Blind Separation of Correlated ECG Signals Using Wold Decomposition

Masoud R. Aghabozorgi, Seyed M.T. Almodarresi

Abstract— **The extraction of the foetal electrocardiogram (FECG) from skin electrode signals recorded from mother body is a problem of concern to signal processing, and blind signal processing (BSS) is a fundamental method for solving this problem. Most proposed BSS techniques for separation of foetal electrocardiogram (FECG) and mother electrocardiogram (MECG) rely on independence of these signals. This paper introduces a novel technique for cases that signals are correlated with each other, i.e. considering a real assumption. The method uses Wold decomposition principle for extracting desired and proper information from the predictable part of the measured data, and exploits approaches based on second-order statistics to estimate the mixing matrix and source signals. Simulation results are provided to illustrate the effectiveness of the method.**

I. Introduction

Foetal electrocardiogram (FECG) extraction is an interesting problem in biomedical engineering which arises when the foetus's heart condition is desired to be monitored. It usually involves measurements from electrodes (sensors) attached to different points of the mother's skin. These electrodes pick up mixtures of foetal electrocardiogram (FECG) and mother electrocardiogram (MECG). The separation of FECG and MECG from these mixtures may be modeled as a blind source separation (BSS) problem [1],[2].

Blind source separation consists of recovering signals from several measured noisy mixtures of them. The problem is called "blind" because no information is available about the mixture, i.e. recovering of source signals is achieved without the knowledge of the characteristics of the transmission channel (mother's body). 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 strong assumption of independence between the source signals. Therefore the goal in these techniques is to achieve a separation process that produces outputs as independent as possible [3],[4]. 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 [5],[6].

In this paper, the aim is to propose a solution to FECG extraction problem using BSS method considering correlated signals which is a real assumption between FECG and MECG signals. This paper is organized as follows: In section 2, the problem of BSS is stated. 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.

Manuscript received June 29, 2006. This work was supported by Yazd University.

Authors are with the Electrical Engineering Department, Yazd University, Yazd, P.O.BOX 89195-741, IRAN (corresponding author to provide phone: +98-351-8211670; e-mail: aghabozorgi@yazduni.ac.ir).

II. BSS Problem Formulation

Assume that d signals $s_1(t),\ldots,s_d(t)$ are transmitted from d sources at different locations. What we receive at m sensors will be instantaneous linear combinations of these signals that construct measured data:

$$\mathbf{x}(t) = \mathbf{a}_1.\mathrm{s}_1(t) + \ldots + \mathbf{a}_\mathrm{d}.\mathrm{s}_\mathrm{d}(t) + \mathbf{n}(t) \qquad (1)$$

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 measured data vector from m sensors, $\mathbf{s}(t) \in \Re^{d\times 1}$ is the signal vector, composed of d unknown source signals, $\mathbf{A} = [\mathbf{a}_1,\ldots,\mathbf{a}_\mathrm{d}] \in \Re^{m\times d}$ 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.

The aim of blind source separation (BSS) is to identify the mixing matrix $\mathbf{A}$ and consequently recovering the source signals from the measurements.

III. Pre-separation Procedure

If main step in our approach for correlated sources is a pre-separation process. The measured 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 are estimated using second order Equation).

A. Wold Decomposition

An arbitrary process can be written as a sum:

$$s(t) = s_\mathrm{r}(t) + s_\mathrm{p}(t) \qquad (4)$$

where $s_\mathrm{r}(t)$ and $s_\mathrm{p}(t)$ are regular and predictable processes. This expansion is called Wold decomposition. In [7] it has been proved that the processes $s_\mathrm{r}(t)$ and $s_\mathrm{p}(t)$ are orthogonal. Furthermore, $s_p(t)$ is comprised of complex exponentials:

$$s_p(t) = \mathbf{c}_0 + \sum_i \mathbf{c}_i.\exp(j\omega_i t) \quad (5)$$

where $\mathbf{c}_i$'s are orthogonal zero-mean random variables. Hence, $s_p(t)$ has a line spectrum:

$$P_{s_p}(\omega) = \sum_i 2\pi\alpha_i\, \delta(\omega - \omega_i) \quad (6)$$

but $s_r(t)$ has a smooth spectrum.

B. Measurements Decomposition

In this subsection a method is proposed for extracting and decomposing some information from the regular and predictable parts of the measured 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 section2:

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \mathbf{A}.\begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \end{bmatrix} \quad (7)$$

where $s_1(t)$ and $s_2(t)$ are the source signals $\mathbf{A}$ is the mixing matrix.

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

$$s_i(t) = s_{ip}(t) + s_{ir}(t) \quad (8)$$

where,

$$s_{1p}(t) = \sum_k \mathbf{a}_k.\exp(j\omega_{1k}t) \quad (9)$$

$$s_{2p}(t) = \sum_l \mathbf{b}_l.\exp(j\omega_{2l}t) \quad (10)$$

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

Using (7),(8) and the fact that regular and predictable parts in each signal are orthogonal ,we get:

$$x_i(t) = x_{ip}(t) + x_{ir}(t) + n_i(t) \quad ; \; i = 1,2 \quad (11)$$

where,

$$x_{ip}(t) = \alpha_i s_{1p}(t) + \beta_i s_{2p}(t) \quad ; \; i = 1,2 \quad (12)$$

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

$$x_{ip}(t) = \sum_q \mathbf{d}_{iq}.\exp(j\omega_q t) \quad ; \; i = 1,2 \quad (13)$$

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

From these equations, each measured 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 measured data are given by: (i,j = 1,2)

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

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

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

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

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

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

where,

$$P_{ijp}^{x}(\omega) = \sum_q 2\pi.\mathrm{E}\{\mathbf{d}_{iq}\mathbf{d}_{jq}^{*}\}\delta(\omega - \omega_q) \quad (18)$$

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

C. Extracting Desired Information from predictable parts

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

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

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

$$P^{x}_{ijp}(\omega)=\sum_{q\neq n}2\pi.\mathrm{E}\{\mathbf{d}_{iq}\mathbf{d}_{jq}{}^{*}\}\delta(\omega-\omega_{q})$$
$$+\sum_{n}2\pi.\mathrm{E}\{\mathbf{d}_{in}\mathbf{d}_{jn}{}^{*}\}\delta(\omega-\Omega_{n}) \quad (24)$$

Removing the terms corresponding to common frequency components results:

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

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

$$\widetilde{\mathrm{r}}^{x}_{ijp}(\tau)=F^{-1}\{\widetilde{P}^{x}_{ijp}(\omega)\}=\mathrm{E}\{\widetilde{x}_{ip}(t+\tau)\widetilde{x}^{*}_{jp}(t)\} \; (26)$$
$$=\mathrm{E}\{(\alpha_{i}\widetilde{s}_{1p}(t)+\beta_{i}\,\widetilde{s}_{2p}(t)).(\alpha_{j}\widetilde{s}_{1p}(t+\tau)+\beta_{j}\,\widetilde{s}_{2p}(t+\tau))^{*}\}$$

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

$$\widetilde{\mathbf{R}}^{x}_{p}(\tau)\overset{\Delta}{=}\begin{bmatrix}\widetilde{\mathrm{r}}^{x}_{11p}(\tau) & \widetilde{\mathrm{r}}^{x}_{12p}(\tau)\\ \widetilde{\mathrm{r}}^{x}_{21p}(\tau) & \widetilde{\mathrm{r}}^{x}_{22p}(\tau)\end{bmatrix}$$

$$=\begin{bmatrix}\alpha_{1} & \beta_{1}\\ \alpha_{2} & \beta_{2}\end{bmatrix}.\begin{bmatrix}\widetilde{\mathrm{r}}^{s}_{11p}(\tau) & 0\\ 0 & \widetilde{\mathrm{r}}^{s}_{22p}(\tau)\end{bmatrix}.\begin{bmatrix}\alpha_{1} & \alpha_{2}\\ \beta_{1} & \beta_{2}\end{bmatrix}^{*}$$
$$=\mathbf{A}.\widetilde{\mathbf{R}}^{s}_{p}(\tau).\mathbf{A}^{H} \quad (27)$$

where H denotes complex conjugate transpose, and

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

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 $\mathbf{A}$.

IV. Blind Source Separation Algorithm

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

$$\widetilde{x}_{ip}(t)=\alpha_{i}\widetilde{s}_{1p}(t)+\beta_{i}\,\widetilde{s}_{2p}(t) \quad ; i=1,2 \quad (29)$$

Steps of the algorithm are following:

A. *Ortogonalization*

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

$$\widetilde{\mathbf{R}}^{s}_{p}(0)=\mathbf{I} \quad (30)$$

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

$$\mathbf{T}=[\frac{1}{\lambda_{1}}\mathbf{v_{1}}\,,\;\frac{1}{\lambda_{2}}\mathbf{v}_{2}]^{H} \quad (31)$$

satisfies:

$$\mathbf{T}.\widetilde{\mathbf{R}}^{x}_{p}(0).\mathbf{T}^{H}=\mathbf{I} \quad (32)$$

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

$$\mathbf{T}.\mathbf{A}.\widetilde{\mathbf{R}}^{s}_{p}(0).\mathbf{A}^{H}.\mathbf{T}^{H}=\mathbf{T}.\mathbf{A}.\mathbf{A}^{H}.\mathbf{T}^{H}=\mathbf{I} \quad (33)$$

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:

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

B. *Estimation of* $\mathbf{U}$

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

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

Hence,

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

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

Since $\mathbf{U}$ is unitary and $\widetilde{\mathbf{R}}^{s}_{p}(\tau)$ is diagonal, equation (36) states that orthogonal correlation matrix $\overline{\mathbf{R}}^{x}_{p}(\tau)$ 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 $\overline{\mathbf{R}}^{x}_{p}(\tau)$ for some lag $\tau\neq 0$.

C. *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

$$\mathbf{s}(t)=\mathbf{A}^{-1}.\mathbf{x}(t) \quad (37)$$

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

V. Simulation Results

In this section, the performance of the proposed method is investigated via computer simulation results. The data (mecg and fecg signals) are from [8].

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

$$\mathrm{PI} = 10.\, log_{10}\, [\frac{1}{\mathrm{G}} \sum_{\mathrm{g}=1}^{\mathrm{G}} \left\| \hat{\mathbf{A}}^{-1}.\mathbf{A} - \mathbf{I} \right\|_F^2] \qquad (38)$$

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

Two experiments we performed and compared: without pre-separation process (experiment #1) and with proposed 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. Some results are illustrated in the following Figures.

Almost in all figures, better performance of proposed algorithm (experiment #2) is evident. In figures 3 and 4, it is obvious that the performances of two experiments become better as the number of the jointly diagonalized correlation matrices is increased. Figures 5 and 6, show improvement in performance by increasing SNR.

VI. Conclusion

In this paper, an approach for separation of MECG and FECG applying BSS method in cases where desired 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.

References

[1] L. De Lathauwer, B. De Moor and J. Vandewalle, “Fetal Electrocardiogram Extraction by Blind Source Subspace Separation”, *IEEE Transactions on biomedical engineering*, pp. 567-572, 2001.

[2] F. Vrins, V. Vigneron, C. Jutten and M. Verleysen, “Abdominal Electrodes Analysis by Statistical Processing for Fetal Electrocardiogram Extraction” *Proceedings of second International Conferenceon Biomedical Engineering*, Innsbruck, Austria, pp. 244-249, Feb. 2004.

[3] S. Choi, A. cichocki, H.M. Park and S.Y. Lee, “Blind Source Separation and Independent Component Analysis: A Review”, *Neural Information Processing*, Vol. 6, No. 1, pp. 1-57, Jan. 2005.

[4] K. Abed-Meraim and et. al., “Blind Source Separation Using Second Order Cyclostationary Statistics”, *IEEE Trans. on Signal Processing*, Vol. 49, No. 4, pp. 694-701, April 2001.

[5] 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.

[6] E. Moreau, “A Generalization of Joint Diagonalization Criteria for Source Separation”, *IEEE Trans. on Signal Processing*, Vol. 49, No. 3, pp. 530-541, March 2001.

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

[8] http://www.physionet.org/

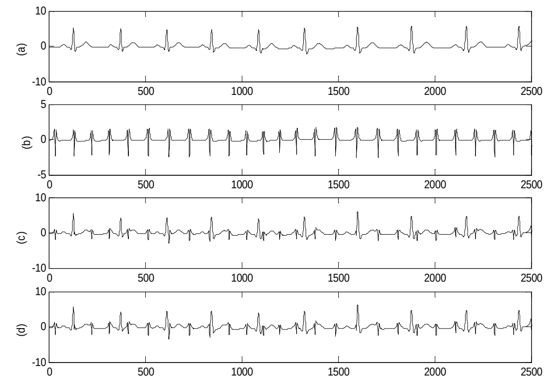

Fig. 1. Signal Samples: (a) mecg, (b) fecg, (c) mixed signal: x1 (c) mixed signal: x2

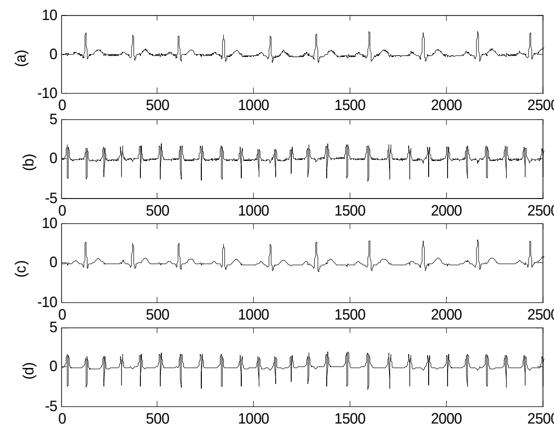

Fig. 2. Signal Samples: (a) extracted signal s1from Experiment #1, (b) extracted signal s2 from Experiment #1, (c) extracted signal s1 from Experiment #2, (d) extracted signal s2 from Experiment#2

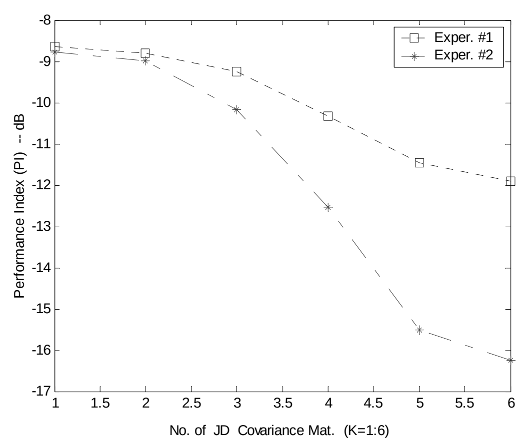

Fig.3. Performance versus number of JD correlation matrices: [Noise Free]

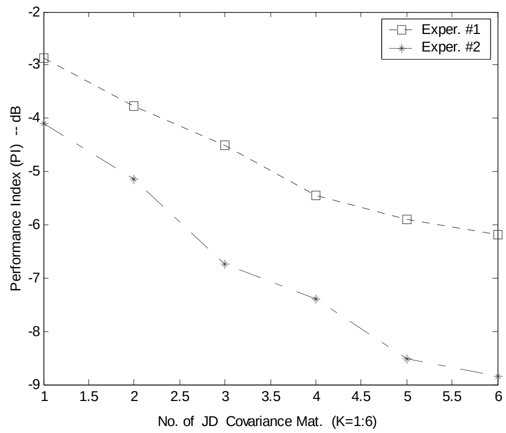

Fig.4. Performance versus number of JD correlation matrices:[SNR=3 dB]

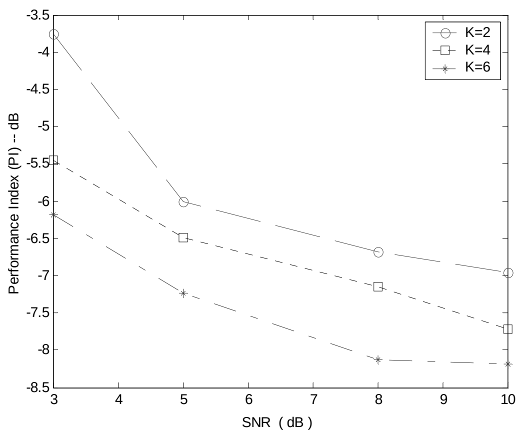

Fig. 5. Performance versus SNR for Experiment #1: [K(No. of JD correlation Mat.)=2,4,6]

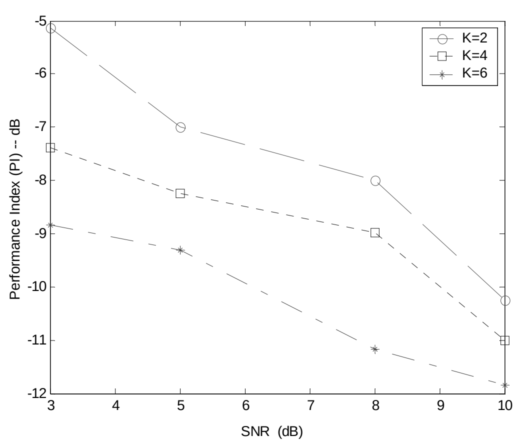

Fig. 6. Performance versus SNR for Experiment #2: [K(No. of JD correlation Mat.)=2,4,6]