USING WOLD DECOMPOSITION PRINCIPLE IN BLIND SEPARATION OF JOINT STATIONARY CORRELATED SOURCES

Masoud R. Aghabozorgi and Ali M. Doost-Hoseini
aghabozorgi@hotmail.com alimdh@cc.iut.ac.ir
Department of Electrical and Computer Engineering,
Isfahan University of Technology, Isfahan, 84154, Iran

ABSTRACT

The separation of unobserved sources from mixed observed data is a fundamental signal processing problem. Most proposed techniques for solving this problem rely on independence or at least uncorrelation assumption of source signals.

In this paper an algorithm is introduced for source signals that are correlated with each other. The method uses a preprocessing technique based on Wold decomposition principle for extracting desired and proper information from the predictable part of the observed data, and exploits approaches based on second-order statistics to estimate the mixing matrix and source signals.

Keywords: blind source separation, wold decomposition, second-order statistics

1. INTRODUCTION

Blind source separation (BSS) consists of recovering source signals from several observed noisy mixtures of them. The observations are obtained from a set of sensors, each receiving a different combination of the source signals.

The lack of prior information can be compensated by considering particular source statistics assumptions. The most popular condition used by BSS techniques is the statistically assumption of independence between the source signals. These techniques assume that the primary sources are statistically independent, and therefore the goal in these techniques is to achieve a separation process that produces outputs as independent as possible[1],[2]. A less stringent condition is uncorrelation of sources. These techniques exploit temporal correlation of each source signal (second-order blind identification), and use a joint diagonalization method of several correlation matrices[3].

Other techniques and algorithms introduced for BSS use special source signals structures, e.g. CMA(Constant Modulus Algorithm) or special structure for combination system such as MUSIC and ESPIRIT[4].

In this paper, the aim is to propose a solution to BSS problem for correlated source signals ,that can be called CS-SOBI, without imposing special structures on signals or mixing matrix. This paper is organized as follows: In section 2, the problem of BSS is stated along with the related assumptions. Proposed pre-separation procedure is introduced in section 3. Section 4 expresses BSS algorithm, and simulation results are presented in section 5. Concluding remarks are given in section 6.

2. PROBLEM FORMULATION

Assume that d signals $s_1(t),\ldots,s_d(t)$ are transmitted from d sources at different locations. By considering a narrowband time-invariant channel, what we receive at m sensors (antennas) will be instantaneous linear combinations of these signals that construct observation data:

$$\mathbf{x}(t) = \mathbf{a}_1 . s_1(t) + \ldots + \mathbf{a}_d . s_d(t) + \mathbf{n}(t) \qquad (1)$$

Thus the model is as follows:

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t) = \mathbf{A}.\mathbf{s}(t) + \mathbf{n}(t) \qquad (2)$$

where $\mathbf{x}(t) \in \Re^{m\times 1}$ is the observed data vector from m sensors, $\mathbf{s}(t) \in \Re^{d\times 1}$ is the signal vector, composed of d unknown source signals, $\mathbf{A} = [\mathbf{a}_1,\ldots,\mathbf{a}_d] \in \Re^{m\times d}$ characterizes the unknown channel and is referred to as "mixing matrix", $\mathbf{n}(t) \in \Re^{m\times 1}$ is the additive noise vector at the sensor array. Following assumptions are considered in the model:

A1) Each element of $\mathbf{s}(t)$ (source signals) is a zero-mean, stationary process.

A2) The additive noise $\mathbf{n}(t)$ is assumed as a stationary, white zero-mean random process, independent of source signals.

A3) Mixing matrix $\mathbf{A}$ has full column rank.

It must be emphasized here that we don't impose any assumption about independence or uncorrelation of source signals. In other words, the source signals can be correlated, and only the following assumption is considered:

A4)Source signals are jointly stationary.

3. PRE-SEPARATION PROCEDURE

The main step in our approach for correlated sources is a pre-separation process. The observed data is decomposed into regular and predictable components, using Wold decomposition. In the predictable component, the combination of uncorrelated contributions of source signals is identified on whose basis $\mathbf{A}$ (and consequently the source signals) is estimated using second order statistics.

3.1. Wold Decomposition

An arbitrary process can be written as a sum:

$$s(t) = s_{\mathrm{r}}(t) + s_{\mathrm{p}}(t) \qquad (3)$$

where $s_{\mathrm{r}}(t)$ and $s_{\mathrm{p}}(t)$ are regular and predictable processes. This expansion is called Wold decomposition. In [6] it has been proved that the processes $s_{\mathrm{p}}(t)$ & $s_{\mathrm{r}}(t)$ are orthogonal. Furthermore, $s_{\mathrm{p}}(t)$ is comprised of complex exponentials:

$$s_{\mathrm{p}}(t) = \mathbf{c}_0 + \sum_{\mathrm{i}} \mathbf{c}_{\mathrm{i}} . \exp(\mathrm{j}\omega_{\mathrm{i}} t) \qquad (4)$$

where $\mathbf{c}_{\mathrm{i}}$'s are orthogonal zero-mean random variables. Hence, $s_{\mathrm{p}}(t)$ has a line spectrum:

$$P_{s_{\mathrm{p}}}(\omega) = \sum_{\mathrm{i}} 2\pi\alpha_{\mathrm{i}}\, \delta(\omega - \omega_{\mathrm{i}}) \qquad (5)$$

but $s_{\mathrm{r}}(t)$ has a smooth spectrum.

3.2. Observation Decomposition

In this subsection a method is proposed for extracting and decomposing some information from the regular and predictable parts of the observation data. For simplicity ,a special case of model (2) with d=2 & m=2 is considered, that can be extended to general cases. So, we have the following model satisfying conditions expressed in section 2:

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \mathbf{A}. \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \end{bmatrix} \qquad (6)$$

where $s_1(t)$ & $s_2(t)$ are the source signals and $\mathbf{A} = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix}$ is the mixing matrix.

Regular and predictable parts of source signal are indicated by $s_{\mathrm{ir}}(t)$ and $s_{\mathrm{ip}}(t)$ ($i = 1,2$):

$$s_{\mathrm{i}}(t) = s_{\mathrm{ip}}(t) + s_{\mathrm{ir}}(t) \qquad (7)$$

where,

$$s_{1\mathrm{p}}(t) = \sum_{\mathrm{k}} \mathbf{a}_{\mathrm{k}} . \exp(\mathrm{j}\omega_{1\mathrm{k}} t) \qquad (9)$$

$$s_{2\mathrm{p}}(t) = \sum_{l} \mathbf{b}_{l} . \exp(\mathrm{j}\omega_{2l} t) \qquad (10)$$

in which $\{\mathbf{a}_{\mathrm{k}}\}$ & $\{\mathbf{b}_l\}$ are sets of orthogonal random variables, and $\{\omega_1\}$ &$\{\omega_2\}$ are proper frequency sets. Also, since source signals are assumed jointly stationary, $\mathbf{a}_{\mathrm{k}}$ & $\mathbf{b}_l$ corresponding to $\omega_{1\mathrm{k}} \neq \omega_{2l}$ are orthogonal.

Using (6) , (7) and the fact that regular and predictable parts in each signal are orthogonal ,we get:

$$x_{\mathrm{i}}(t) = x_{\mathrm{ip}}(t) + x_{\mathrm{ir}}(t) + n_{\mathrm{i}}(t) \quad ; \ \mathrm{i} = 1,2 \qquad (11)$$

where

$$x_{\mathrm{ip}}(t) = \alpha_{\mathrm{i}} s_{1\mathrm{p}}(t) + \beta_{\mathrm{i}} s_{2\mathrm{p}}(t) \quad ; \ \mathrm{i} = 1,2 \qquad (12)$$

From (9),(10),(12) obtains:

$$x_{\mathrm{ip}}(t) = \sum_{\mathrm{q}} \mathbf{d}_{\mathrm{iq}} . \exp(\mathrm{j}\omega_{\mathrm{q}} t) \quad ; \ \mathrm{i} = 1,2 \qquad (13)$$

where $\{\mathbf{d}_{\mathrm{iq}}\}$ are orthogonal random variables and $\{\omega_{\mathrm{q}}\} = \{\omega_1\} \cup \{\omega_2\}$.

From these equations, each observation signal has regular and predictable components, each corresponding to the combination of individual regular and predictable parts of source signals. A spectral method for separating these parts follows.

The correlation functions of the observation data are given by: (i,j = 1,2)

$$\mathrm{r}_{\mathrm{ij}}^{\mathrm{x}}(\tau) = \mathrm{r}_{\mathrm{ijr}}^{\mathrm{x}}(\tau) + \mathrm{r}_{\mathrm{ijp}}^{\mathrm{x}}(\tau) + \mathrm{N}_0 . \delta(\tau) \qquad (14)$$

where N_0 is the noise variance and $\mathrm{r}_{\mathrm{ijr}}^{\mathrm{x}}(\tau)$ & $\mathrm{r}_{\mathrm{ijp}}^{\mathrm{x}}(\tau)$ are correlation functions of regular and predictable parts:

$$\mathrm{r}_{\mathrm{ijr}}^{\mathrm{x}}(\tau) = \mathrm{E}\{x_{\mathrm{ir}}(t+\tau) x_{\mathrm{jr}}^{*}(t)\} \qquad (15)$$

$$\mathrm{r}_{\mathrm{ijp}}^{\mathrm{x}}(\tau) = \mathrm{E}\{x_{\mathrm{ip}}(t+\tau) x_{\mathrm{jp}}^{*}(t)\} = \sum_{\mathrm{q}} \mathrm{E}\{\mathbf{d}_{\mathrm{iq}} \mathbf{d}_{\mathrm{iq}}^{*}\} . \exp(\mathrm{j}\omega_{\mathrm{q}} t) \qquad (16)$$

Hence power spectral density(psd) and cross spectral density(csd) functions of observations have the forms:

$$P_{\mathrm{ij}}^{\mathrm{x}}(\omega) = P_{\mathrm{ijr}}^{\mathrm{x}}(\omega) + P_{\mathrm{ijp}}^{\mathrm{x}}(\omega) + \mathrm{N}_0 \qquad (17)$$

where,

$$P_{\mathrm{ijp}}^{\mathrm{x}}(\omega) = \sum_{\mathrm{q}} 2\pi . \mathrm{E}\{\mathbf{d}_{\mathrm{iq}} \mathbf{d}_{\mathrm{jq}}^{*}\} . \delta(\omega - \omega_{\mathrm{q}}) \qquad (18)$$

As expected, the spectra of the predictable parts are pure impulsive. So, it is possible to detect and separate these components in the observation spectra. Consequently, correlation functions ($\mathrm{r}_{\mathrm{ijp}}^{\mathrm{x}}(\tau)$) of the predictable parts are obtained that will be used next.

3.3. Extracting Desired Information From Predictable Part

Rewriting predictable parts of source signals (9)-(10), considering $\{\Omega_n\}$ as the common frequency set, we obtain:

$$s_{1p}(t) = \sum_{k \neq n} \mathbf{a}_k . \exp(j\omega_{1k} t) + \sum_n \mathbf{a}_n . \exp(j\Omega_n t) \quad (19)$$

$$s_{2p}(t) = \sum_{l \neq n} \mathbf{b}_l . \exp(j\omega_{2l} t) + \sum_n \mathbf{b}_n . \exp(j\Omega_n t) \quad (20)$$

where:

1)Random variables $\mathbf{a}_k$ & $\mathbf{b}_l$ corresponding to $\omega_{1k} \neq \omega_{2l}$ are orthogonal.

2)Correlation of predictable signals $s_{1p}(t)$ & $s_{2p}(t)$ arises from correlation of random variables $\mathbf{a}_n$ & $\mathbf{b}_n$ corresponding to $\{\Omega_n\}$ (common frequency components of source signals).

3)Removing common frequency components of source signals from $s_{1p}(t)$ & $s_{2p}(t)$ result in two residue signals, $\tilde{s}_{1p}(t)$ & $\tilde{s}_{2p}(t)$, that are uncorrelated:

$$\tilde{s}_{1p}(t) = \sum_{k \neq n} \mathbf{a}_k . \exp(j\omega_{1k} t) \quad (21)$$

$$\tilde{s}_{2p}(t) = \sum_{l \neq n} \mathbf{b}_l . \exp(j\omega_{2l} t) \quad (22)$$

Hence,(for i=1,2)

$$x_{ip}(t) = [\alpha_i \tilde{s}_{1p}(t) + \beta_i \tilde{s}_{2p}(t)] + \sum_n (\alpha_i \mathbf{a}_n + \beta_i \mathbf{b}_n) . \exp(j\Omega_n t)$$
$$= \tilde{x}_{ip}(t) + \sum_n (\alpha_i \mathbf{a}_n + \beta_i \mathbf{b}_n) . \exp(j\Omega_n t) \quad (23)$$

and relation (18) can be rewritten as:

$$P^x_{ijp}(\omega) = \sum_{q \neq n} 2\pi . E\{\mathbf{d}_{iq} \mathbf{d}_{jq}^*\} . \delta(\omega - \omega_q) + \sum_n 2\pi . E\{\mathbf{d}_{in} \mathbf{d}_{jn}^*\} . \delta(\Omega - \Omega_n) \quad (24)$$

Removing the terms corresponding to common frequency components (see Appendix-A):

$$\tilde{P}^x_{ijp}(\omega) = \sum_{q \neq n} 2\pi . E\{\mathbf{d}_{iq} \mathbf{d}_{jq}^*\} . \delta(\omega - \omega_q) \quad (25)$$

from which the desired correlation functions are obtained: (for i = 1,2)

$$\tilde{r}^x_{ijp}(\tau) = F^{-1}\{\tilde{P}^x_{ijp}(\omega)\} = E\{\tilde{x}_{ip}(t+\tau) \tilde{x}^*_{jp}(t)\} \quad (26)$$
$$= E\{(\alpha_i \tilde{s}_{1p}(t) + \beta_i \tilde{s}_{2p}(t)) . (\alpha_j \tilde{s}_{1p}(t+\tau) + \beta_j \tilde{s}_{2p}(t+\tau))^*\}$$

At last, because $\tilde{s}_{1p}(t)$ & $\tilde{s}_{2p}(t)$, are uncorrelated, we get the following matrix form:

$$\tilde{\mathbf{R}}^x_p(\tau) \overset{\Delta}{=} \begin{bmatrix} \tilde{r}^x_{11p}(\tau) & \tilde{r}^x_{12p}(\tau) \\ \tilde{r}^x_{21p}(\tau) & \tilde{r}^x_{22p}(\tau) \end{bmatrix}$$
$$= \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} . \begin{bmatrix} \tilde{r}^s_{11p}(\tau) & 0 \\ 0 & \tilde{r}^s_{22p}(\tau) \end{bmatrix} . \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix}^*$$
$$= \mathbf{A} . \tilde{\mathbf{R}}^s_p(\tau) . \mathbf{A}^H \quad (27)$$

where H denotes complex conjugate transpose, and

$$\tilde{r}^s_{iip}(\tau) = E\{\tilde{s}_{ip}(t+\tau) \tilde{s}^*_{ip}(t)\} \quad ; \ i = 1,2 \quad (28)$$

It is seen that in (27) the matrix which is related to source signals is diagonal (a desired condition). This representation is the basis of an algorithm for estimating mixing matrix $\mathbf{A}$.

4. BLIND SOURCE SEPARATION ALGORITHM

In this section, an algorithm for estimating $\mathbf{A}$ (and recovering source signals) is used which is based on the model embedded in eq.(23) and restated in (29) using second order statistics.

$$\tilde{x}_{ip}(t) = \alpha_i \tilde{s}_{1p}(t) + \beta_i \tilde{s}_{2p}(t) \quad ; \ i = 1,2 \quad (29)$$

Steps of the algorithm are following:

4.1. Orthogonalization

Although $\tilde{s}_{1p}(t)$ & $\tilde{s}_{2p}(t)$ are uncorrelated and based on the discussion in section 2, we can assume that:

$$\tilde{\mathbf{R}}^s_p(0) = \mathbf{I} \quad (30)$$

according to eq. (29), $\tilde{x}_{1p}(t)$ & $\tilde{x}_{2p}(t)$ are correlated. Hence, we apply an orthogonalization transformation on $\tilde{x}_{1p}(t)$ & $\tilde{x}_{2p}(t)$. The orthogonalizer matrix is obtained from eigendecomposition of matrix $\tilde{\mathbf{R}}^x_p(\tau)$ at $\tau = 0$. If eigenvalues of $\tilde{\mathbf{R}}^x_p(0)$ are denoted by μ_1 & μ_2 and $\mathbf{v_1}$ & $\mathbf{v}_2$ are the corresponding eigenvectors, the orthogonalization matrix $\mathbf{T}$, defined by:

$$\mathbf{T} = [\frac{1}{\sqrt{\mu_1}} \mathbf{v_1}, \frac{1}{\sqrt{\mu_2}} \mathbf{v}_2]^H \quad (31)$$

satisfies:

$$\mathbf{T} . \tilde{\mathbf{R}}^x_p(0) . \mathbf{T}^H = \mathbf{I} \quad (32)$$

Also from (27),(30),(32), it is seen that :

$$\mathbf{T} . \mathbf{A} . \tilde{\mathbf{R}}^s_p(0) . \mathbf{A}^H . \mathbf{T}^H = \mathbf{T} . \mathbf{A} . \mathbf{A}^H . \mathbf{T}^H = \mathbf{I} \quad (33)$$

This equation shows that matrix $\mathbf{U} = \mathbf{T}.\mathbf{A}$, is a unitary matrix. As a consequence mixing matrix $\mathbf{A}$ can be factored as:

$$\mathbf{A} = \mathbf{T}^{-1}.\mathbf{U} \qquad (34)$$

4.2. Estimation of $\mathbf{U}$

By applying orthogonalization matrix $\mathbf{T}$ to equation (27) for some $\tau \neq 0$,

$$\overline{\mathbf{R}}_{\mathrm{p}}^{\mathrm{x}}(\tau) \overset{\Delta}{=} \mathbf{T}.\mathbf{A}.\widetilde{\mathbf{R}}_{\mathrm{p}}^{\mathrm{s}}(\tau).\mathbf{A}^{H}.\mathbf{T}^{H} \qquad (35)$$

Hence,

$$\overline{\mathbf{R}}_{\mathrm{p}}^{\mathrm{x}}(\tau) = \mathbf{U}.\widetilde{\mathbf{R}}_{\mathrm{p}}^{\mathrm{s}}(\tau).\mathbf{U}^{H} \qquad (36)$$

where matrix $\overline{\mathbf{R}}_{\mathrm{p}}^{\mathrm{x}}(\tau)$ is called orthogonal correlation matrix.

Since $\mathbf{U}$ is unitary and $\widetilde{\mathbf{R}}_{\mathrm{p}}^{\mathrm{s}}(\tau)$ is diagonal, equation (36) states that orthogonal correlation matrix $\overline{\mathbf{R}}_{\mathrm{p}}^{\mathrm{x}}(\tau)$ is diagonalized by the unitary transformation $\mathbf{U}$ (unitary diagonalization). In other words unitary matrix $\mathbf{U}$ can be specified by unitary diagonalizing of orthogonal correlation matrix $\overline{\mathbf{R}}_{\mathrm{p}}^{\mathrm{x}}(\tau)$ for some lag $\tau \neq 0$.

The essential aim is finding a unique unitary matrix $\mathbf{U}$ that diagonalizes $\overline{\mathbf{R}}_{\mathrm{p}}^{\mathrm{x}}(\tau)$ for all time lags τ. A method for attaining this aim, as used in most BSS approaches that exploit statistical properties, is joint diagonalization(JD) method which operates as simultaneous diagonalization of the set $\{\overline{\mathbf{R}}_{\mathrm{p}}^{\mathrm{x}}(\tau_{\mathrm{i}}) \mid \mathrm{i} = 1,2,...,\mathrm{K}\}$ of K orthogonal correlation matrices and is described in the following theorem[3]: (In our case n=2)

THEOREM. Let $\tau_1, \tau_2, ..., \tau_{\mathrm{K}}$ be K nonzero time lags, and let $\mathbf{V}$ be a unitary matrix such that :

$$\forall 1 \le \mathrm{k} \le \mathrm{K} \quad \mathbf{V}^{H}.\overline{\mathbf{R}}_{\mathrm{p}}^{\mathrm{x}}(\tau).\mathbf{V} = \mathrm{diag}.[\mathrm{d}_1(\mathrm{k}), ..., \mathrm{d}_{\mathrm{n}}(\mathrm{k})] \qquad (37)$$

$$\forall 1 \le \mathrm{i} \ne \mathrm{j} \le \mathrm{n} \quad \exists \mathrm{k}, \ni \mathrm{d}_{\mathrm{i}}(\mathrm{k}) \ne \mathrm{d}_{\mathrm{j}}(\mathrm{k}) \qquad (38)$$

then:

- $\mathbf{V}$ is essentially equal to $\mathbf{U}$ (desired unique unitary matrix)
- A permutation can be operated on diagonal elements of $\mathrm{diag}.[\mathrm{d}_1(\mathrm{k}), ..., \mathrm{d}_{\mathrm{n}}(\mathrm{k})]$ □

which states that a unique unitary matrix $\mathbf{U}$ can be determined if for at least a $\tau \neq 0$,eigenvalues of $\overline{\mathbf{R}}_{\mathrm{p}}^{\mathrm{x}}(\tau)$ are distinct, a condition that is surely satisfied for sources with different spectra.

4.3. Computing $\mathbf{A}$ and $\mathbf{S}(t)$

After determination of a unique unitary matrix $\mathbf{U}$, $\mathbf{A}$ can be computed from $\mathbf{A} = \mathbf{T}^{-1}.\mathbf{U}$, and consequently the source signals are estimated as

$$\mathbf{s}(t) = \mathbf{A}^{-1}.\mathbf{x}(t) \qquad (39)$$

It is important to note that for computing $\mathbf{s}(t)$, we use observation data $\mathbf{x}(t)$ (not $\widetilde{\mathbf{x}}(t)$), so there isn't any information loss.

5. SIMULATION RESULTS

In this section, the performance of the proposed method is investigated via computer simulation results.

Source signals $(s_{\mathrm{i}}(t)\ ;\ \mathrm{i} = 1,2)$ are composed of regular and predictable parts. Regular component in each source consists of a zero-mean normal process (independent from the other) and a uniform process (common between two sources). Predictable parts consist of random amplitudes sinusoidal functions with some common frequency and correlated amplitude components. These signals are mixed by an arbitrary 2×2 mixing matrix $\mathbf{A}$, and corrupted by AWGN, to obtain the observation signals $(x_{\mathrm{i}}(t)\ ;\ \mathrm{i} = 1,2)$. Then algorithm is applied on observed data and the estimation of $\mathbf{A}$, $\hat{\mathbf{A}}$, is obtained.
This procedure is repeated for G=500 independent trials.

For evaluate the approach, the following performance index (PI) is introduced,

$$\mathrm{PI} = 10.\mathit{log}_{10}\left[\frac{1}{\mathrm{G}}\sum_{\mathrm{g}=1}^{\mathrm{G}}\left\|\hat{\mathbf{A}}^{-1}.\mathbf{A} - \mathbf{I}\right\|_{F}^{2}\right] \qquad (40)$$

where $\|.\|_F$ is the Frobenius norm.

Two experiments we performed and compared: without pre-separation process (experiment #1) and with pre-separation process (experiment #2). The experiments were executed under noise free and SNR=3,5,8,10 (dB) conditions for various number of correlation matrices used in JD algorithm, and with different correlation coefficient of original source signals. Some results are illustrated in Figures 1-6.
Figures 1 & 2 show performance index(PI)(in dB) versus the number of jointly diagonalized correlation matrices for experiments #1 and #2. In figures 3 & 4, corresponding to experiment #1 and #2, the performance indexes (in dB) versus several SNR (in dB) have been plotted for some constant number of jointly diagonalized correlation matrices with correlation coefficient 0.5. In figures 5 & 6, SNR and the number of jointly diagonalized correlation matrices have been kept constant, and correlation coefficient of source signals has been varied from 0.1 to 0.9.

Almost in all figures, better performance of proposed algorithm is evident. In figures 1 & 2, it is obvious that the

performances of two experiments become better as the number of the jointly diagonalized correlation matrices is increased, but this has a limit (as it is seen from small different between PI for K=5 and K=6). Figures 3 & 4 , show improvement in performance by increasing SNR. From figures 5 & 6, it is revealed that the performance index of two experiments is close for small correlation coefficients, and as the correlation coefficient is increased, the PI of two experiments is decreased, but the performance in experiment #2 is better than that in experiment #1, particularly for intermediate correlation coefficients.

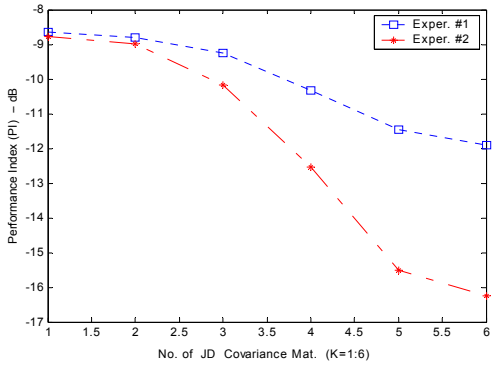

Fig.1. Performance versus number of JD covariance matrices:[Noise Free]&[Correlation Coef.=0.5]

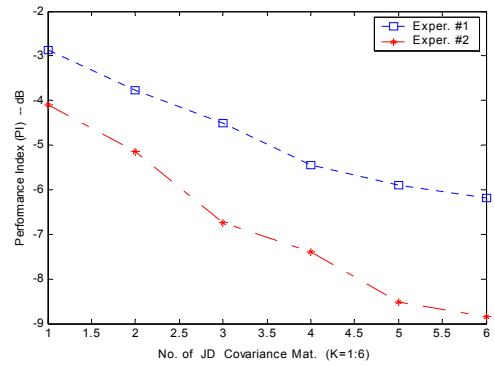

Fig.2. Performance versus number of JD covariance matrices:[SNR=3 dB]&[Correlation Coef.=0.5]

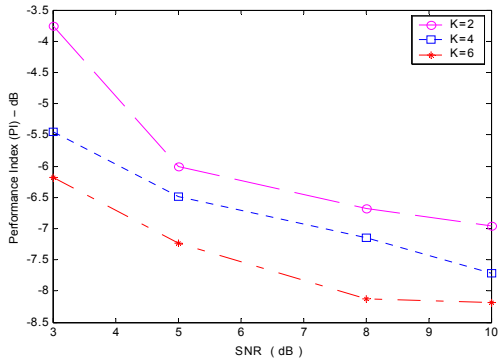

Fig. 3. Performance versus SNR for Experiment #1: [K(No. of JD Covariance Mat.)=2,4,6]&[Correlation coef.=0.5]

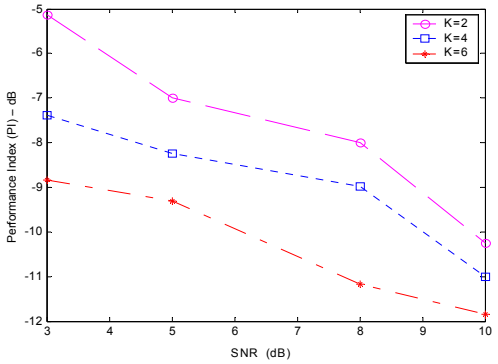

Fig. 4. Performance versus SNR for Experiment #2: [K(No. of JD Covariance Mat.)=2,4,6]&[Correlation Coef.=0.5]

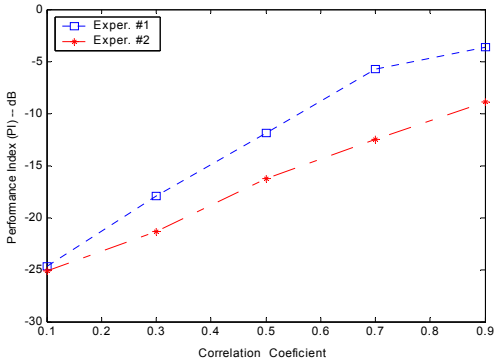

Fig. 5. Performance versus Correlation Coefficient: [Noise Free]&[K(No. of JD Covariance Mat.)=6]

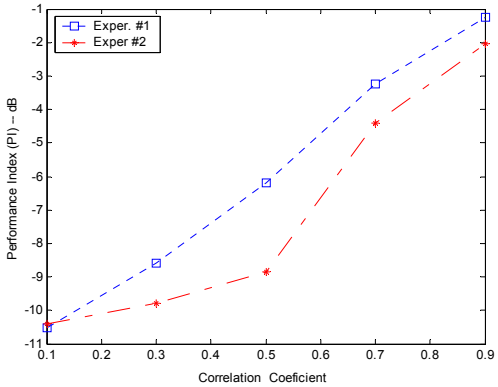

Fig. 6. Performance versus Correlation Coefficient: [SNR=3 dB]&[K(No. of JD Covariance Mat.)=6]

6. CONCLUTION

In this paper, an approach (called CS-SOBI) for solving BSS problem in cases where source signals are correlated, is introduced without additional assumptions on signal or mixing matrix structures.

An important step of this BSS algorithm is a pre-separation procedure where based on Wold decomposition principle, the information of predictable part of source

signals (i.e. uncorrelated parts of predictable signals) is derived. The diagonal structure of the correlation matrix of this parts is essential for next step of algorithm where using second-order based method and JD technique, separation process is completed by estimating $\hat{\mathbf{A}}$ and recovering $\hat{\mathbf{s}}(t)$. Simulation results show effectiveness of algorithm.

APPENDIX-A. A METHOD FOR SEPARATION OF COMMON AND UNCOMMON FREQUENCIES

For identifying the terms corresponding to common frequency components from the terms corresponding to uncommon frequency components in $P_{\mathrm{ijp}}^{\mathrm{x}}(\omega)$, following technique has been used:

Using (21)-(23) relation (24) can be rewritten as the following extended relation for the self and cross spectral density functions of the observations:

$$\begin{aligned} P_{\mathrm{ijp}}^{\mathrm{x}}(\omega) = & \sum_{\mathrm{k}} 2\pi.\alpha_i\alpha_j.\mathrm{E}\{\mathbf{a}_{\mathrm{k}}\mathbf{a}_{\mathrm{k}}^{*}\}.\delta(\omega-\omega_{1\mathrm{k}}) + \\ & \sum_{l} 2\pi.\beta_i\beta_j.\mathrm{E}\{\mathbf{b}_l\mathbf{b}_l^{*}\}.\delta(\omega-\omega_{2l}) + \\ & [\sum_{\mathrm{n}} 2\pi.\alpha_i\alpha_j.\mathrm{E}\{\mathbf{a}_{\mathrm{n}}\mathbf{a}_{\mathrm{n}}^{*}\} + \\ & \sum_{\mathrm{n}} 2\pi.\beta_i\beta_j.\mathrm{E}\{\mathbf{b}_{\mathrm{n}}\mathbf{b}_{\mathrm{n}}^{*}\} + \\ & \sum_{\mathrm{n}} 2\pi.\alpha_i\beta_j.\mathrm{E}\{\mathbf{a}_{\mathrm{n}}\mathbf{b}_{\mathrm{n}}^{*}\} + \\ & \sum_{\mathrm{n}} 2\pi.\alpha_j\beta_i.\mathrm{E}\{\mathbf{b}_{\mathrm{n}}\mathbf{a}_{\mathrm{n}}^{*}\}\,\delta(\Omega-\Omega_{\mathrm{n}})\,] \end{aligned} \quad (41)$$

Here we define two ratios at a specified frequency $\omega=\omega_0$:

$$\lambda_1(\omega_0) = \frac{P_{11\mathrm{p}}^{\mathrm{x}}(\omega_0)}{P_{12\mathrm{p}}^{\mathrm{x}}(\omega_0)} \quad (42)$$

$$\lambda_2(\omega_0) = \frac{P_{21\mathrm{p}}^{\mathrm{x}}(\omega_0)}{P_{22\mathrm{p}}^{\mathrm{x}}(\omega_0)} \quad (43)$$

If ω_0 be an uncommon frequency i.e. $\omega_{1\mathrm{k}}$ or ω_{2l}, then it is seen that $\lambda_1(\omega_0)$ and $\lambda_2(\omega_0)$ are equal:

But if ω_0 is a common frequency i.e. Ω_{n}, then $\lambda_1(\omega_0)$ and $\lambda_2(\omega_0)$ will not be equal:

Fact: If $\omega_0=\Omega_{\mathrm{n}}$ *in eq.(42) and eq.(43), then* $\lambda_1(\omega_0)\neq\lambda_2(\omega_0)$.

Proof: we use the contradiction concept. We assume $\lambda_1(\omega_0)=\lambda_2(\omega_0)$ ($\omega_0=\Omega_{\mathrm{n}}$), then we will have following equality:

$(\alpha_1\beta_2-\alpha_2\beta_1)^2.(\mathrm{E}\{\mathbf{a}_{\mathrm{n}}\mathbf{a}_{\mathrm{n}}^{*}\}.\mathrm{E}\{\mathbf{b}_{\mathrm{n}}\mathbf{b}_{\mathrm{n}}^{*}\}-\mathrm{E}\{\mathbf{a}_{\mathrm{n}}\mathbf{b}_{\mathrm{n}}^{*}\}.\mathrm{E}\{\mathbf{b}_{\mathrm{n}}\mathbf{a}_{\mathrm{n}}^{*}\})=0$

and this results that:

$$\alpha_1\beta_2-\alpha_2\beta_1=0 \quad (44)$$

But mixing matrix is $\mathbf{A}=\begin{bmatrix}\alpha_1 & \beta_1\\ \alpha_2 & \beta_2\end{bmatrix}$, therefore eq. (44) shows that $|\mathbf{A}|=0$, and this result has contradiction with assumption A3. So $\lambda_1(\omega_0)\neq\lambda_2(\omega_0)$.

Therefore, using relations (42) & (43) a criterion for distinguishing uncommon frequencies from common frequencies is:

$$\begin{cases} \text{If } \lambda_1(\omega_0)=\lambda_2(\omega_0) \Rightarrow \omega_0 \text{ is an uncommonfreq.} \\ \text{If } \lambda_1(\omega_0)\neq\lambda_2(\omega_0) \Rightarrow \omega_0 \text{ is a commonfreq.} \end{cases}$$

7. REFERENCES

[1] C. Jutten and J. Herault, "Blind Source Separation I. An Adaptive Algorithm Based on Neuromimetic Architecture", Signal Processing, vol. 24, no. 1, pp. 1-10, July 1991.

[2] J.-F. Cardoso, "Blind Signal Separation: Statistical Principles", Proceedings of the IEEE, vol. 86, no. 10, pp. 2009-2025, Oct. 1998.

[3] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso and E. Moulines, "A Blind Source Separation Technique Using Second-Order Statistics", IEEE Trans. on Signal Processing, vol. 45, no. 2, pp. 434-444, Feb. 1997.

[4] A.-J. Van Der Veen, "Algebraic Methods for Deterministic Blind Beamforming", Proceedings of the IEEE, vol. 86, no. 10, pp. 1987-2008. Oct. 1998.

[5] L. Tong, R.-W. Liu, V.C. Soon and Y.-F. Huang, "Indeterminacy and Identifiability of Blind Identification", IEEE Trans. Circuits and Systems, vol. 38, no. 5, pp. 499-509, May 1991.

[6] A. Papoulis, Probability, Random Variables, and Statistic process, McGraw-Hill, Third Ed., 1991.