

Blind Source Separation of Underdetermined Mixtures Based on Local Mean Decomposition and Conjugate Gradient Algorithm

Wei Li and Huizhong Yang*

Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education) of Jiangnan University, Wuxi 214122, Jiangsu, PR China

Email: liweilwlcw@sina.com; yhz@jiangnan.edu.cn

Abstract—Most of the existing underdetermined blind source separation (BSS) approaches assume that the source signals are strictly or partially sparse. This paper, however, presents a new BSS method in underdetermined mixing situation for non-sparse signals. The proposed method first introduces the local mean decomposition algorithm into the BSS problem to rebuild some extra mixing signals. Such signals are then combined with the initial mixtures such that the underdetermined BSS problem is transformed to a determined one and the difficulty of the deficiency of the mixtures is overcome. For the rebuilt mixtures and the newly formed determined BSS problem, the minimum mutual information principle is employed as the BSS cost function. A conjugate gradient learning algorithm is then derived for training the separating matrix. In each update step of the algorithm, the term of score function is estimated by a kernel function estimation algorithm. The simulation results have demonstrated the efficacy of the proposed underdetermined BSS method.

Index Terms—Blind source separation, underdetermined model, local mean decomposition, conjugate gradient

I. INTRODUCTION

The objective of blind source separation (BSS) is to recover the latent source signals from their mixtures without knowing a prior knowledge of mixing system. BSS has recently received wide attention in literatures, because of its appealing applications in signal denoising, audio and image processing, feature extraction, electromagnetic and biomagnetic analysis problem and so on [1]-[3].

We consider the following linear noisy BSS model,

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} \quad (1)$$

in which, the mixing matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is unknown, $\mathbf{N} \in \mathbb{R}^{m \times T}$ is a white Gaussian noise matrix, $\mathbf{S} \in \mathbb{R}^{n \times T}$ contains n unknown statistical mutual independent source signals and $\mathbf{X} \in \mathbb{R}^{m \times T}$ is the observation matrix containing m mixtures of the sources, where T is the

total number of time index. In this paper, we cope with the underdetermined BSS problem in which the number of sources is greater than the number of observations, i.e. $m < n$. The observation matrix \mathbf{X} in mixing model (1) can also be expressed as vectors $\mathbf{x}(t)$ for a set of time index: $t = 1, 2, \dots, T$, as well as the sources \mathbf{S} and the noise \mathbf{N} . Then (1) can be reformulated in a vector form as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad t = 1, 2, \dots, T \quad (2)$$

where $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$ is the vector of sources at index t , $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ and $\mathbf{n}(t) = [n_1(t), n_2(t), \dots, n_m(t)]^T$ is the corresponding observation and noise vector.

Most of the existing underdetermined BSS approaches assume that the source signals are sparse, i.e., the source signals have one nonzero element at most in each time of the available samples. In this case, the blind separation task can be addressed in two stages: first, the mixing matrix \mathbf{A} is estimated using higher-order statistics based methods [4], [5], time-frequency distributions [6], [7], wavelet packet transform [8], overcomplete representation [9], and some clustering algorithms such as K-Means clustering [10], [11], Median-based clustering [12] and discriminative clustering [13] etc.; then the source signals $s(t)$ is recovered in the light of the estimate of \mathbf{A} in the second stage. Consequently, if the mixing matrix is not well estimated, it is impossible to restore the sources. Georgiev *et al.* [14] presented another sparse component analysis method and pointed out that the underdetermined BSS problem is solvable if the number of nonzero elements in each time of the sources is smaller than the number of observation signals. However, if the strictly or partially sparsity assumption of the source signals is not satisfied, the above approaches will fail in the end.

In this paper, we propose a novel underdetermined blind source separation method. The method uses the local mean decomposition (LMD) algorithm to reconstruct several new observation signals; thereby the underdetermined BSS problem is transformed to a

Manuscript received December 18, 2013; revised April 15, 2014.

This work was supported by the National Natural Science Foundation of China under Grant No. 61273070, Doctor Candidate Foundation of Jiangnan University under Grant No. 1252050205135130 and a Project Funded by the Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions.

*Corresponding author email: yhz@jiangnan.edu.cn.
doi:10.12720/jcm.9.5.425-432

determined BSS problem which is much easier to cope with. Then the minimum mutual information criterion is employed for the blind separation task, and conjugate gradient is used for deriving the training equations of the separating matrix. In each update of the learning algorithm, the score function is estimated by a kernel density estimation method directly. Theoretically, since the local mean decomposition algorithm can be applied to various types of signals, the proposed method does not resort to the sparsity constraint which is included in most existing underdetermined BSS methods. The simulation results have confirmed the efficacy of the proposed underdetermined BSS method.

The rest of this paper is organized as follows: The local mean decomposition algorithm is introduced for generating product functions of mixture signals in section 2. In section 2, the conjugate gradient optimization algorithm is used to solve the newly formed determined BSS problem. Section 3 presents some simulations to demonstrate the performance of the proposed algorithm. Some conclusions are drawn in Section 4.

II. LOCAL MEAN DECOMPOSITION ALGORITHM

Local mean decomposition is a robust and conceptually simple iterative method which is developed for analyzing complicated signals in terms of time-varying frequency, phase and energy [15]-[17]. LMD decomposes a rough signal into a set of product functions, each of which is a product of an envelope signal and a frequency modulated signal. Therefore, LMD can be used to analyze a wide variety of natural signals, but is of particular relevance with regard to the analysis of amplitude and frequency modulated signals, such as electrocardiograms, functional magnetic resonance imaging data, and earthquake data, etc.

In practice, LMD is performed through separating a frequency modulated signal from an amplitude modulated envelope signal progressively. Briefly, this separation scheme is realized by firstly smoothing the original rough signal, secondly subtracting the smoothed signal from the original signal and finally demodulating the result using an envelope estimate in amplitude. With the observation signal $x(t)$, the LMD algorithm can be described as follows:

1) Detect all local extrema $\{n_i\}$ of the observation signal, where $i=1,2,\dots,I$ is the index of the successive extrema and I is the total number of the extrema.

2) Compute the local mean function m_{jk} , where j represents the index of the decomposed product function component, and k is the iteration number in each decomposition process.

- Calculate the local averages of the successive maximum and minimum points of x . So the i th mean value m_i of every two adjacent extrema n_i and n_{i+1} is given by $m_i = (n_i + n_{i+1})/2$.

- The local averages are then smoothed using the moving averaging approach to form a smoothly varying continuous local mean function m_{jk} .

3) Compute the corresponding local envelope function a_{jk} .

- Calculate the local envelopes of the successive maximum and minimum points of x . Then the i th local envelope magnitude a_i of each half-wave oscillation can be calculated as $a_i = |n_i - n_{i+1}|/2$.

- The local envelope magnitudes are then smoothed in the same way as the local means to form a smoothly varying continuous local envelope function a_{jk} .

4) Compute the corresponding frequency modulated signal v_{jk} .

- The local mean signal m_{jk} is subtracted from the original data x , and the resulting signal is denoted by h_{jk} .

- Divide h_{jk} by a_{jk} , to produce a frequency modulated signal v_{jk} .

5) Compute the envelope function $a_{j(k+1)}$ of v_{jk} as the same as the step 3 to decide the direction of the procedure.

- If $|a_{j(k+1)} - 1| \approx 0$, i.e. the signal v_{jk} is a purely frequency modulated signal which has a flat envelope, then go to step 6.

- If $|a_{j(k+1)} - 1| \gg 0$, then replace x by v_{jk} , let k plus 1 and go to the first step.

- The above procedures will be repeated for K times until $|a_{j(K+1)} - 1| \approx 0$ holds. Therefore, a set of equations can be obtained as

$$\begin{cases} h_{j1} = x - m_{j1}, & v_{j1} = h_{j1}/a_{j1} \\ h_{j2} = v_{j1} - m_{j2}, & v_{j2} = h_{j2}/a_{j2} \\ \vdots & \vdots \\ h_{jK} = v_{j(K-1)} - m_{jK}, & v_{jK} = h_{jK}/a_{jK} \end{cases} \quad (3)$$

6) Calculate the final envelope signal a_j of the corresponding product function component.

- The envelope signal a_j can be derived by multiplying together all the successive envelope estimate functions a_{jk} , $k=1,2,\dots,K$, which are acquired during the above iterative process, i.e.,

$$a_j = a_{j1} * a_{j2} * \dots * a_{jK} = \prod_{k=1}^K a_{jk} \quad (4)$$

7) Compute the j th product function component P_j by multiplying the frequency modulated signal v_{jK} by the final envelop signal a_j : $P_j = v_{jK} * a_j$.

8) The product function component is subtracted from the original observation signal: $u_j = x - P_j$.

(9) Judge whether the decomposition procedure has been accomplished.

- If the new obtained signal u_j is a constant or contains no more oscillations, then replace x by u_j as a new observation signal, let j plus 1 and go to the first step for deriving the next product function component. The procedure will be repeated for J times until u_j becomes a monotonic function. Then we can obtain

$$\begin{cases} u_1 = x - P_1 \\ u_2 = u_1 - P_2 \\ \vdots \\ u_J = u_{J-1} - P_J \end{cases} \quad (5)$$

From Equ. (5), it can be seen that the original observation signal x can be expressed as a linear combination of the product function components $\{P_j\}$, $j=1,2,\dots,J$ and the remaining difference u_j . The product functions will be used to solve a blind source separation problem in the next section.

III. UNDERDETERMINED BLIND SOURCE SEPARATION WITH CONJUGATE GRADIENT ALGORITHM

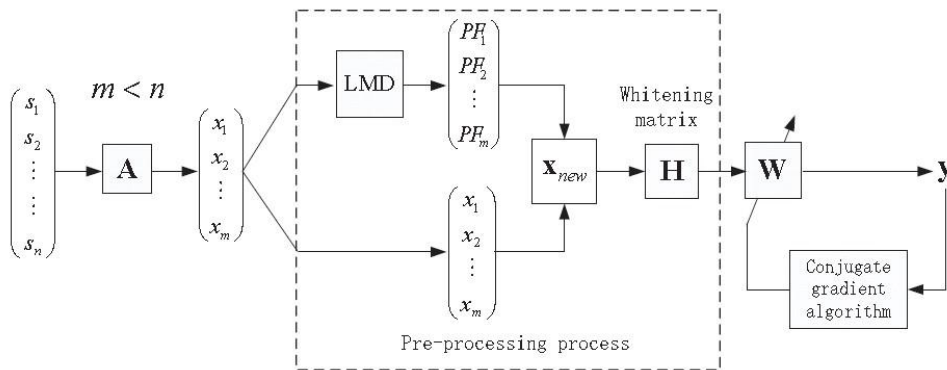


Fig. 1. The schematic diagram of the proposed underdetermined BSS algorithm

A. Generate New Mixture Signals

Before presenting the underdetermined BSS algorithm, we explain how to construct additional mixed observation signals. The following proposition demonstrates why the product functions from LMD can be used as the extra observations.

Proposition 1. Suppose a vector \mathbf{x} is a linear combination of a set of vectors $\{\mathbf{y}_i\}$, $i=1,2,\dots,n$, with the coefficient vector $\mathbf{a}=[a_1,a_2,\dots,a_n]$, and is also a linear combination of an another group of vectors $\{\mathbf{z}_j\}$, $j=1,2,\dots,m$, with the coefficient vector $\mathbf{b}=[b_1,b_2,\dots,b_m]$. Then each component of $\{\mathbf{y}_i\}$ ($\{\mathbf{z}_j\}$) can be represented as a linear combination of $\{\mathbf{z}_j\}$ ($\{\mathbf{y}_i\}$), if the matrix of the Kronecker product between the coefficient vector \mathbf{a} (\mathbf{b}) and its transpose is non-singular.

As is well known, blind source separation is an ill-conditioned problem under the underdetermined condition. Since there are fewer observation signals than sources, the inverse or pseudo-inverse matrix of the underdetermined mixing matrix can not be computed. In this case, to recover the source signals is a difficult and intractable task.

In this section, an underdetermined blind source separation approach is proposed, which aims at solving the disadvantage of lack of observations of the underdetermined model. Some product functions are generated by performing the local mean decomposition algorithm on the observation signals. Then a new group of observations is formed by lumping the generated signals and the original observations together. So the underdetermined BSS problem can be transformed to a much easier determined one. After the observation reconstruction, the minimum mutual information principle is used as the criterion for the BSS, and a conjugate gradient based BSS algorithm with score function estimation is derived for recovering the independent source signals. The schematic diagram of the proposed algorithm is shown in Fig. 1.

Proof. Without loss of generality, suppose that

$$\mathbf{x} = a_1\mathbf{y}_1 + a_2\mathbf{y}_2 + \dots + a_n\mathbf{y}_n = \mathbf{a}[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n]^T \quad (6)$$

and also

$$\mathbf{x} = b_1\mathbf{z}_1 + b_2\mathbf{z}_2 + \dots + b_m\mathbf{z}_m = \mathbf{b}[\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m]^T \quad (7)$$

So we have

$$\mathbf{a}[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n]^T = \mathbf{b}[\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m]^T \quad (8)$$

Denote $\mathbf{A} \doteq \mathbf{a}^T \otimes \mathbf{a}$ and $\mathbf{B} \doteq \mathbf{a}^T \otimes \mathbf{b}$, where \otimes is the Kronecker product operator.

To proceed, left multiply the column vector \mathbf{a}^T to the both sides of Equ. (8), we obtain that

$$\mathbf{A}[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n]^T = \mathbf{B}[\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m]^T \quad (9)$$

Therefore, if \mathbf{A} is non-singular, it then holds that

$$[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n]^T = \mathbf{A}^{-1}\mathbf{B}[\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m]^T \quad (10)$$

Obviously, $\mathbf{A}^{-1}\mathbf{B}$ is a $n \times m$ linear transform matrix. In other words, each component within the set $\{\mathbf{y}_i\}$ can be represented as a linear combination of the set $\{\mathbf{z}_j\}$, vice versa.

From section 2, It is known that each observation signal x_i , $i=1,2,\dots,m$ can expressed as a linear combination of the source signals $\{s_k\}$, $k=1,2,\dots,n$. In the meanwhile, x_i can also be represented as a linear combination of a serial product functions $\{P_j\}$, $j=1,2,\dots,J$ regardless of the remaining difference, exploiting the LMD method introduced in the previous section. Therefore, according to Proposition 1, each product function component P_j can be represented as a linear combination of the source signals. Then a new set of observation signals can be constructed by the original observations together with the obtained product functions. An underdetermined BSS problem is transformed to a much easier (over-)determined one, which will be solved by the method proposed in the next subsection.

B. The Conjugate Gradient Based BSS Algorithm

The proposed underdetermined BSS algorithm mainly contains two stages: the pre-processing of the original underdetermined mixtures and the blind separation of the newly built determined model.

Pre-processing of the underdetermined mixtures: In the pre-processing stage, we need to extend the dimension number of the observation signals and to transform the underdetermined mixing model to a determined one.

First, the local mean decomposition algorithm is performed on each component of the original underdetermined mixtures \mathbf{x} . Thus we can obtain a number of product functions P_{pq} , where p denotes the index of the original observation signal and q is the numerical order of the product function component.

Secondly, we should choose several proper product functions, which inherit vast majority information of the original observations, as the complementary observations. To achieve this goal, the cross-correlation coefficient r_{pq} between the product function P_{pq} and the corresponding observation signal x_p is calculated as

$$r_{pq} = \frac{T \sum_{t=1}^T P_{pq}(t)x_p(t) - \sum_{t=1}^T P_{pq}(t) \sum_{t=1}^T x_p(t)}{\sqrt{T \sum_{t=1}^T P_{pq}^2(t) - \left[\sum_{t=1}^T P_{pq}(t) \right]^2} \sqrt{T \sum_{t=1}^T x_p^2(t) - \left[\sum_{t=1}^T x_p(t) \right]^2}} \quad (11)$$

where T is the total number of sample points. After r_{pq} are calculated out, we reserve the product function components with large numerical value of r_{pq} . These components, denoted by P'_{pq} , are considered to have

more information related to x_p , while the rest components with small numerical value of r_{pq} are removed.

Thirdly, we lump the product functions P'_{pq} together with the original observation \mathbf{x} to construct a new set of observations \mathbf{x}_{new} . It should be noted that the number of the selected product functions should be adequate to ensure that the new set of observations \mathbf{x}_{new} have not fewer components than the source signals \mathbf{s} . Thus the underdetermined BSS problem is transformed to an (over-)determined one.

Finally, since the number of the new observations is often greater than that of the sources, a pre-whitening and dimension reduction process is needed. As soon as the samples of \mathbf{x}_{new} are obtained, the sample correlation matrix is computed as $\hat{\mathbf{R}} = E\{\mathbf{x}_{new}\mathbf{x}_{new}^T\}$. An eigenvalue decomposition is then performed on $\hat{\mathbf{R}}$:

$$\hat{\mathbf{R}} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T = \mathbf{V}_s\mathbf{\Lambda}_s\mathbf{V}_s^T + \mathbf{V}_n\mathbf{\Lambda}_n\mathbf{V}_n^T \quad (12)$$

where the subscript \mathbf{s} and \mathbf{n} indicate the source and the noise separately, \mathbf{V} and $\mathbf{\Lambda}$ present the corresponding eigenvector matrix and the eigenvalue matrix respectively. The dimension of $\mathbf{\Lambda}_s$ is selected in accordance with the number of the sources. Thus the whitened observations are obtained by

$$\bar{\mathbf{x}}_{new} = \mathbf{\Lambda}_s^{-\frac{1}{2}}\mathbf{V}_s^T\mathbf{x}_{new} \quad (13)$$

Blind separation with conjugate gradient algorithm: When the pre-processing process is finished, the underdetermined BSS model is transformed to a determined one. For such model, a minimum mutual information based conjugate gradient BSS algorithm is developed to accomplish the blind separation task.

The basic idea of minimum mutual information is to minimize the statistical dependence among the components of the output signals $\mathbf{y}(t)$ of the separating system [18,19]. Mutual information is usually approximated by the Kullback-Leibler divergence between the joint probability density function (PDF) and the product of the marginal PDF of a random vector. In view of the relation between Kullback-Leibler divergence and differential entropy, the minimum mutual information cost function $I(\mathbf{y}, \mathbf{W})$ can be formulated as

$$I(\mathbf{y}, \mathbf{W}) = \sum_{i=1}^n H(y_i) - H(\mathbf{y}) = \sum_{i=1}^n H(y_i) - \ln|\det(\mathbf{W})| \quad (14)$$

where $H(\bullet)$ denotes differential entropy and $\det(\bullet)$ is the determinant of a non-singular matrix. The nature gradient of $I(\mathbf{y}, \mathbf{W})$ with respect to the separating matrix \mathbf{W} is given by

$$\nabla_{\mathbf{W}}I(\mathbf{y}, \mathbf{W}) = \left(E\{\Phi(\mathbf{y})\mathbf{y}^T\} - \mathbf{I} \right) \mathbf{W} \quad (15)$$

where $\Phi(\mathbf{y}) = [\phi_1(y_1), \phi_2(y_2), \dots, \phi_n(y_n)]^T$, whose components $\phi_i(y_i) = -p'(y_i)/p(y_i)$, $i = 1, 2, \dots, n$ are defined as the score functions of the output signals.

The following definition is helpful to develop the conjugate gradient BSS algorithm.

Definition 1. Suppose that there exists a space of $n \times n$ dimensional matrices. When two arbitrary points of the space, i.e., two matrices \mathbf{A} and \mathbf{B} , are not very far from each other, the shortest trajectory from \mathbf{A} to \mathbf{B} is defined as ‘geodesic’ denoted by $\mathbf{G}(\xi)$, which can be formulated by

$$\mathbf{G}(\xi) = \exp(\xi \mathbf{T}_A \mathbf{A}^{-1}) \mathbf{A} \quad (16)$$

where the parameter ξ is bounded in the interval $[0,1]$ such that $\mathbf{G}(0) = \mathbf{A}$ and $\mathbf{G}(1) = \mathbf{B}$, and $\mathbf{T}_A = \mathbf{G}'(0)$ is a tangent vector at the point \mathbf{A} , which indicates the direction of the geodesic.

The conjugate gradient algorithm contains two key procedures in each update [20], [21]: 1) Calculate the tangent vector of the current solution point, which is conjugate to the former searching direction, thus the next searching direction is determined; 2) Solve a one-dimensional optimization problem to find the new iterative solution along the newly formed trajectory of geodesic.

Denote by \mathbf{W}_k the k th searching result, by $\mathbf{T}_{\mathbf{W}_k}$ the current searching direction at the point \mathbf{W}_k . Since the separating matrix space is a Riemannian space, the tangent vector $\mathbf{T}_{\mathbf{W}_k}$ can be calculated by

$$\mathbf{T}_{\mathbf{W}_k} = \mathbf{T}_{\mathbf{W}_{k-1}} \mathbf{W}_{k-1}^{-1} \mathbf{W}_k \quad (17)$$

In order to find the new searching direction that is conjugate to the former direction, the nature gradient $\nabla I(\mathbf{y}, \mathbf{W}_k)$ at the point \mathbf{W}_k is computed as in Equ. (15). The conjugate gradient direction is thus given by the tangent vector

$$\mathbf{T}_{\mathbf{W}_{k+1}} = \nabla I(\mathbf{y}, \mathbf{W}_k) + \lambda_{k+1} \mathbf{T}_{\mathbf{W}_k} \quad (18)$$

where the parameter λ_{k+1} is selected to ensure $\mathbf{T}_{\mathbf{W}_{k+1}}$ is the conjugate direction. In practice, λ_{k+1} can be calculated by the finite difference approximation as

$$\lambda_{k+1} = \frac{\text{tr}\left\{\left(\nabla I(\mathbf{y}, \mathbf{W}_k) - \nabla I(\mathbf{y}, \mathbf{W}_{k-1})\right) \nabla I(\mathbf{y}, \mathbf{W}_k)^T\right\}}{\text{tr}\left\{\nabla I(\mathbf{y}, \mathbf{W}_{k-1}) \nabla I(\mathbf{y}, \mathbf{W}_{k-1})^T\right\}} \quad (19)$$

After the new searching direction $\mathbf{T}_{\mathbf{W}_{k+1}}$ is determined, the next iterative point \mathbf{W}_{k+1} can be obtained by solving the following one-dimensional linear optimization problem along the geodesic

$$\mathbf{W}_{k+1} = \arg \min_{\xi} I(\mathbf{y}, \mathbf{G}_{\mathbf{W}_k}(\xi)) \quad (20)$$

where the geodesic is given by $\mathbf{G}_{\mathbf{W}_k}(\xi) = \exp(\xi \mathbf{T}_{\mathbf{W}_k} \mathbf{W}_k^{-1}) \mathbf{W}_k$, as defined before.

The score functions $\Phi(\mathbf{y})$ in each update of the proposed algorithm are estimated by a kernel density estimation method. It is assumed that there are T realizations $\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(T)$ of the separation signals. Therefore, the kernel density estimator [22], [23] $\hat{p}_{i,h}(y_i)$, exploited to estimate the true marginal PDF $p_i(y_i)$, is given by

$$\hat{p}_{i,h}(y_i) = \frac{1}{T} \sum_{t=1}^T K_h(y_i - y_i(t)) \quad (21)$$

where $K_h(u) = \frac{1}{h} K\left(\frac{u}{h}\right)$, $K(\bullet)$ is a kernel function, and

h is the bandwidth. In this paper, The kernel function is chosen to be Gaussian kernel function

$$K_G(v) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right).$$

To fix the bandwidth h , the asymptotic mean integrated squared error $M(h)$, given by

$$M(h) = \frac{\sigma_K^4 h^4 R(p_i''(y_i))}{4} + \frac{R(K(y_i))}{Th} \quad (22)$$

is employed to measure the gap between the true density $p_i(y_i)$ and the estimator $\hat{p}_{i,h}(y_i)$, in which, $R(K(y_i)) = \int K^2(y_i) dy_i$ and $\sigma_K^2 = \int y_i^2 K(y_i) dy_i$. The optimal bandwidth h_{opt} is obtained as

$$\hat{h}_{opt} = T^{-\frac{1}{5}} \left\{ \frac{R(K)}{\sigma_K^4 R(\hat{p}_i''(g_i(\hat{h}_{opt})))} \right\}^{\frac{1}{5}} \quad (23)$$

by minimizing the measure $M(h)$, where

$g_i(h_{opt}) = C(K(y_i)) D(p_i(y_i)) h_{opt}^{\frac{5}{7}}$ for some appropriate $C(K(y_i))$ and $D(p_i(y_i))$. $C(K(y_i))$ can be just simply replaced by an proper constant. Usually,

$C(K(y_i))$ is selected as $(6\sqrt{2})^{\frac{1}{7}}$. $D(p_i(y_i))$ is a function of $R(p_i''(y_i))/R(p_i^{(3)}(y_i))$ and can be estimated by the plug-in method [23]. Substituting the bandwidth h in (21) with the solution value of \hat{h}_{opt} , the kernel density estimator $\hat{p}_{i,h_{opt}}(y_i)$ is obtained. Therefore,

the score function can be estimated as $\hat{\phi}_i(y_i) = -\hat{p}'_{i,h_{opt}}(y_i)/\hat{p}_{i,h_{opt}}(y_i)$, where $\hat{p}'_{i,h_{opt}}(y_i)$ is the first-order derivative of $\hat{p}_{i,h_{opt}}(y_i)$.

IV. SIMULATIONS

In order to demonstrate the validity of the proposed underdetermined blind source separation method, several simulations are performed in this section. The simulations considered the underdetermined BSS problem in the case of four speech sources and three mixtures. The 3×4 dimensional mixing matrix \mathbf{A} is selected randomly, whose elements subject to uniform distribution in the interval $[0,1]$. The observation signals are obtained according to (2) with 1% Gaussian noise. The sources and the noisy mixtures are shown in Fig. 2 and Fig. 3, respectively.

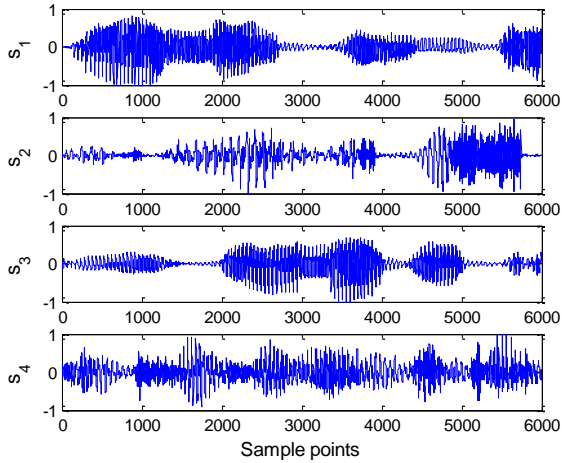


Fig. 2 Four speech signals as the sources

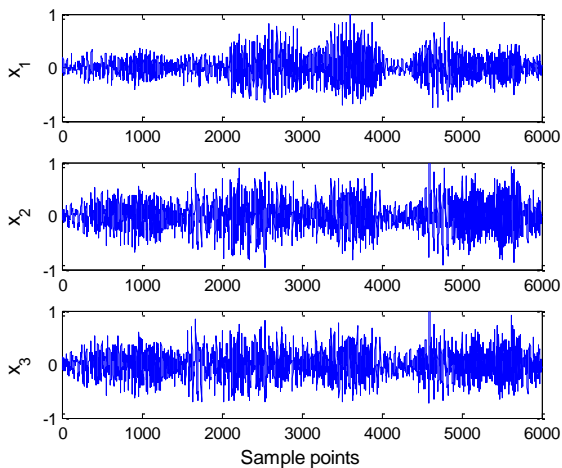
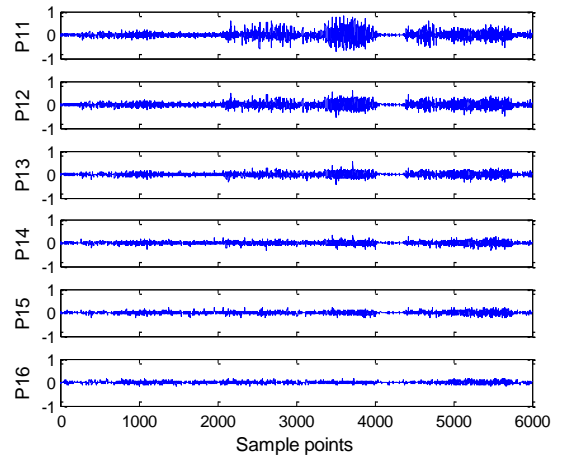


Fig. 3 The generated underdetermined mixtures

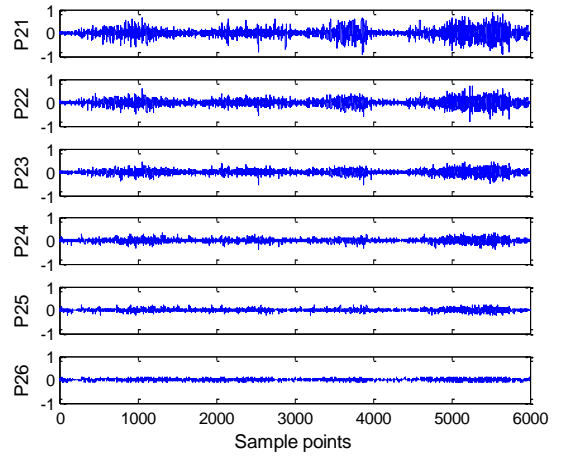
The local mean decomposition algorithm is performed on the mixtures to generate the product function components $\{P_{pq}\}$, which can be seen in Fig.4. For simplicity, only several main components are displayed. The correlation coefficients r_{pq} between each product function P_{pq} and the corresponding observation signal x_p are then calculated out and listed in Table 1.

It can be seen, from Table 1, that the correlation coefficients decrease as the second subscript q of the product functions increases. Thus several product functions with large correlation coefficient are reserved

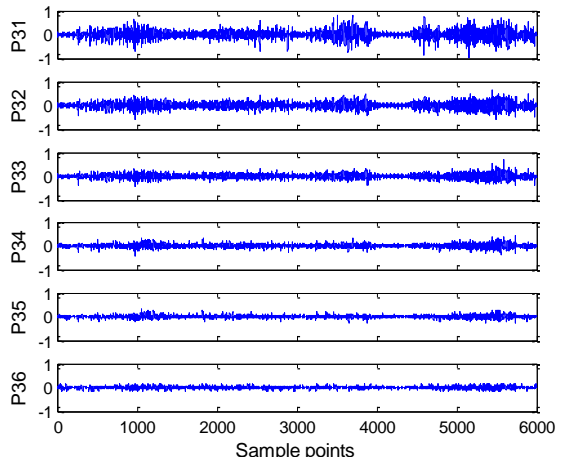
for constructing the new observation signals $\mathbf{x}_{new}(t)$. The proposed conjugate gradient based BSS algorithm is finally implemented on $\mathbf{x}_{new}(t)$. The estimated source signals are computed out, as is shown in Fig. 5. Comparing Fig. 5 with Fig. 2, it can be seen that the proposed method can recover the waveform of the sources very clear in the underdetermined cases up to permutation and scalar indeterminacy.



(a)



(b)



(c)

Fig. 4 Product functions of the mixtures by LMD (a. product functions of x_1 ; b. product functions of x_2 ; c. product functions of x_3)

TABLE I: THE CORRELATION COEFFICIENTS BETWEEN EACH PRODUCT FUNCTION AND THE CORRESPONDING OBSERVATION SIGNAL

r_{1q}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}
x_1	0.7062	0.3484	0.1932	0.1224	0.0895
r_{2q}	P_{21}	P_{22}	P_{23}	P_{24}	P_{25}
x_2	0.7515	0.4208	0.2057	0.1533	0.1107
r_{3q}	P_{31}	P_{32}	P_{33}	P_{34}	P_{35}
x_3	0.7446	0.3637	0.1678	0.1004	0.0838

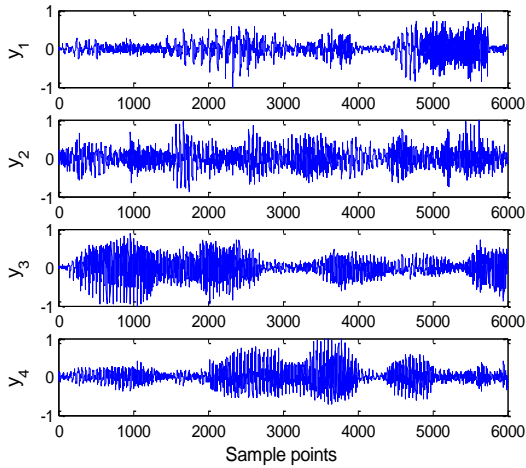


Fig. 5 The separated signals from the proposed underdetermined BSS algorithm

For the sake of quantitatively evaluating the performance of the proposed algorithm, the output results from the proposed LMD based algorithm (LMDBSS) together with an underdetermined BSS algorithm for non-sparse sources based on spatial time-frequency distribution (STFDBSS) [24] and an algorithm resorting to sparsity assumption [11] are presented for comparison. The separation performance of the three algorithms, measured by the mean squared error (MSE) criterion averaged over T independent runs, is defined as:

$$MSE = \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{T} \sum_{j=1}^T 10 \log_{10} \left(\frac{\|\hat{s}_{i,j} - s_i\|^2}{\|s_i\|^2} \right) \right) \quad (24)$$

where n is the number of source signals and $\hat{s}_{i,j}$ is the estimation of the normalized source signal s_i in the j th independent run. The mixtures are generated with the same scheme as the previous simulation. The additive noise are generated to be white and Gaussian with uncorrelated samples whose variance was assumed to be uniform. The algorithms are performed under different signal-to-noise ratio (SNR) varied from 0 to 30 dB by a step of 2.5 dB. The variation of MSEs with respect to SNR is shown in Fig. 6.

From Fig. 6, it can be seen that the sparsity-based underdetermined BSS algorithm is not able to extract the non-sparse source signals, and the proposed LMD based underdetermined BSS algorithm has a superior separation

performance, which always maintains a lower value of MSE than the other two algorithms.

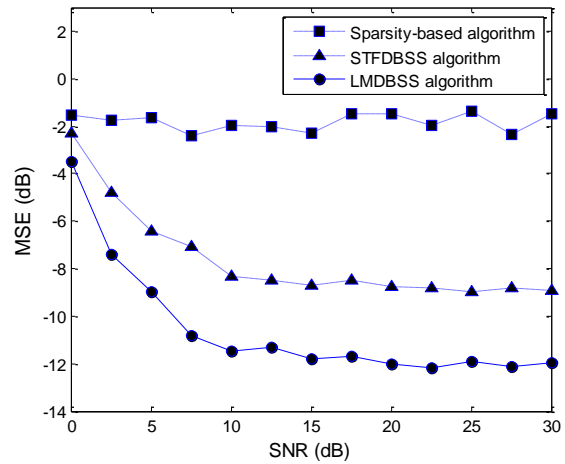


Fig. 6 The performance comparison by average MSE of the proposed LMDBSS algorithm with two previous underdetermined BSS algorithms

V. CONCLUSIONS

In this paper, a novel underdetermined blind source separation method is presented based on local mean decomposition and conjugate gradient algorithm. To make the underdetermined BSS problem become more tractable, some additional observation signals are firstly constructed through performing the local mean decomposition algorithm on the rough observations. Hence, the underdetermined mixing model is transformed to a determined model. Subsequently, the minimum mutual information principle is used for solving the regenerated determined BSS problem. The separating matrix is trained by the conjugate gradient learning algorithm in Riemannian space, and the kernel probability density estimation is employed to estimate the score function of the separated signals instead of selecting some nonlinear functions. The numerical simulations have shown the effectiveness of the proposed underdetermined BSS method. From the simulation results, it can be seen that local mean decomposition makes the underdetermined BSS problem become much easier, and the great advantages of the proposed underdetermined BSS method is that it does not resort to the sparsity constraint of the source signals, and it does not need to employ complicated approaches to estimate the unknown underdetermined mixing matrix, which are included in many former researches.

REFERENCES

- [1] A. Budipriyanto, "Blind source separation based dynamic parameter identification of a multi-story moment-resisting frame building under seismic ground motions," *Procedia Engineering*, vol. 54, pp. 299-307, 2013.
- [2] F. Gu, H. Zhang, and D. Zhu, "Blind separation of non-stationary sources using continuous density hidden Markov models," *Digital Signal Process.*, vol. 23, no. 5, pp. 1549-1564, Sep. 2013.

- [3] Y. Wang, Ö. Yılmaz, and Z. Zhou, "Phase aliasing correction for robust blind source separation using DUET," *Applied and Computational Harmonic Analysis*, vol. 35, no. 2, pp. 341-349, Sep. 2013.
- [4] L. D. Lathauwer, J. Castaing, and J. F. Cardoso, "Fourth-order cumulant-based blind identification of underdetermined mixtures," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2965-2973, 2007.
- [5] J. Thomas, Y. Deville, and H. Shahram, "Differential fast fixed-point algorithm for underdetermined instantaneous and convolutive partial blind source separation," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3717-3729, 2007.
- [6] Ö. Yılmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1830-1847, 2004.
- [7] P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," *Signal Process.*, vol. 81, no. 11, pp. 2353-2362, 2001.
- [8] A. Sadhu, B. Hazra, and S. Narasimhan, "Decentralized modal identification of structures using parallel factor decomposition and sparse blind source separation," *Mech. Syst. Signal Process.*, vol. 41, no. 1-2, pp. 396-419, Dec. 2013.
- [9] Y. Q. Li, A. Cichocki, and S. Amari, "Analysis of sparse representation and blind source separation," *Neural Computation*, vol. 16, no. 6, pp. 1193-1234, 2004.
- [10] M. Zibulevsky and B. A. Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural Computation*, vol. 13, no. 4, pp. 863-882, 2001.
- [11] Y. Q. Li, S. Amari, A. Cichocki, D. W. C. Ho, and S. L. Xie, "Underdetermined blind source separation based on sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 423-437, 2006.
- [12] F. J. Theis, C. G. Puntonet, and E. W. Lang, "Median-based clustering for underdetermined blind signal processing," *IEEE Signal Process. Letters*, vol. 13, no. 2, pp. 96-99, 2006.
- [13] J. J. Thiagarajan, K. N. Ramamurthy, and A. Spanias, "Mixing matrix estimation using discriminative clustering for blind source separation," *Digital Signal Process.*, vol. 23, no. 1, pp. 9-18, Jan. 2013.
- [14] P. G. Georgiev, F. Theis, and A. Cichocki, "Sparse component analysis and blind source separation of underdetermined mixtures," *IEEE Trans. Neural Networks*, vol. 16, no. 4, pp. 992-996, Jan. 2005.
- [15] J. S. Smith, "The local mean decomposition and its application to EEG perception data," *J. The Royal Society Interface*, vol. 2, no. 5, pp. 443-454, 2005.
- [16] Y. X. Wang, Z. J. He, J. W. Xiang, and Y. Y. Zi, "Application of local mean decomposition to the surveillance and diagnostics of low-speed helical gearbox," *Mech. and Machine Theory*, vol. 47, pp. 62-73, 2012.
- [17] Y. Yang, J. S. Cheng, and K. Zhang, "An ensemble local means decomposition method and its application to local rub-impact fault diagnosis of the rotor systems," *Measurement*, vol. 45, no. 3, pp. 561-570, 2012.
- [18] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley, New York, 2001.
- [19] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, 2010.
- [20] Y. Nishimori, S. Akaho, and M. D. Plumbley, "Natural conjugate gradient on complex flag manifolds for complex independent subspace," in *Proc. 18th International Conf. Artificial Neural Networks, Lecture Notes in Computer Science*, 2008, pp. 165-174.
- [21] E. Alan, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Analysis and Applications*, vol. 20, no. 2, pp. 303-353, 1998.
- [22] J. Karvanen and V. Koivunen, "Blind separation methods based on Pearson system and its extensions," *Signal Process.*, vol. 82, no. 4, pp. 663-673, 2002.
- [23] J. Karvanen, J. Eriksson, and V. Koivunen, "Pearson system based method for blind separation," in *Proc. 2nd Int. Workshop on Independent Component Analysis and Blind Signal Separation*, 2000, pp. 585-590.
- [24] D. Peng and Y. Xiang, "Underdetermined blind separation of non-sparse sources using spatial time-frequency distributions," *Digital Signal Process.*, vol. 20, no. 2, pp. 581-596, 2010.



Wei Li Ph.D student in control theory and control engineering of Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University. He received the M.S. degree in detection technology and automatic equipment from Anhui Polytechnic University in 2010. His current research interest covers data analysis and signal processing.



Huizhong Yang professor of Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University. She received his PhD degree in control theory and control engineering from East China University of Science and Technology in 2001. Her research interest covers modeling and analysis of complex industrial process.