Proceedings of the 2013 Federated Conference on
Computer Science and Information Systems pp. 19–25

# Underdetermined Blind Separation of an Unknown Number of Sources Based on Fourier Transform and Matrix Factorization

Ossama S. Alshabrawy*, Mohamed E. Ghoneim*
Damietta University, faculty of Science, Department of Mathematics & Computer Science, 34517, Damietta, Egypt

A. A. Salama
Port Said University, faculty of Science, Department of Mathematics & Computer Science, Port Said, Egypt

Aboul Ella Hassanien*
Cairo University, faculty of Computers and Information, Department of Information Technology, Cairo, Egypt

* Scientific Research Group in Egypt (SRGE), http://www.egyptscience.net

***Abstract*—This paper presents an approach for underdetermined blind source separation that can be applied even if the number of sources is unknown. Moreover, the proposed approach is applicable in the case of separating I+3 sources from I mixtures without additive noise. This situation is more challenging and suitable to practical real world problems. Also, the sparsity conditions are not imposed unlike to those employed by some conventional approaches. Firstly, the number of source signals are estimated followed by the estimation of the mixing matrix based on the use of short time Fourier transform and rough-fuzzy clustering. Then, source signals are normalized and recovered using modified Lin's projected gradient algorithm with modified Armijo rule. The simulation results show that the proposed approach can separate I+3 source signals from I mixed signals, and it has superior evaluation performance compared to conventional approaches.**

***Keywords*—Underdetermined Blind Source Separation; Rough Fuzzy clustering; Short Time Fourier transform; Lin's Projected Gradient; Armijo rule**

## I. Introduction

BLIND Signal Separation (or Blind Source Separation, BSS) has received a great deal of attention in the fields of digital communication systems, speech processing, medical imaging, water marking, biomedical engineering, and data mining [3]-[7] in recent years in combination with artificial neural networks, information theory, and computer science applications. Blindness or blind separation means that no or very little information is known about the source signals or the mixing system [1].

The objective of BSS is to extract original source signals using only the information gathered from observed signals with no or very limited knowledge about the source signals or the mixing system. The approaches developed by researchers in the last few years can be classified into two methodologies, namely over-determined BSS and underdetermined BSS, according to the number of source signals and observable mixed signals [20]. BSS that has fewer sensors or observable mixed signals than source signals is called underdetermined BSS while a BSS that has more sensors than sources is called over-determined BSS, Underdetermined BSS is challenging and is more realistic in practical situations. However, most approaches for BSS rarely involve underdetermined BSS cases. The classical independent component analysis (ICA) approach fails to solve underdetermined BSS problems [10]. Moreover in many practical problems, there are a large number of source signals but a few numbers of sensors that means the underdetermined case. Another major difficulty of ICA is that the mixing matrix and the magnitude of original source signals cannot be estimated due to its ambiguities and that the order, sign, and the variances of the independent components cannot be determined [2].

Most of the current traditional BSS methods assume that the source signals are as statistically independent as possible given the observed data and that the mixing matrix is of full column rank. In many real-world situations, however, this hypothesis is not valid. Consequently, recovering the source signals by multiplying the observable data mixtures by the pseudo inverse of the mixing matrix cannot be used. This makes recovering the source signals a very challenging task [8]. In practical terms, the over-determined mixture assumption does not always hold (e.g., in radio communications the probability of receiving more sources than sensors increases with increase of reception bandwidth), thus it is necessary to solve the problem of underdetermined blind source separation (UBSS) [9].

Nonnegative Matrix Factorization (NMF) has been widely applied to BSS problems. However, the separation results are sensitive to the initialization of parameters, also the additive parts by NMF are not necessarily localized, and consequently the solution is not unique. Avoiding the subjectivity of choosing parameters, we use general matrix factorization (GMF), which completely relaxes the non-negativity constraints from its factors with the Alternative Least Squares (ALS) method as an initialization to the

Ossama S. Alshabrawy. Corresponding autor.
Tel. +2-012-211-76423
Email. ossama_alshabrawy87@yahoo.com

source signals instead of random initial values. GMF is a generalization of the well-known NMF where the NMF is constrained by non-negativity on all its factors, is not necessarily localize, has low convergence and, does not provide a unique solution in some cases without additive constraints and parameters. However, GMF has no constraints of non-negativity and is fast convergent with the ALS method used for initialization and improvement.

The motivation of this research is to separate sparse, and super and sub-Gaussian signals in the underdetermined case with an unknown number of source signals without resorting to any sparsity conditions, and to increase the performance of the separation.

The rest of the paper is organized as follows. Section II formulates the problem. In Section III, we introduce an overview, background, and the basic concepts of Projected Gradient and GMF, alternative least square, and rough fuzzy clustering. In Section IV, we present the details of the proposed approach. In section V, we show the analysis of typical experiments and the results obtained by different BSS methods, where the simulation results show the effectiveness and high performance of the proposed algorithm. Finally, a short conclusion and future work are presented in Section 6.

## II. Problem Statement

The problem considered in this paper is an underdetermined instantaneous BSS with an unknown number of source signals but without background noise, which can be mathematically formulated as follows:

Assume that for I unobservable components $X(t) = tr[X_1(t), X_2(t), ...., X_J(t)]$, where J is the number of source signals, and *X(t)* is a zero-mean vector. The available sensor vector $Y(t) = tr[Y_1(t), Y_2(t), ...., Y_I(t)]$, where I is the number of sensors and *tr* is the transpose of the vector, is given by

$$Y(t) = AX(t).$$

Here $A \in R^{I \times J}$ is a non-singular and unobservable matrix and has a non-zero determinant, and the rank of A is I. $X \in R^{J \times T}$, $Y \in R^{I \times T}$. t=0,….,T-1 are the sampling instant time points.

## III. Preliminary Topics

This section provides a brief explanation of the basic technologies used in this paper including projected gradient and GMF, alternative least squares, and rough fuzzy clustering.

### *A. Projected Gradient and General Matrix Factorization*

GMF is a generalization of NMF where there are no nonnegative constraints on all of the factors [12] and is the focus of a great deal of attention in Mathematics and Computer Science. NMF has been widely used in many areas including BSS [11], [13], [14]. However, the solution is not unique since NMF is non-convex programming, and in most algorithms it frequently converges to local optima. Unlike NMF, GMF is convergent and has good local optima avoidance when initialized with ALS. In this paper, GMF is regarded as a good tool for solving the problem of UBSS. The novelty in this paper is that GMF is to solve the UBSS problem for the first time.

The basic GMF decomposition model for BSS is as follows:

$$Y = AX \qquad (1)$$

where, $Y \in R_{\pm}^{I \times T}$ represents the observable mixtures, $A \in R_{\pm}^{I \times J}$ is the mixing matrix, and $X \in R_{\pm}^{J \times T}$ is the source signals matrix. Hence, Y, A, and X have both signs unlike NMF where Y, A, and X are non-negative. For BSS, I is the number of mixtures or sensors, T is the number of sample time points, and J is the number of sources. With only the data observable mixtures (Y) as the only known variable, the mixing matrix A and the source signals X are estimated using Equation (2).

We will use the projected gradient based update rules in GMF. These updates take the following generalized form of iterative rules [11]:

$$X^{(n+1)} = X^{(n)} - \alpha_X P_X \qquad (2)$$

$$A^{(n+1)} = A^{(n)} - \alpha_A P_A \qquad (3)$$

where $P_A$ and $P_X$ are the descent directions, and $\alpha_A$ and $\alpha_x$ are the learning rates, of A and X respectively.

The projected gradient algorithms for GMF are based on the alternating minimization technique which can be written in the matrix form as follows:

$$\min_{x_{ij}} Cost(Y \,||\, AX) = \frac{1}{2}||Y - AX||_F^2 \qquad (4)$$

$$\min_{a_{ij}} Cost(Y^T \,||\, X^T A^T) = \frac{1}{2}||Y^T - X^T A^T||_F^2 \qquad (5)$$

Basically, the matrix A is assumed to be full rank. Consequently, this provides the existence of a unique solution $X^* \in R^{J \times T}$. The gradient matrix for A and X is given by the following equations:

$$Grad_X(X) = \nabla_X Cost(Y \,||\, AX) = A^T(AX - Y) \qquad (6)$$

$$Grad_A(A) = \nabla_A Cost(Y^T \,||\, X^T A^T) = (AX - Y)X^T \qquad (7)$$

One of the projected gradients based approaches, and will be applied in this paper in a modified version, is Lin's projected gradient algorithm [15]. Lin's projected gradient (LPG) algorithm can be induced by the iterative formulas (2) and (3) with $P_A$ and $P_X$ expressed by the equations (6) and (7). Moreover, the projection on the subspace of non-negative real numbers is not considered.

*B. Alternative least squares (ALS)*

The minimization of cost function in equations (4) and (5) which represent the standard squared Euclidean distance can be formulated as follows:

$$\begin{aligned} Cost(Y \parallel AX) &= \frac{1}{2}\|Y - AX\|_F^2 \\ &= \frac{1}{2} tr(Y - AX)(Y - AX) \end{aligned} \tag{8}$$

where *tr* stands for the transpose of the matrix. The above cost function can be alternately minimized with respect to the two factors A and X [11]. Moreover, each time during the optimization process of one factor while keeping the other one fixed [18],[19] and finding the stationary or critical points, which are obtained by equating the gradients to zero. This corresponds to the following two minimization problems:

$$\begin{aligned} &A^{(k+1)} = \min_A \|Y - AX^{(k)}\|_F^2, \text{ and} \\ &X^{(k+1)} = \min_X \|Y^T - X^T[A^{(k+1)}]^T\|_F^2 \end{aligned} \tag{9}$$

The gradients after equating them by zero according to the Karush-Kuhn-Tucker (KKT) optimality conditions are:

$$\frac{\partial D_F(Y \parallel AX)}{\partial a_{ij}} = [-YX^T + AXX^T]_{ij} = 0,$$

$$\frac{\partial D_F(Y \parallel AX)}{\partial x_{ij}} = [-A^T Y + A^T AX]_{ij} \; \forall ij. \tag{10}$$

Consequently,

$$A = YX^T(XX^T)^{-1} \text{ and, } X = (A^T A)^{-1} A^T Y. \tag{11}$$

This method will be used as an initialization in our proposed system.

*C. Rough fuzzy clustering*

In fuzzy c-means (FCM) algorithm developed by Dunn in 1973, improved by Bezdek in 1981, and is the best known method for fuzzy clustering, based on optimizing objective function, the concept of traditional k-means clustering algorithm is extended which for each data point a degree of membership or membership function $\zeta_{ij} \in [0,1]$ of clusters is calculated.

$$\zeta_{ij} = \frac{1}{\sum_{j=1}^{c} \left( \frac{d_{ik}}{d_{jk}} \right)^{2/\delta - 1}} \tag{12}$$

, where $\delta$ is the degree of fuzziness.

In contrast to fuzzy clustering, in rough c-means (RCM), the concept of k-means is extended by considering each cluster as an interval or rough set Y [16]. It is characterized by the lower approximation $\underline{BY}$ and the upper approximations $\overline{BY}$ with the following properties: (i) an object or a sample $y_k$ can be part of at most one lower approximation; (ii) if $y_k \in \underline{BY}$ of cluster X, then simultaneously $y_k \in \overline{BY}$ ; and (iii) if $y_k$ is not a part of any lower approximation, then it belongs to two or more upper approximations. This permits overlaps between clusters.

A rough–fuzzy c-means algorithm which involves the integration of fuzzy and rough sets has been developed [17]. This allows incorporating the fuzzy membership value $\zeta_{ij}$ of a sample $y_k$ to cluster center $\beta_i$. Moreover, instead of absolute individual distance $d_{ik}$ from the centroid, the membership to the cluster center $\beta_i$ is relative to the other centers $\beta_j \; \forall i \neq j$. Consequently, the robustness of the clustering will be enhanced with respect to different choices of the parameters. The centroid $\beta_i$ of cluster $U_i$ can be determined by the following equation:

$$\beta_i = \begin{cases} Z & \text{if } \overline{BU_i} - \underline{BU_i} \neq \varnothing, \; \underline{BU_i} \neq \varnothing, \\ \dfrac{\sum_{y_k \in (\overline{BU_i} - \underline{BU_i})} y_k}{|\overline{BU_i} - \underline{BU_i}|} & \text{if } \overline{BU_i} - \underline{BU_i} \neq \varnothing, \underline{BU_i} = \varnothing, \\ \dfrac{\sum_{y_k \in (\underline{BY})} y_k}{|\underline{BU_i}|} & \text{otherwise}. \end{cases} \tag{13}$$

where,

$$Z = W_{upper} \frac{\sum_{y_k \in (\overline{BU_i} - \underline{BU_i})} y_k}{|\overline{BU_i} - \underline{BU_i}|} + W_{lower} \frac{\sum_{y_k \in (\underline{BU_i})} y_k}{|\underline{BU_i}|}$$

The algorithm of rough fuzzy c-means is stated below in Algorithm 1.

**Algorithm 1** Rough fuzzy c-means clustering

Step1: Assign initial means $\beta_i$ for c clusters.

Step 2: Compute the fuzzy membership $\zeta_{ij}$ for c clusters and N data objects according to equation (12) and Normalize the distances used for fuzzy membership in [0,1].

Step 3: Assign each data object $y_k$ to the lower or upper approximation of cluster pair $U_i$ and $U_j$.

Step 4: Compute the difference $\zeta_{ik} - \zeta_{jk}$ to cluster centroids $\beta_i$ and $\beta_j$

Step 5: Let $\zeta_{ik}$ be maximum and $\zeta_{jk}$ be the next to maximum

Step 6: If $abs(\zeta_{ik} - \zeta_{jk})$ is less than some *threshold*

Then,

$y_k \in \overline{BU_i}$ and $y_k \in \overline{BU_i}$ and $y_k$ cannot be a member of any lower approximation,

Else,

$y_k \in \underline{BU_i}$ such that the membership $\zeta_{ik}$ is the maximum over c clusters
Step 7: Compute the new centroid for each cluster using equation (13)
Step 8: Repeat Steps 2–7 until convergence or until there are no new assignments.

## IV. Proposed UBSS Algorithm

In this section, the proposed approach is presented briefly starting with estimating the mixing matrix knowing only the observable mixtures matrix. Also, a method for GMF gradient-based update rules initialized with ALS is introduced.

### *A. Mixing matrix estimation based on short time Fourier transform and rough fuzzy clustering*

Conventional algorithms estimate the mixing matrix based on clustering algorithms such as the k-means algorithm require that the source signals to be very sparse in the time domain and this is unavailable in many practical real world problems. Other algorithms are based on an assumption that there exist many TF points of single source occupancy (SSO), or require that there exists at least one small region in the TF plain with only a single source and such a TF region must exist for each source. All aforementioned approaches require that for each source there exist many TF points of SSO. However, single source detection (SSD) requires that there exists at least one TF point of SSO and is hence less restrictive than the other approaches [20].

The short time Fourier transform (STFT) of the $i^{\text{th}}$ observed signal is defined by the following equation:

$$Y_i^{Fourier}(t,r) = \sum_{l=0}^{\infty} h(l-t) X_i(l) e^{-jrl} \qquad (14)$$

at frame *t* and frequency bin *r* where *h(l)* is a window sequence. In equation (14), *i=1,2,...,I; t=0,1,...,T-1* are the sampling points over the time domain and *r=0,1,...,T-1* are the sampling points over the frequency domain. The SSD is based on the ratio of the TF transforms and finds a set of TF points where a single source is active for each source. Therefore, for a given $\varepsilon > 0$, a set

$$\chi_F = \left\{ (t,r) \mid \| \mathrm{Im}\left[ \frac{Y^S(t,r)}{Y_1^S(t,r)} \right] \|_F < \varepsilon,\ Y_1^S(t,r) \neq 0 \right\} \qquad (15)$$

where, Im[ ] denotes the imaginary part. We can choose any of the mixture instead of $Y_1$.

During clustering the observable mixtures after incorporating STFT, we need to determine the number of source signals. Since there is an overlap between the data objects, estimating the number of sources require an efficient validity index [21] and can be given by the following equation:

$$V(\beta, c) = Scat(c) + \frac{Sep(c)}{Sep(C_{\max})} \qquad (16)$$

where, $\beta$ is the cluster centers, c is the number of clusters, and $C_{max}$ is the chosen maximum number of clusters. Here,

$$Scat(c) = \frac{\frac{1}{c}\sum_{i=1}^{c} \| \sigma(\beta_i) \|}{\| \sigma(Y) \|} \qquad (17)$$

Also, the value of *Scat(c)* varies from 0 to 1. The term that represents the separation between clusters is defined by

$$sep(c) = \frac{D_{\max}^2}{D_{\min}^2} \sum_{i=1}^{c} \left( \sum_{j=1}^{c} \| \beta_i - \beta_j \|^2 \right)^{-1} \qquad (18)$$

where,

$D_{\min} = \min_{i \neq j} \| \beta_i - \beta_j \|, D_{\max} = \max_{i,j} \| \beta_i - \beta_j \|$

After clustering, and determining the number of source signals, the *i*th column vector of A, denoted as $\hat{a}_i$, is estimated as

$$\hat{a}_i = \frac{1}{|\chi_{C_i}|} \sum_{(t,r) \in \chi_{C_i}} \mathrm{Re}\,[Y^F(t,r)]. \qquad (19)$$

Here $|\chi_{C_i}|$ represents the number of TF points in cluster $C_i$ for *i=1,2,...,J*.

**Algorithm 2** Mixing matrix estimation and determining the number of source signals

**Input:** the observable mixtures *Y= [Y1, Y2,……,YI]*
**Output:** number of source signals, the mixing matrix A

Step1: Calculate STFT Y using equations (14)
Step 2: Calculate $\chi_F$ using equation (15)
Step 3: Cluster $\chi_S$ using rough fuzzy c-means clustering stated in Algorithm 1 for different number of clusters by choosing $C_{min}$, $C_{max}$ (i.e. min and max chosen number of clusters, respectively) using equations (16)-(18) and the cluster number that minimizes *V* is considered to be the optimal value for number of source signals.
Step 4: Determine the TF points and their quantity in each cluster.
Step 5: Calculate the columns of the mixing matrix A using equation (19)

### *B. Lin's Projected Gradient (LPG) with Armijo rule based GMF*

In Lin's projected gradient algorithm the learning rates $\alpha_A$ and $\alpha_x$ are not fixed diagonal matrices in the inner iterations but are scalars. These learning rates are computing by inexact estimation techniques. Lin considered two options to estimate the learning rates. The first option is the Armijo rule along the projective arc of the algorithm proposed by Bertsekas [23]. The value of the learning rate $\alpha_x$, for every iterative step of the algorithm, is given by:

$$\alpha_x^{(k)} = \rho^{m_k}, \qquad (20)$$

where $m^k$ is the first non-negative integer m for which

$$Cost(Y \| AX^{(k+1)}) - Cost(Y \| AX^{(k)}) \leq \upsilon tr\left\{ \nabla_X Cost(Y \| AX^{(k)})^T \left( X^{(k+1)} - X^{(k)} \right) \right\} \qquad (21)$$

with $\rho \in (0,1) and\ \upsilon \in (0,1)$. The value of the learning rate $\alpha_A$ is computed in a similar way.

The second option is the modified Armijo rule. Lin and More [24] noticed that $\alpha_A$ and $\alpha_x$ might be very similar, and they proposed to start from $\alpha_x^{(k-1)}$ and to increase or decrease the learning rate according to condition (20). Here in this paper LPG algorithm with Armijo rule is extended, different from [11], for general matrix factorization relaxing the non-negativity constraints. Moreover the value of $\rho$ and $\upsilon$ are changed to be $\rho \in (-1,1) and\ \upsilon \in (-1,1)$. The algorithm of the modified LPG algorithm with Armijo rule is listed below in Algorithm 3.

**Algorithm 3** Modified LPG based GMF

**Input:** the observable mixtures Y= [Y1, Y2,……..,Yt], number of components (source signals) J, Maximum number of iterations N, the mixing matrix A estimated from algorithm 2

**Output:** the estimated source signals X

Step 1: Initialize the matrix X by ALS according to equation (11)

Step 2: set $\alpha_x$ =1

Step 3: Assign $X^{(n+1)} = X^{(n)} - \alpha_X P_X$

Step 4: Repeat 4-10 until the stopping criteria is met

Step 5: If condition in (21) is met, then

Step 6: Repeat steps 5,6 until condition (21) does not hold

Step 7: Assign $\alpha_x = \frac{\alpha_x}{\rho}$

Step 8: Update $X^{(n+1)} = X^{(n)} - \alpha_X P_X$

Step 9: Else

Repeat steps 8,9 until condition (21) is met

Step 10: Set $\alpha_x = \alpha_x \rho$

Step 11: Update $X^{(n+1)} = X^{(n)} - \alpha_X P_X$

The source signals are then rescaled and normalized

## V. Experiments and Simulation

In this section, the effectiveness of the proposed approach will be discussed by comparing results of experiments and stimulations. Experiments and simulations were performed on synthetically generated signals using the proposed approach and other conventional approaches. In the simulations, sparse, super- and sub-Gaussian signals were separated from the underdetermined mixtures in the challenging case where the true number of source signals is unknown.

The parameter inputs of the modified LPG algorithm are the observable mixtures matrix Y, and the mixing matrix A obtained from algorithm 2. We choose the maximum number of iterations to be only 25 iterations. We investigate the performance of the proposed UBSS approach in the above mentioned cases by comparing its results with the results of approaches in Khor (2006) [22], Kim and Yoo (2009) [20], Xiang and Peng (2010) [8]. Here, the simulation of the separation of sparse and Gaussian signals is provided followed by some discussion. Then, the cases of a variety of sparse, non-sparse, and super- and sub-Gaussian signals are stated.

### *A. Sparse and Gaussian signals*

The separation of J synthetic Gaussian and sparse signals from I=3 mixtures was performed in the time domain for J=4, 5, and 6 source signals. In this simulation, the mixing matrix was estimated using algorithm 2. The proposed approach was compared to the abovementioned algorithms. The simulation settings were as follows. Synthetic sparse signals were generated by generating 5000 Gaussian samples using the *randn* command of Matlab and substituting 80% of the samples chosen randomly by zeros for each source. The results show that the proposed approach can separate I+3 source signals from I mixtures, unlike the previous approaches. This is confirmed in the next simulations. The analysis aims at comparing mainly the reconstruction index Signal-to-Interference Ratio (SIR) to evaluate the performance of the proposed approach. Given original source signals X and its estimations $\hat{X}$ obtained by the proposed approach, SIR in decibels is defined as

$$SIR = -10\log\left(\frac{\| \hat{X}_i - X_i \|_F^2}{\| X_i \|_F^2}\right), \quad i = 1,2,....,J \qquad (15)$$

Fig. 1 illustrates the averaged SIRs when the number of the sources increases from 4 to 6 signals.

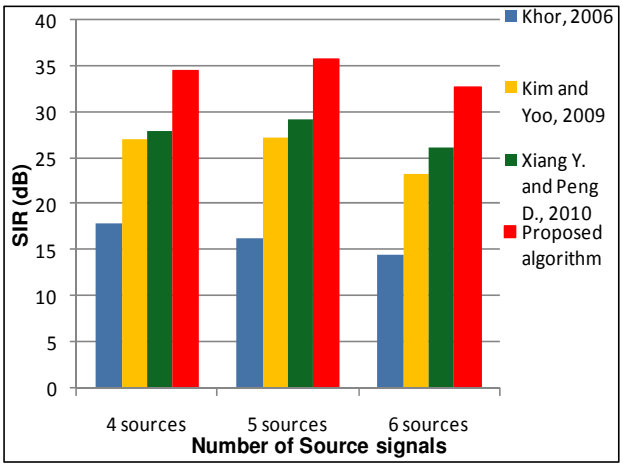



Fig. 1. performance estimation of the source signals according to the number of sources from 3 observable mixtures (in case of Sparse and Gaussian sources).

*B. Synthetic signals*

We investigated the effectiveness of the proposed UBSS approach by comparing with the methods mentioned above. We chose the number of mixtures to be only 2 and the number of sources to be 5 in order to prove that the proposed algorithm can separate I+3 source signals from I mixtures. The mixture signals that we perform our experiments on are mixed by the following randomly generated mixing matrix:

$$\begin{bmatrix} 0.5377 & -2.2588 & 0.3188 & -0.4336 & 3.5784 \\ 1.8339 & 0.8622 & -1.3077 & 0.3426 & 2.7694 \end{bmatrix}$$

The mixing matrix was once again estimated using algorithm 2. The true and estimated values of A are shown in Table 1. The six source signals, two observable mixtures, and estimated source signals are plotted in Fig. 2. The number of sampling time points is 10,000. The simulation results of the proposed approach in addition to those of the five different UBSS methods are shown in Fig.4. The performance of the source recovery method can be evaluated by Eqs. (15) and (16).

$$SNR = \frac{1}{J}\sum_i 10\log\left(\frac{\| X_i \|_F^2}{\| \hat{X}_i - X_i \|_F^2}\right) \qquad (16)$$

Where J is the number of sources and $\|.\|_F^2$ is the Frobenius norm. The efficiency of the separation results is good when $SNR \geq 25$ [10].

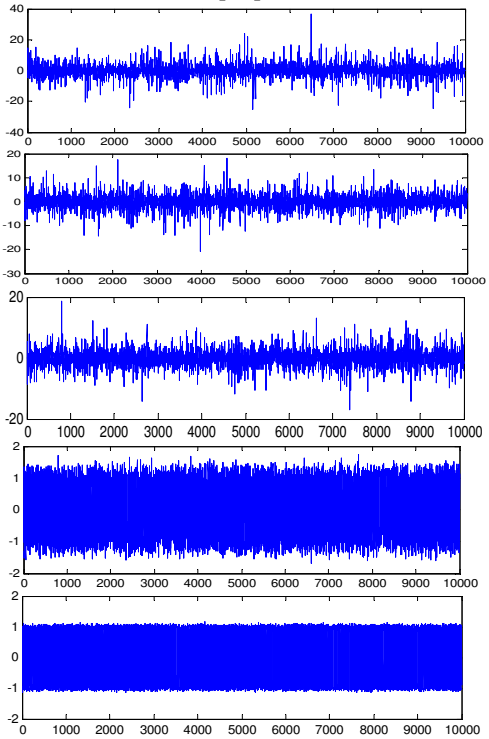
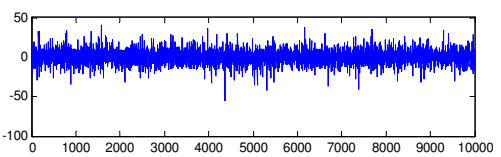


Fig. 2. (a) Source signals



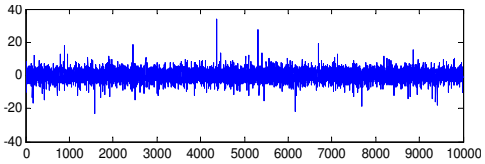


Fig. 2. (b) Observable mixtures

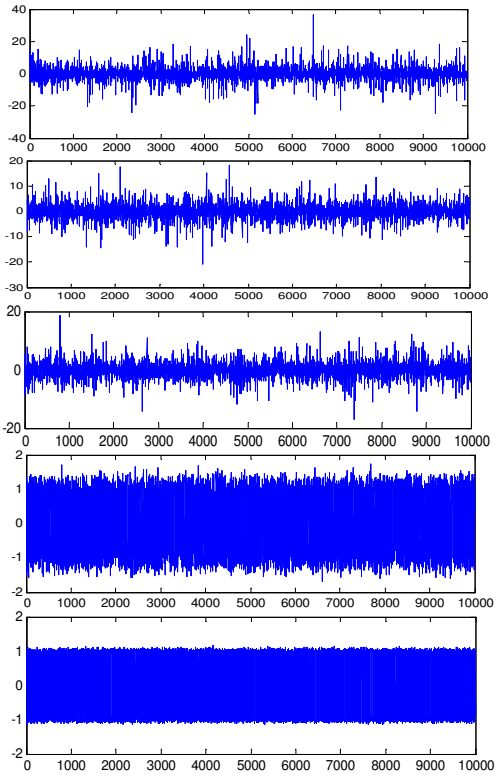


Fig. 2. (c) Estimated source signals

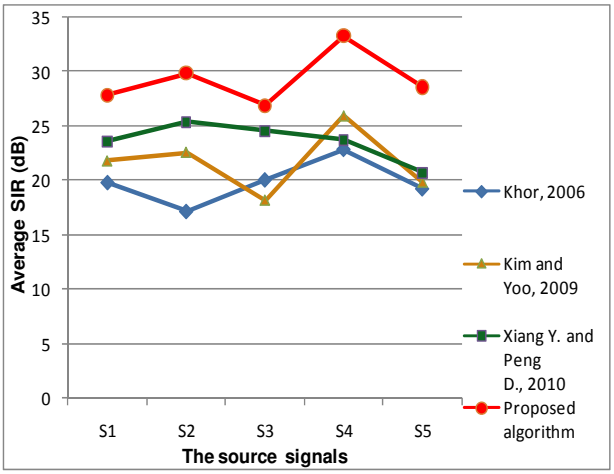



Fig. 3. Performance estimation of the source signals from 2 observable mixtures.

We note from Fig. 3 that the proposed approach achieves about 5.7 dB higher SIR for J=5 sources with only two mixtures than the highest performance algorithm among the other three approaches. Another comparison of the proposed approach with the other five approaches is presented using

the following examples with 2 observable mixtures where $X_i$ is the chosen source signal shown in Fig. 4.

**Example 1.** $X=\{X_1, X_2, X_5\}$

**Example2**. $X=\{X_1, X_2, X_3, X_5\}$

**Example 3**. $X=\{X_1, X_2, X_3, X_4, X_5\}$

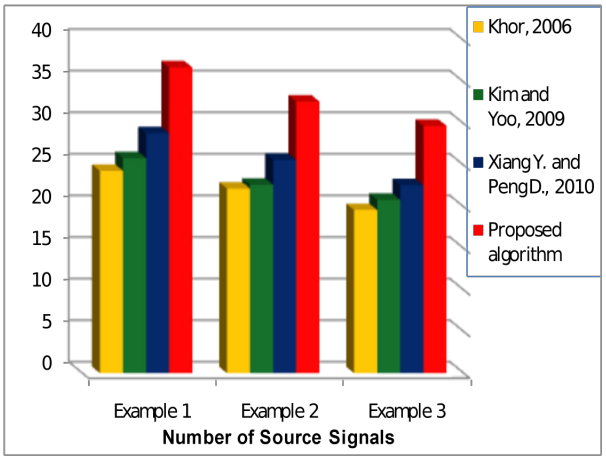



Fig. 4. Performance estimation of Examples 1–3 from 2 observable mixtures.

From the results in Figs. 3and 4, we can conclude that the separation performance of the proposed approach is very high, has faster convergence, and can separate I+3 source signals from I mixed signals.

## II. Conclusion

In this paper, we addressed the problem of underdetermined blind source separation with the challenging case that the true number of source signals is unknown. A new two-step approach for optimum estimation of the source signals without additive noise. In this approach, STFT is combined with rough fuzzy c-means clustering to estimate the mixing matrix and determine the number of source signals. Then the source signals are estimated by a modified LPG algorithm with Armijo rule based general matrix factorization. Simulation experiments demonstrated the validity and superior performance of the proposed approach.

## References

[1] Ossama S. Alshabrawy, M. E. Ghoniem, W. A. Awad and Aboul ella hassanien, Underdetermined Blind Source Separation based on Fuzzy C-Means and Semi-Nonnegative Matrix Factorization. *IEEE FEDERATED CONFERENCE ON COMPUTER SCIENCE AND INFORMATION SYSTEMS*, Wroclaw, Poland, 9-12 September, 2012, pp. 695-700

[2] A. Hyvarinen, J. Karhunen, E. Oja, "Independent Component Analysis, Wiley," New York, 2001.

[3] Yadong, Liu, Zongtan, Zhou, Dewen, Hu. "A novel method for spatio temporal pattern analysis of brain fMRI data," Science in China Series F: Information Sciences, vol. 48, no. 2, pp. 151–160, 2005.

[4] Araki, S., Makino, S., Blin, A., "Underdetermined blind separation for speech in real environment with sparseness and ICA," In Proceedings of the ICASSP'04, Montreal, Canada, pp.881–884, 2004.

[5] Ohnishi, Naoya, Imiya, Atsushi, "Independent component analysis of optical flow for robot navigation," Neurocomputing, vol.7, nos. 10–12, pp. 2140–2163, 2008.

[6] Tonazzini, Anna, Bedini, Luigi, Salerno, Emanuele, "A Markov model for blind image separation by a mean-field EM algorithm," IEEE Transactions on Image Processing, vol. 15, no.2, pp.473–482, 2005.

[7] Er-Wei, Bai, QingYu, Li, Zhiyong ,Zhang, "Blind source separation channel equalization of nonlinear channels with binary inputs," IEEE Transactions on Signal Processing, vol. 53, no. 7, pp.2315–2323, 2005.

[8] Dezhong Peng, Yong Xiang, "Underdetermined blind separation of nonsparse sources using spatial time-frequency distributions," Digital Signal Processing, vol. 20, pp. 581–596, 2010.

[9] Fengbo Lu, Zhitao Huang,Wenli Jiang, "Underdetermined blind separation of non-disjoint signals in time–frequency domain based on matrix diagonalization," Signal Processing, vol. 91, pp. 1568–1577, January 2011.

[10] Chaozhu Zhang, Cui Zheng, "Underdetermined Blind Source Separation Based on Fuzzy C-Means Clustering and Sparse Representation," International Conference on Graphic and Image Processing (ICGIP), Proc. of SPIE vol. 8285, 2011.

[11] Andrzej Cichocki, Rafal Zdunek, Anh Huy Phan, Shun-ichi Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, John Wiley, 2009

[12] Vladimir Nikulin, Tian-Hsiang Huangb, Shu-Kay Ngc, Suren I. Rathnayake, Geoffrey J. McLachlan, "A very fast algorithm for matrix factorization," Statistics and Probability Letters, vol. 81, pp. 773–782, February 2011.

[13] Tuomas Virtanen, "Monaural Sound Source Separation by Nonnegative Matrix Factorization with Temporal Continuity and Sparse Criteria," IEEE Transactions on Audio, Speech, and Language Processing, vol. 15, no. 3, March 2007.

[14] A. Cichocki, R. Zdunek, S. Amari, "New algorithms for non-negative matrix factorization in applications to blind source separation, " In Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP2006, vol. 5, pages 621–624, Toulouse, France, May 14–19, 2006.

[15] Ch-J. Lin. Projected gradient methods for non-negative matrix factorization. Neural Computation, 19(10), pp. 2756–2779, October 2007.

[16] Sushmita Mitra, Witold Pedryczb, Bishal Barman, "Shadowed c-means: Integrating fuzzy and rough clustering," Pattern Recognition, vol. 43, pp. 1282–1291, 2010.

[17] S. Mitra, H. Banka, W. Pedrycz, "Rough–fuzzy collaborative clustering, IEEE Transactions on Systems," Man, and Cybernetics—Part B36, pp. 795-805, 2006.

[18] A. Cichocki, S. Amari. *Adaptive Blind Signal and Image Processing*, John Wiley & Sons Ltd, New York, 2003.

[19] D.D. Lee, H.S. Seung. "Algorithms for Nonnegative Matrix Factorization," vol. 13, MIT Press, 2001.

[20] SangGyun Kim, Chang D. Yoo, "Underdetermined Blind Source Separation Based on Subspace Representation," IEEE Transaction on Signal Processing, vol. 57, no. 7, July 2009.

[21] Sun, Haojun, Wang, Shengrui, Jiang, Qiangshan. FCM-based model selection algorithms for determining the number of clusters. Pattern Recognition 37 (10), pp. 2027–2037, 2004.

[22] Khor, L.C. "Robust adaptive blind signal estimation algorithm for under- determined mixture," IEE Proceedings—Circuits, Devices and Systems, vol. 153, no. 4, pp. 320–331, 2006.

[23] D. P. Bertsekas. Nonlinear Programming. Athena Scientific, Belmont, MA, 1999.

[24] Ch. J. Lin and J.J. Mor´e. Newton's method for large bound-constrained optimization problems. SIAM Journal on Optimization, 9(4), pp. 1100–1127, 1999.