A New Second Order Method for Blind Signal Separation from Convolutive Mixtures

Y. Xiang, S. Nahavandi, H. Trinh, and H. Zheng

School of Engineering and Technology,
Deakin University, Geelong, VIC 3217, Australia
{yxiang,nahavand,hmt,hzheng}@deakin.edu.au

Abstract

This paper presents a new approach to separate colored signals mixed by FIR (finite impulse response) and MIMO (multiple-input multiple-output) channels. A cost function is proposed by employing linear constraint to the demixing vectors. The linear constraint is shown to be sufficient for avoiding trivial solution. The minimization of the cost function is performed using the Lagrangian method. Simulation results demonstrate the performance of the algorithm.

1 Introduction

Blind signal separation (BSS) from FIR MIMO channel outputs is a fundamental problem encountered in wireless communications, speech enhancement, image reconstruction and medical signal analysis. It is blind in the sense that neither the channel nor the signals are known. While some methods try to retrieve the input signals (up to scaling and permutation), complete recovery of input signals is not necessary for some practical applications. For example, in the *Cocktail Party* problem, speech enhancement can be achieved by separating the desired speech signals from interfering sources. Although the separated signals may be convolutive distorted versions of the original speech signals, this distortion is to some extent not detectable by human ears as long as the order of convolution is not high.

Among the BSS methods, most of them are based on higher order statistics (HOS) because they only exploit spatial diversity via multiple sensors [5, 7]. This paper focus on blind signal separation using second order statistics (SOS). Compared with the HOS-based methods, the algorithms based on second order statistics are more reliable and efficient if specific assumptions about channel and signals are met. Prior work on this topic includes [1, 2, 6, 8]. While the methods reported in [2, 8] separate signals mixed by two-input two-output (TITO) channels, the method by Sahlim and Broman [6] and the method by Ahmed *et al.* [1] can perform signal separation from multiple-input multiple-output (MIMO) channels. The methods shown in [1, 6] use Newton based optimization techniques, and require the main diagonal elements of the channel matrix to be unity which is a strong condition.

In this paper, we develop a new method for blind separation of colored signals distorted by FIR MIMO channels. Different from the method known as blind identification via decorrelating subchannels (BIDS) [3, 4], the proposed method minimizes a cost function, which is the sum of the squared cross correlations of the signal estimates, subject to a linear constraint by using the Lagrangian function. It does not impose the condition of unity main diagonal elements on the channel matrix. Simulation examples show the effectiveness of the new approach.

2 Problem Formulation

The FIR MIMO system we consider is described as follows

$$\mathbf{y}(n) = \mathbf{H}(z)\mathbf{x}(n) + \mathbf{w}(n) \quad (1)$$

where $\mathbf{x}(n)$ denotes the $I \times 1$ *unknown* input vector, $\mathbf{y}(n)$ the $I \times 1$ output vector, $\mathbf{H}(z)$ the $I \times I$ channel matrix and $\mathbf{w}(n)$ the $I \times 1$ noise vector. We assume, without loss of generality, that all the data in the time domain are real valued, and the noise is spatially and temporally white and uncorrelated with input signals. The noise variance can be estimated by technique shown in [3] and hence be used to reduce (or remove asymptotically) the noise contribution to any forms of the second order statistics. So we will ignore in the rest of this paper the noise term for convenience and use the following simplified model:

$$\mathbf{y}(n) = \mathbf{H}(z)\mathbf{x}(n). \tag{2}$$

The objective of BSS is to find an $I \times I$ separator $\mathbf{G}(z)$ such that

$$\mathbf{C}(z) \triangleq \mathbf{G}(z)\mathbf{H}(z) = \mathbf{P}\mathbf{\Lambda}(z)$$

where $\mathbf{P}$ is a permutation matrix and $\mathbf{\Lambda}(z)$ is a diagonal polynomial matrix. The polynomial matrices $\mathbf{C}(z)$, $\mathbf{G}(z)$ and $\mathbf{H}(z)$ have degrees L_C, L_G and L_H respectively, and $L_C = L_G + L_H$.

Let

$$\mathbf{v}(n) \triangleq \mathbf{G}(z)\mathbf{y}(n) = \mathbf{C}(z)\mathbf{x}(n). \tag{3}$$

Denote the correlation matrix of $\mathbf{x}(n)$ by

$$\mathbf{R}_{\mathbf{xx}}(\tau) \triangleq E[\mathbf{x}(n+\tau)\mathbf{x}(n)^T]$$

and the power spectral matrix of $\mathbf{x}(n)$ (the Z-transform of $\mathbf{R}_{\mathbf{xx}}(\tau)$) by $\mathbf{S}_{\mathbf{xx}}(z)$, where E denotes the expectation operator and ^{T} the transpose operator. Similar definitions are applied to other vectors. Then, (2) and (3) imply

$$\mathbf{S}_{\mathbf{yy}}(z) = \mathbf{H}(z)\mathbf{S}_{\mathbf{xx}}(z)\mathbf{H}(z^{-1})^T$$

and

$$\mathbf{S}_{\mathbf{vv}}(z) = \mathbf{C}(z)\mathbf{S}_{\mathbf{xx}}(z)\mathbf{C}(z^{-1})^T$$

respectively. To perform signal separation, we assume

(a) The input power spectral matrix $\mathbf{S}_{\mathbf{xx}}(z)$ is diagonal and the diversity of $\mathbf{S}_{\mathbf{xx}}(z)$ is larger than the degree of $\mathbf{C}(z)$.
(b) The channel matrix $\mathbf{H}(z)$ is nonsingular (i.e., full rank almost everywhere).

The notation of diversity is an important concept and is defined in [3]. To ensure $\mathbf{G}(z)\mathbf{H}(z)$ to be diagonalizable by $\mathbf{G}(z)$, it is sufficient if the number of linear and scalar equations corresponding to all off-diagonal polynomial elements of $\mathbf{G}(z)\mathbf{H}(z)$ is less than the total number of the scalar parameters in $\mathbf{G}(z)$, i.e., $I^2(L_G+1) > (I^2-I)(L_G+L_H+1)$. This leads to the following condition

$$L_G \geq (I-1)L_H. \tag{4}$$

Given $\mathbf{C}(z)$ a nonsingular polynomial matrix, it is shown [3] that the diagonalization of $\mathbf{S}_{\mathbf{vv}}(z)$ leads to that of $\mathbf{C}(z)$ if *(a)* and *(b)* hold.

3 Principle of Signal Separation

Define

$$\bar{\mathbf{C}} \triangleq [\mathbf{C}(0), \mathbf{C}(1), \cdots, \mathbf{C}(L_C)]$$

$$\bar{\mathbf{G}} \triangleq [\mathbf{G}(0), \mathbf{G}(1), \cdots, \mathbf{G}(L_G)]$$

where $\mathbf{G}(l_1)$, $l_1 = 0, 1, \cdots, L_G$ and $\mathbf{C}(l_2)$, $l_2 = 0, 1, \cdots, L_C$ represent the coefficients of $\mathbf{G}(z)$ at z^{-l_1} and $\mathbf{C}(z)$ at z^{-l_2}, respectively. Then (3) can be rewritten as

$$\mathbf{v}(n) = \bar{\mathbf{G}}\bar{\mathbf{y}}(n) = \bar{\mathbf{C}}\bar{\mathbf{x}}(n)$$

where

$$\bar{\mathbf{x}}(n)^T \triangleq [\mathbf{x}(n)^T, \mathbf{x}(n-1)^T, \cdots, \mathbf{x}(n-L_C)^T]$$

$$\bar{\mathbf{y}}(n)^T \triangleq [\mathbf{y}(n)^T, \mathbf{y}(n-1)^T, \cdots, \mathbf{y}(n-L_G)^T].$$

It follows

$$\mathbf{R}_{\mathbf{vv}}(\tau) = \bar{\mathbf{G}}\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau)\bar{\mathbf{G}}^T.$$

In practice, $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau)$ will be replaced by its estimate $\hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau)$ with finite sample size N, which is defined as

$$\begin{aligned} \hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau) &\triangleq \frac{1}{N}\sum_{n=0}^{N-1}\bar{\mathbf{y}}(n+\tau)\bar{\mathbf{y}}(n)^T \\ &= \frac{1}{N}\bar{\mathbf{Y}}(\tau)\bar{\mathbf{Y}}(0)^T \end{aligned} \tag{5}$$

where

$$\bar{\mathbf{Y}}(\tau) = [\bar{\mathbf{y}}(\tau), \bar{\mathbf{y}}(\tau+1), \cdots, \bar{\mathbf{y}}(\tau+N-1)].$$

Similarly, we denote

$$\bar{\mathbf{X}}(\tau) = [\bar{\mathbf{x}}(\tau), \bar{\mathbf{x}}(\tau+1), \cdots, \bar{\mathbf{x}}(\tau+N-1)].$$

Then we have

$$\bar{\mathbf{y}}(n) = \mathbf{H}_{sw}\bar{\mathbf{x}}(n) \tag{6}$$

$$\bar{\mathbf{Y}}(\tau) = \mathbf{H}_{sw}\bar{\mathbf{X}}(\tau) \tag{7}$$

where $\mathbf{H}_{sw}$ represents the generalized Sylvester matrix associated to $\mathbf{H}(z)$ defined by

$$\mathbf{H}_{sw} \triangleq \begin{bmatrix} \mathbf{H}(0) & \cdots & \cdots & \mathbf{H}(L_H) & & \\ & \ddots & & & \ddots & \\ & & \mathbf{H}(0) & \cdots & \cdots & \mathbf{H}(L_H) \end{bmatrix} \tag{8}$$

and $\mathbf{H}(l)$ the coefficient of $\mathbf{H}(z)$ at z^{-l}.

3.1 Cost Function

Denote $\bar{\mathbf{G}}^T = [\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_I]$. The diagonalization of $\mathbf{S}_{\mathbf{vv}}(z)$ is implemented by minimizing the squared sum of the off-diagonal elements of $\mathbf{S}_{\mathbf{vv}}(z)$. It is equivalent to minimize the following cost function

$$\mathbf{J} = 2\mathbf{g}_i^T \mathbf{R}_i \mathbf{g}_i + \mathbf{J}(\mathbf{g}_p, p \neq i) \tag{9}$$

where

$$\mathbf{R}_i = \sum_{j \neq i} \sum_{\tau=-K}^{K} \hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau) \mathbf{g}_j \mathbf{g}_j^T \hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau)^T$$

$$\mathbf{J}(\mathbf{g}_p, p \neq i) = \sum_{p,q \neq i,\ p \neq q} \sum_{\tau=-K}^{K} \left(\mathbf{g}_p^T \hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau) \mathbf{g}_q \right)^2.$$

The cost function $\mathbf{J}$ is a non-quadratic function of $\bar{\mathbf{G}}$. But it is quadratic with respect to each (individual) row of $\bar{\mathbf{G}}$. To avoid trivial solution, a linear constraint

$$\mathbf{d}^T \mathbf{g}_i = 1 \tag{10}$$

is imposed on $\mathbf{g}_i$ for $i = 1, 2, \cdots, I$, where $\mathbf{d} = [1, 1, \cdots, 1]^T$.

3.2 Existence of Solution

From (5) and (7), we have

$$\mathbf{g}_i^T \hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau) \mathbf{g}_j = \frac{1}{N} \mathbf{g}_i^T \mathbf{H}_{sw} \bar{\mathbf{X}}(\tau) \bar{\mathbf{X}}(0)^T \mathbf{H}_{sw}^T \mathbf{g}_j.$$

It is obvious that if $\mathbf{H}_{sw}$ is not full row rank, there exist nonzero $\mathbf{g}_i$ satisfying $\mathbf{d}^T \mathbf{g}_i = 1$, $i = 1, 2, \cdots, I$ such that $\mathbf{g}_i^T \mathbf{H}_{sw} = \mathbf{0}$ which leads to

$$\mathbf{g}_i^T \hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau) \mathbf{g}_j = 0$$

$$\mathbf{g}_i^T \hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau) \mathbf{g}_i = 0$$

at the same time. To avoid trivial solution, $\mathbf{H}_{sw}$ needs to be of full row rank. We will show below that the generalized Sylvester matrix associated to a nonsingular square channel matrix has full row rank.

Let $\mathbf{q}_i$, $i = 0, 1, \cdots, L_G$ be $I \times 1$ vectors and $\mathbf{q}^T \triangleq [\mathbf{q}_0^T, \mathbf{q}_1^T, \cdots, \mathbf{q}_{L_G}^T]$. Denote $\mathbf{q}(z) = \sum_{i=0}^{L_G} \mathbf{q}_i z^{-i}$. If $\mathbf{H}_{sw}$ is not of full row rank, there exists $\mathbf{q} \neq \mathbf{0}$ such that

$$\mathbf{q}^T \mathbf{H}_{sw} = \mathbf{0}.$$

This is equivalent to

$$\mathbf{q}(z)^T \mathbf{H}(z) = \mathbf{0} \quad \textit{for each } z \tag{11}$$

(i.e., $\mathbf{H}(z)$ is singular) which contradicts the nonsingularity of $\mathbf{H}(z)$. So, $\mathbf{H}_{sw}$ must be of full row rank.

4 Algorithm Implementation

When $\mathbf{g}_p$, $p = 1, \cdots, i-1, i+1, \cdots, I$ are fixed, the minimization of criterion (9) subject to constraint (10) can be performed using the Lagrange method:

$$\min_{\mathbf{g}_i, \lambda_i} \mathcal{L}(\mathbf{g}_i, \lambda_i) = \mathbf{g}_i^T \mathbf{R}_i \mathbf{g}_i + \lambda_i (\mathbf{d}^T \mathbf{g}_i - 1) \tag{12}$$

with $i = 1, 2, \cdots, I$. The minimization of the above equation with respect to $\mathbf{g}_i$ and λ_i leads to the following equations:

$$2\mathbf{R}_i \mathbf{g}_i + \lambda_i \mathbf{d} = \mathbf{0} \tag{13}$$

$$\mathbf{d}^T \mathbf{g}_i = 1 \tag{14}$$

where $\mathbf{0} = [0, 0, \cdots, 0]^T$ in this section. From (13) and (14), we have

$$\begin{bmatrix} \mathbf{g}_i \\ \lambda_i \end{bmatrix} = \begin{bmatrix} 2\mathbf{R}_i & \mathbf{d} \\ \mathbf{d}^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}. \tag{15}$$

Since $\begin{bmatrix} 2\mathbf{R}_i & \mathbf{d} \\ \mathbf{d}^T & 0 \end{bmatrix}$ is a Hermitian matrix, its inverse is also a Hermitian matrix with the form $\begin{bmatrix} \mathbf{Q} & \mathbf{q} \\ \mathbf{q}^T & q \end{bmatrix}$. By exploiting this property of Hermitian matrix, one has

$$\begin{bmatrix} 2\mathbf{R}_i & \mathbf{d} \\ \mathbf{d}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{Q} & \mathbf{q} \\ \mathbf{q}^T & q \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{16}$$

where $\mathbf{I}$ is the identity matrix. (16) can be replaced by the following equations:

$$2\mathbf{R}_i \mathbf{Q} + \mathbf{d}\mathbf{q}^T = \mathbf{I}$$

$$2\mathbf{R}_i \mathbf{q} + q\mathbf{d} = \mathbf{0}$$

$$\mathbf{d}^T \mathbf{Q} = \mathbf{0}^T$$

$$\mathbf{d}^T \mathbf{q} = 1.$$

After algebraic calculation and manipulation, we obtain

$$\mathbf{g}_i = \frac{\mathbf{R}_i^{-1} \mathbf{d}}{\mathbf{d}^T \mathbf{R}_i^{-1} \mathbf{d}} \tag{17}$$

$$\lambda_i = -\frac{2}{\mathbf{d}^T \mathbf{R}_i^{-1} \mathbf{d}}. \tag{18}$$

Thus, a recursive quadratic minimization algorithm can be developed to update a row vector of $\bar{\mathbf{G}}$ at each time. The proposed algorithm is formulated as follows:

- Step 1. Randomly initialize $\mathbf{g}_i^{(0)}, i = 1, 2, \cdots, I$ and compute $\mathbf{J}^{(0)}$ using the following equations:

$$\mathbf{J}^{(l)} \triangleq \sum_{1 \leq i \neq j \leq I} \sum_{\tau=-K}^{K} \left(\mathbf{g}_i^{(l)T} \hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau) \mathbf{g}_j^{(l)} \right)^2. \tag{19}$$

- Step 2. Set iteration index $l = 1$.
- Step 3. For $i = 1, 2, \cdots, I$, update $\mathbf{R}_i^{(l)}$ and $\mathbf{g}_i^{(l)}$ as follows:

$$\mathbf{R}_i^{(l)} = \sum_{j>i} \sum_{\tau=-K}^{K} \hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau)\mathbf{g}_j^{(l-1)}\mathbf{g}_j^{(l-1)T}\hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau)^T + \sum_{j<i} \sum_{\tau=-K}^{K} \hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau)\mathbf{g}_j^{(l)}\mathbf{g}_j^{(l)T}\hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}(\tau)^T$$

$$\mathbf{g}_i^{(l)} = \frac{(\mathbf{R}_i^{(l)})^{-1}\mathbf{d}}{\mathbf{d}^T(\mathbf{R}_i^{(l)})^{-1}\mathbf{d}}.$$

- Step 4. Compute $\mathbf{J}^{(l)}$ from (19). If $|\mathbf{J}^{(l)} - \mathbf{J}^{(l-1)}| > \epsilon_1$ (ϵ_1 denotes a selected threshold), then $l = l+1$, go to Step 3. Otherwise, go to Step 5.
- Step 5. If $\mathbf{J}^{(l)} > \epsilon_2$ (ϵ_2 denotes another threshold), stop. Otherwise, go to Step 1.

Remark: Under ideal condition, $\mathbf{J}$ should be zero if the global minimum is reached. This implies that when $\mathbf{g}_i$, $i = 1, 2, \cdots, I$ approach their true values, $\mathbf{R}_i$ approaches to be singular, which leads to λ_i approaching to zero. From (13) we can see that the solution of (17) approaches to be a null vector of $\mathbf{R}_i$.

5 Simulations

This section gives some numerical simulations to evaluate the performance of the new approach. In our simulations, the FIR MIMO systems had two inputs and two outputs. White Gaussian noise of variance σ^2 was added to the channel outputs and the signal to noise ratio (SNR) was defined as SNR$\triangleq -10log_{10}\sigma^2$. The separator $\mathbf{G}(z)$ was chosen to be of degree $L_G = L_H$. The performance was measured by the mean rejection level (we assume here that the permutation indeterminacy is $\mathbf{P} = \mathbf{I}$):

$$\mathcal{I}perf \triangleq \sum_{i \neq j} \sum_{\tau=0}^{L_G+L_H} \frac{E|(\mathbf{C}(\tau))_{ij}|^2}{E|(\mathbf{C}(\tau))_{ii}|^2}.$$

The mean rejection level was estimated by averaging 60 independent runs.

Example 1: An FIR MIMO channel of degree 3 was considered, whose parameters are tabulated in Table 1. Two colored signals of unit variance were generated by independent filters driven by Gaussian white inputs. The two filters are polynomials of degree 150 with all coefficients independently chosen from Gaussian distribution of zero mean and unit variance. Figure 1 and Figure 2 demonstrate the performance of the proposed method via SNR and sample size, respectively. In Figure 1, the sample size was 44000 and SNR was varied from 0dB to 35dB. The technique shown in [3] was used to compensate the influence of noise at low SNR levels. In Figure 2, the sample size varied from 10000 to 50000 while the SNR was consistently at 30dB.

Table 1. The parameters of $\mathbf{H}(z)$ in *Example 1.*

z^{-1}	$[\mathbf{H}(z)]_{11}$	$[\mathbf{H}(z)]_{12}$	$[\mathbf{H}(z)]_{21}$	$[\mathbf{H}(z)]_{22}$
$l = 0$	0.161	0.601	0.883	0.362
$l = 1$	0.45	0.468	0.643	0.497
$l = 2$	0.767	0.777	0.045	0.208
$l = 3$	0.397	0.625	0.481	0.626

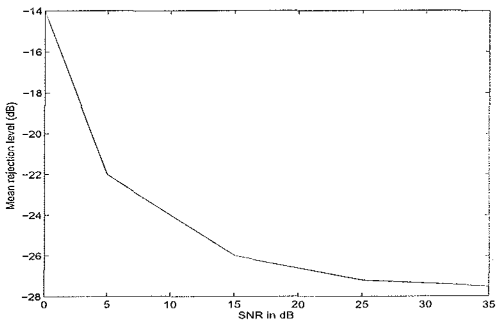

Figure 1. Performance versus SNR.

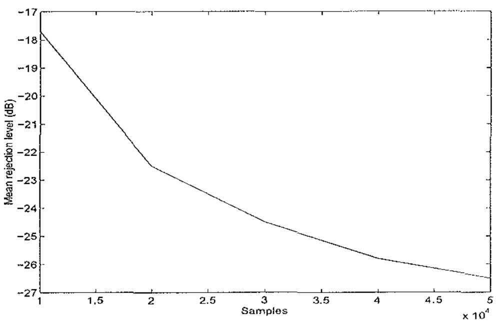

Figure 2. Performance versus sample size.

Table 2. The parameters of $\mathbf{H}(z)$ in *Example 2.*

z^{-1}	$[\mathbf{H}(z)]_{11}$	$[\mathbf{H}(z)]_{12}$	$[\mathbf{H}(z)]_{21}$	$[\mathbf{H}(z)]_{22}$
$l = 0$	-1.348	0.386	-1.523	-2.524
$l = 1$	1.126	0.178	-1.401	1.123
$l = 2$	0.844	-0.668	1.432	-1.68

Example 2: The proposed method was used to separate real speech signals from the outputs of an FIR MIMO channel of degree 2. The channel parameters are shown in Table 2. The signal to noise ratio was kept at 40dB. Figure 3 contains two real speech signals sampled at 7000 Hz, which were obtained from the Linguistic Data Consortium. Speech signal 1 and 2 are the sentences "She had your dark suit in greasy wash water all year" and "Don't ask me to carry an oily rag like that", respectively. The outputs of two (simulated) microphones are shown in Figure 4. Figure 5 shows the estimates of the original speech signals. Slight distortion can be observed in the reconstructed signals in comparison with the original ones. However, by playing these reconstructed signals, we found that the convolutive distortion could not be detected by ears.

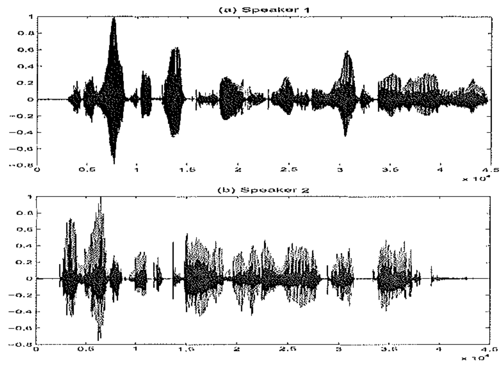

Figure 3. The original speech signals.

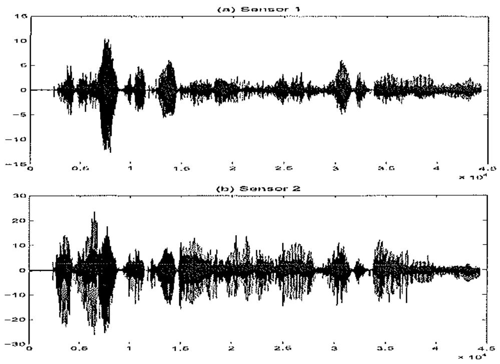

Figure 4. The mixed speech signals.

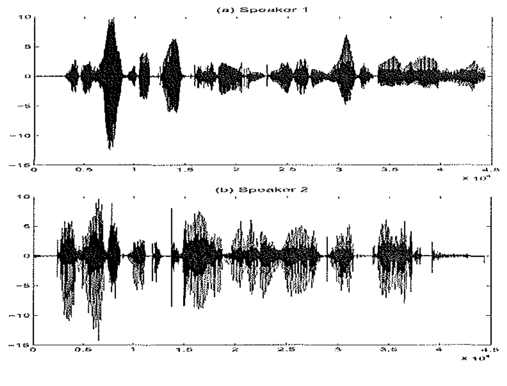

Figure 5. The reconstructed speech signals.

6 Conclusion

In this paper, we have presented a new approach to separate colored signals from their convolutive mixtures. Unlike most existing signal separation methods, it exploits only the second order statistics of the measurable signals and requires weaker condition on the FIR MIMO channel system. The channel matrix is allowed to have non-unit main diagonal elements. Based on a linear constraint on the row vectors of the separator, a Lagrangian function was constructed and minimized. The performance of the algorithm was demonstrated by numerical simulation results.

References

[1] A. Ahmed, P. J. W. Rayner, and S. J. Godsill, Recursive decorrelation for blind convolutive signal separation, *Proc. of Int. Workshop on Independent Component Analysis and Blind Signal Separation*, Aussois, France, Jan. 1999.

[2] S. V. Gerven and D. V. Compernolle, Signal separation by symmetric adaptive decorrelation: sta-

bility, convergence, and uniqueness, *IEEE Trans. on Signal Processing*, Vol. 43, pp. 1602–1612, July 1995.

[3] Y. Hua, Blind identification and equalization of channels driven by colored signals, Chapter 4, *Signal Processing Advances in Wireless Communications*, Prentice-Hall, 2000.

[4] Y. Hua, S. An, and Y. Xiang, Blind Identification and Equalization of FIR MIMO Channels by BIDS, *Proc. of 2001 IEEE International Conference on Acoustics, Speech and Signal Processing*, Salt Lake City, USA, May 2001.

[5] A. Mansour and C. Jutten, Fourth order criteria for blind separation of sources, *IEEE Trans. on Signal Processing*, vol.43, No.8, pp. 2022–2025, Aug. 1995.

[6] H. Sahlin and H. Broman, MIMO signal separation for FIR channels: a criterion and performance analysis, *IEEE Trans. on Signal Processing*, vol. 48, No. 3, pp. 642–649, Mar. 2000.

[7] H.N. Thi and C. Jutten, Blind source separation for convolutive mixtures, *Signal Separation*, vol.45, pp. 209–229, 1995.

[8] E. Weinstein, M. Feder, and A. V. Oppenheim, Multi-channel signal separation by decorrelation, *IEEE Trans. on Speech Audio Processing*, Vol. 1, pp. 405–413, Oct. 1993.