 420, No. 2-3, Jan. 2007, pp. 540-552.

[19] Pierre Comon, "Tensors", *IEEE Signal Processing Magazine*, 2014, pp. 44-53.

[20] LD Lathauwer, BD Moor, J Vandewalle, "On the best rank-1 and rank-(*R1, R2, ..., RN*) approximation of higher-order tensor", SIAM J. Matrix Anal. Appl., Vol. 21, 2000, pp. 1324-1342.

[21] M. Rajih, P.Comon, "Enhanced line search: A novel method to accelerate PARAFAC", *SIAM J. Matrix Anal. & Appl.,* Vol. 30, No. 3, 2008, pp. 1128-1147.

[22] Ammar Mesloub, Karim Abed-Meraim, and Adel Belouchrani, "A New Algorithm for Complex Non-Orthognal Joint Diagonalization Based on Shear and Givens Rotations", *IEEE Trans. Signal Process.*, Vol. 62, No. 8, Apr. 2014, pp. 1913-1925.

[23] A. Yeredor, "Non-orthognal joint diagonalization in the least-squares sense with application in blind source separation", *IEEE Trans. Signal Process.*, Vol. 50, No. 7, Jul. 2002, pp. 1545-1552.

[24] P. Tichavsky and A. Yeredor, "Fast approximate joint diagonalization incorporating weight matrices", *IEEE Trans. Signal Process.* , Vol. 57, No. 3, Mar. 2009, pp. 878-891.

[25] K. Wang, X. F Gong, and Q. Lin, "Complex non-orthogonal joint diagonalization based on LU and LQ decomposition", *in Proc. 10th Int. Conf. Latent Variable Anal. Signal Separat.*, Mar. 2012, pp. 50-57.