Blind Source Separation Using Second-Order Cyclostationary Statistics

Karim Abed-Meraim, Yong Xiang, Jonathan H. Manton, and Yingbo Hua, *Senior Member, IEEE*

***Abstract*—This paper studies the blind source separation (BSS) problem with the assumption that the source signals are cyclostationary. Identifiability and separability criteria based on second-order cyclostationary statistics (SOCS) alone are derived. The identifiability condition is used to define an appropriate contrast function. An iterative algorithm (ATH2) is derived to minimize this contrast function. This algorithm separates the sources even when they do not have distinct cycle frequencies.**

I. Introduction

BLIND source separation has recently become an intense research topic in many applications such as remote sensing, speech processing, medical diagnosis, and wireless communications. It is motivated by practical scenarios which involve multisources and multisensors. A basic model for BSS is that of m statistically independent signals whose n (possibly noisy) linear combinations are observed. Given these observations, BSS aims to estimate both the structure of the linear combinations and the source signals. For BSS to be possible, something extra must be known about the source signals. In this paper, the extra assumption is that the source signals are *cyclostationary* [1]. This assumption is reasonable since many man-made signals encountered in communications, telemetry, radar, and sonar systems are cyclostationary. Other papers that perform BSS based on this assumption include [2]–[5].

This paper restricts its attention to methods based on second-order cyclostationary statistics (SOCS). It derives necessary and sufficient conditions for successful BSS based on SOCS alone. Iterative and noniterative optimization algorithms for BSS are derived. Simulation results illustrate the validity of these methods.

This work can be seen as a new contribution above and beyond the contributions in [4]. More precisely and comparatively to [4], we can say the following.

- Whereas [4] considers only a scalar signal, our paper considers a vector of signals. This is a significant extension.
- Indeed, when considering a vector of signals, it is necessary to perform source separation. Necessary and sufficient conditions for being able to separate the sources are given in our paper. These conditions cannot be derived from [4].
- Our paper presents methods for 1) recovering a single source signal of interest and 2) simultaneously recovering all signals. Since [4] considers only the scalar signal case, it is not possible to derive a method for [4, Eq. (2)].
- Moreover, different "cost functions" are used in the two papers. Roughly speaking, our paper finds the matrix that whitens the cyclic correlations at various lags, whereas [4] finds the vector that minimizes the least square error between the signal and its cyclically shifted version.

The rest of this paper is organized as follows. Section II indicates some definitions and introduces the problem of BSS together with relevant hypothesis. In Section III, a necessary and sufficient condition for BSS using a set of cyclic correlation coefficients is given. Under this condition, two separation criteria are introduced: first in the case of sources with distinct cycle frequencies and then in the general case of sources sharing same cycle frequencies. In Section IV, we consider the case where only one or few signals are of interest. Condition for partial identifiability, separation criteria, and new noniterative separation algorithms are given. In Section V, new iterative (and possibly adaptive) optimization algorithms for BSS using SOCS are presented. These algorithms minimize a certain contrast function derived from the separation criteria of Section III using the technique of natural gradient [10]. Simulation results and concluding remarks are given in Sections VI and VII, respectively.

II. Problem Formulation

Assume that m source signals impinge on an array of n sensors, where $n \geq m$. The output of each sensor is modeled as a weighted sum of the source signals corrupted by additive noise. This can be expressed in vector form as follows:

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{w}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{w}(t).$$

Here, $\mathbf{s}(t) = [s_1(t), \cdots, s_m(t)]^T$ is the $m \times 1$ complex *source vector*, $\mathbf{w}(t) = [w_1(t), \cdots, w_n(t)]^T$ is the $n \times 1$ complex *noise vector*, $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_m]$ is the unknown $n \times m$ full column rank *mixing matrix*, and the superscript T denotes the transpose of a vector. The source signal vector $\mathbf{s}(t)$ is modeled as a cyclostationary complex stochastic process. The component processes $s_i(t)$, $1 \leq i \leq m$ are assumed to be mutually independent with zero mean. In particular, we assume that

$$\left\langle e^{\mathrm{J}\beta_i t} s_i(t+\tau) s_j^*(t) \right\rangle = 0, \qquad \text{if } i \neq j \tag{1}$$

$$\left\langle e^{\mathrm{J}\beta_j t} s_i(t+\tau) s_i^*(t) \right\rangle = 0, \qquad \text{if } \beta_i \neq \beta_j \tag{2}$$

$$\left\langle e^{\mathrm{J}\beta_i t} s_i(t) s_i^*(t) \right\rangle > 0, \qquad \forall\, i. \tag{3}$$

Manuscript received January 19, 1999; revised December 27, 2000. This work was supported by the Australian Research Council.The associate editor coordinating the review of this paper and approving it for publication was Prof. Bjorn Ottersten.

K. Abed-Meraim is with the Signal and Image Processing Department, Ecole Nationale Supérieure des Télécommunications, Paris, France (e-mail: abed@tsi.enst.fr).

Y. Xiang is with Bandspeed Pty. Ltd., Melbourne, Australia.

J. H. Manton is with the Department of Electrical Engineering, University of Melbourne, Parkville, Australia (e-mail: jon@ee.mu.oz.au).

Y. Hua is with the University of California, Riverside, CA 92502 USA.

Publisher Item Identifier S 1053-587X(01)02243-7.

Here, $J = \sqrt{-1}$, and $\langle\cdot\rangle$ denotes the time averaging operator (see [1])

$$\left\langle e^{\mathrm{J}\beta_j t} s_i(t+\tau)s_i^*(t)\right\rangle \stackrel{\text{def}}{=} \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1} e^{\mathrm{J}\beta_j t} s_i(t+\tau)s_i^*(t).$$

Furthermore, each β_i is a nonzero cycle frequency of source i. The cyclic autocorrelation function $\rho_i(\tau)$ is defined to be $\rho_i(\tau) \stackrel{\text{def}}{=} \langle s_i(t+\tau)s_i^*(t)e^{\mathrm{J}\beta_i t}\rangle$ and satisfies $\rho_i(0) > 0$. The superscript $*$ denotes complex conjugate, whereas the superscript $\star$ denotes the complex conjugate transpose of a vector. The additive noise $\mathbf{w}(t)$ is modeled as a stationary complex random process so that [1]

$$\left\langle e^{\mathrm{J}\beta_i t} w_i(t+\tau)w_i^*(t)\right\rangle = 0 \quad \forall\, i,\, \tau.$$

The output cyclic correlation matrix $\mathbf{R}_x^{(\beta_i)}(\tau)$ is defined to be

$$\mathbf{R}_x^{(\beta_i)}(\tau) \stackrel{\text{def}}{=} \left\langle e^{\mathrm{J}\beta_i t}\mathbf{x}(t+\tau)\mathbf{x}^\star(t)\right\rangle.$$

Under the above assumptions, the cyclic correlation matrices of the array output take the following structure:

$$\mathbf{R}_x^{(\beta_i)}(\tau) = \sum_{\{j|\beta_j=\beta_i\}} \rho_j(\tau)\mathbf{a}_j\mathbf{a}_j^\star \tag{4}$$

where the sum is over all sources with cycle frequency β_i. In particular, if all sources have distinct cycle frequencies, i.e., $\beta_i \neq \beta_j$ for $i \neq j$, then only source i contributes to $\mathbf{R}_x^{(\beta_i)}(\tau)$ which becomes

$$\mathbf{R}_x^{(\beta_i)}(\tau) = \rho_i(\tau)\mathbf{a}_i\mathbf{a}_i^\star. \tag{5}$$

The aim of blind source separation is to find an $m \times n$ separating matrix $\mathbf{B} = [\mathbf{b}_1, \cdots, \mathbf{b}_n]$ such that $\hat{\mathbf{s}}(t) = \mathbf{B}\mathbf{x}(t)$ is an estimate of the source signals. Note that it is not possible to uniquely identify the separating matrix $\mathbf{B}$ (or, equivalently, the mixing matrix $\mathbf{A}$) because the exchange of a fixed scalar between the source signal and the corresponding column of $\mathbf{A}$ leaves the observations unaffected. We take advantage of this indetermination to assume that the emitter signals have unit-norm zero-lag cyclic autocorrelation coefficients, i.e.,

$$\rho_i(0) = \left\langle e^{\mathrm{J}\beta_i t} s_i(t)s_i^*(t)\right\rangle = 1. \tag{6}$$

In addition, the *numbering* of signals with the same cycle frequency is immaterial. The best that can be done then is to determine $\mathbf{B}$ up to a permutation and scaling of its columns [6]. That is, $\mathbf{B}$ is a separating matrix if

$$\mathbf{B}\mathbf{y}(t) = \mathbf{P}\mathbf{\Lambda}\mathbf{s}(t)$$

where $\mathbf{P}$ is a permutation matrix and $\mathbf{\Lambda}$ a *unitary* diagonal matrix. Note that if all sources have distinct cycle frequencies, then the numbering of signals is possible according to the numbering of the cycle frequencies. In this case, $\mathbf{B}$ is a separating matrix if

$$\mathbf{B}\mathbf{y}(t) = \mathbf{\Lambda}\mathbf{s}(t)$$

for a given unitary diagonal matrix $\mathbf{\Lambda}$.

Remarks:

1) For simplicity, we have adopted here the definition of second order cyclostationarity given in [1]. A more rigorous definition can be used as follows: A zero-mean second order cyclostationary process $s(t)$ is characterized by the property that its time-varying autocorrelation $r_s(t,\,\tau) = E(s(t+\tau)s^*(t))$ varies periodically with respect to time t. Thus, it accepts a Fourier series representation, i.e.,

$$r_s(t,\,\tau) = \sum_{\beta\in C} r_s^\beta(\tau)e^{-j\beta t}$$

$$r_s^\beta(\tau) = \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1} r_s(t,\,\tau)e^{j\beta t}$$

where the Fourier coefficients $r_s^\beta(\tau)$ are called the cyclic autocorrelation at cycle frequency β, and

$$C = \{\beta | 0 \le \beta < 2\pi \quad \text{and} \quad r_s^\beta(\tau) \neq 0\}$$

is called the cycle frequency set of $s(t)$. It is shown in [14] that if $s(t)$ is a mixing process, a consistent and asymptotically normal estimator of $r_s^\beta(\tau)$ is given by (the time averaging operator)

$$\hat{r}_s^\beta(\tau) = \frac{1}{T}\sum_{t=0}^{T-1} s(t+\tau)s^*(t)e^{j\beta t}.$$

2) In this presentation, we have considered one cycle frequency for each source signal. In practice, the sources' energy may be distributed to more than one cycle frequency. In that case, we can replace $\mathbf{R}_x^{(\beta_i)}(\tau)$ by a linear combination of cyclic correlation matrices that adds coherently[1] the energy of the considered source over its different cycle frequencies. Another possibility, is to use several cycle frequencies (i.e., several cyclic correlation matrices) for each source signal. The important point is that such data preprocessing does not affect the algorithm derivations given in the sequel.
3) The mutual independence of the sources expressed by (1) is a fundamental condition for blind source separation. On the other hand, conditions (2) and (3) are not necessary and can be relaxed.

   In fact, condition (2) is only used to select one particular source signal by selecting its corresponding cycle frequency. The case where (2) is not satisfied is equivalent to that where several sources share a same cycle frequency. This case is treated in Theorems 2 and 4 of this paper.

   Condition (3) is only used to constrain the separating matrix output to have nonzero cyclic correlation coefficients. Thus, it can be replaced by the condition

$$\left|\left\langle e^{\mathrm{J}\beta_i t} s_i(t)s_i^*(t)\right\rangle\right| > 0 \quad \forall\, i \tag{7}$$

which is always satisfied if $s_i(t)$ is cyclostationary with cycle frequency β_i. For simplicity, we keep using (3) but

[1]For example, if C_i denotes the cycle frequency set of source i and if $C_i \cap C_j = \emptyset$ for $i \neq j$, then we can coherently combine cyclic correlation matrices according to $\mathbf{R}_x^{(i)}(\tau) = \sum_{\beta\in C_i} \lambda_\beta^* \mathbf{R}_x^{(\beta)}(\tau)$, where λ_β is the largest (nonzero) eigenvalue of $\mathbf{R}_x^{(\beta)}(\tau)$.

will indicate briefly in Sections IV and V how the proposed algorithms are modified if we replace (3) with (7).

III. Condition for Identifiability

This section states and proves a necessary and sufficient condition for blind source separation via second-order cyclostationary statistics of the array output. The definitions and notation in the previous section are used.

Recall that $\rho_i(\tau)$ is the cyclic autocorrelation function of the ith source signal. For a given set of nonzero time lags $\tau_1, \cdots, \tau_K$, the $1 \times (K+1)$ cyclic autocorrelation vector $\boldsymbol{\rho}_i$ is defined to be

$$\boldsymbol{\rho}_i = [\rho_i(0), \rho_i(\tau_1), \cdots, \rho_i(\tau_K)].$$

The following is a necessary and sufficient condition for BSS using only the cyclic correlation matrices $\mathbf{R}_x^{(\beta_i)}(\tau)$, $i = 1, \cdots, m$ at time lags $0, \tau_1, \cdots, \tau_K$.

Identifiability Condition: For any $K \geq 0$, blind source separation can be achieved using the output cyclic correlation matrices $\{\mathbf{R}_x^{(\beta_i)}(\tau) | i = 1, \cdots, m;\ \tau = 0, \tau_1, \cdots, \tau_K\}$ if and only if there do not exist two distinct source signals $s_i(t)$ and $s_j(t)$ whose cycle frequencies are the same ($\beta_i = \beta_j$) and whose cyclic autocorrelation vectors ρ_i and ρ_j are linearly dependent.[2]

The sufficiency of the above condition follows from Theorems 1 and 2 below. A proof of its necessity is now given.

Proof: Without loss of generality, assume that the source vector $\mathbf{s}(t)$ is such that $\beta_1 = \beta_2$ and $\boldsymbol{\rho}_1 = \boldsymbol{\rho}_2$. For any mixing matrix $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_m]$, define another mixing matrix $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \mathbf{a}_3, \cdots, \mathbf{a}_m]$, where

$$[\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2] = [\mathbf{a}_1, \mathbf{a}_2] \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

Similarly, define another source vector $\tilde{\mathbf{s}}(t)$ by $\tilde{\mathbf{s}}(t) = [\tilde{s}_1(t), \tilde{s}_2(t), s_3(t), \cdots, s_m(t)]^T$, where

$$\begin{bmatrix} \tilde{s}_1(t) \\ \tilde{s}_2(t) \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix}.$$

Then, it is readily verified that the output vectors $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{w}(t)$ and $\tilde{\mathbf{x}}(t) = \tilde{\mathbf{A}}\tilde{\mathbf{s}}(t) + \mathbf{w}(t)$ have the same cyclic correlation matrices at time lags $0, \tau_1, \cdots, \tau_K$ as well as the source vectors $\mathbf{s}(t)$ and $\tilde{\mathbf{s}}(t)$. □

The following theorem gives a separation criterion for when the source vector has distinct cycle frequencies.

Theorem 1: Assume that the cycle frequencies of the source signals are distinct. For any matrix $\mathbf{B}$, define $\mathbf{z}(t)$ to be the $m \times 1$ vector given by $\mathbf{z}(t) = \mathbf{B}\mathbf{x}(t)$. In addition, define its cyclic cross-correlation $r_{ij}(\tau) \stackrel{\text{def}}{=} \langle z_i(t+\tau) z_j^*(t) e^{\mathrm{J}\beta_i t} \rangle$. Then, $\mathbf{B}$ is a separating matrix if and only if

$$r_{ij}(0) = 0 \quad \text{and} \quad r_{ii}(0) = 1 \tag{8}$$

for all $1 \leq i \neq j \leq m$.

[2]Note that with the scaling convention (6), $\boldsymbol{\rho}_i$ and $\boldsymbol{\rho}_j$ are linearly dependent iff $\boldsymbol{\rho}_i = \boldsymbol{\rho}_j$.

Proof: Define $\mathbf{C} = \mathbf{B}\mathbf{A} = [c_{ij}]_{1 \leq i,\, j \leq m}$ so that $\mathbf{z}(t) = \mathbf{C}\mathbf{s}(t) + \mathbf{B}\mathbf{w}(t)$. It follows from the mutual independence of the source signals and the stationarity of $\mathbf{w}(t)$ that $r_{ij}(\tau) = c_{ii} c_{ji}^* \rho_i(\tau)$. Because $\mathbf{B}$ is a separating matrix if and only if $\mathbf{C}$ is a unitary diagonal matrix, it is sufficient to show that (8) is equivalent to $\mathbf{C}$ being unitary and diagonal. This readily follows from the fact that $\rho_i(0) = 1$ for all i. In particular, if $\mathbf{C}$ is unitary and diagonal, then $c_{ii} c_{ji}^* = 0$ for all $i \neq j$, and $c_{ii} c_{ji}^* = |c_{ii}|^2 = 1$ if $i = j$. This implies (8). Conversely, if (8) is true, then $|c_{ii}|^2 = 1$, and $c_{ji}^* = 0$ for all $j \neq i$, that is, $\mathbf{C}$ is unitary and diagonal. □

Theorem 2 is a generalization of Theorem 1.

Theorem 2: Assume that the identifiability condition is satisfied, that is, if $\beta_i = \beta_j$, then $\boldsymbol{\rho}_i$, and $\boldsymbol{\rho}_j$ are linearly independent. Then, $\mathbf{B}$ is a separating matrix if and only if

$$r_{ij}(k) = 0 \quad \text{and} \quad r_{ii}(0) = 1 \tag{9}$$

for all $1 \leq i \neq j \leq m$ and $k = 0, \tau_1, \cdots, \tau_K$.

Proof: We extend the proof of Theorem 1. Recall that $\mathbf{C} = \mathbf{B}\mathbf{A}$. Write it as $\mathbf{C} = [\mathbf{c}_1, \cdots, \mathbf{c}_m]$. Consider any source i. Let $i_1, \cdots, i_I$ be all the sources (including i) with the same cyclic frequency as source i. Define $\mathbf{s}_i(t) = [s_{i_1}(t), \cdots, s_{i_I}(t)]^T$ as the corresponding column vectors of $\mathbf{C}$ i.e., $\mathbf{C}_i = [\mathbf{c}_{i_1}, \cdots, \mathbf{c}_{i_I}]$, and

$$\mathbf{S}_i(\tau) = \langle \mathbf{s}_i(t+\tau) \mathbf{s}_i^\star(t) e^{\mathrm{J}\beta_i t} \rangle = \mathrm{diag}(\rho_{i_1}(\tau), \cdots, \rho_{i_I}(\tau))$$

$$\mathbf{R}_z^{\beta_i}(\tau) = \langle \mathbf{z}(t+\tau) \mathbf{z}^\star(t) e^{\mathrm{J}\beta_i t} \rangle = \mathbf{C}_i \mathbf{S}_i(\tau) \mathbf{C}_i^H$$

where the last equality comes from (4). The superscript H denotes the complex conjugate transpose of a matrix. Using these matrix notations, (9) leads to

$$\sum_{\beta_i} \mathbf{R}_z^{\beta_i}(0) = \mathbf{C}\mathbf{C}^H = \mathbf{I}.$$

Otherwise, $\mathbf{C}$ is unitary (and thus, in particular, $\mathbf{C}_i$ is full column rank), and

$$\mathbf{R}_z^{\beta_i}(k) = \mathbf{C}_i \mathbf{S}_i(k) \mathbf{C}_i^H \ \text{ is diagonal for } \ k = \tau_1, \cdots, \tau_K.$$

We can then conclude that $\mathbf{C}_i = \mathbf{P}\boldsymbol{\Lambda}$, where $\mathbf{P}$ is a permutation matrix, and $\boldsymbol{\Lambda} = [\boldsymbol{\Lambda}_i^T, \mathbf{0}^T]^T$, where $\boldsymbol{\Lambda}_i$ is an $I \times I$ unitary diagonal matrix by using [6, Th. 2]. □

Iterative algorithms based on the criteria in Theorems 1 and 2 are derived in Section V.

IV. Conditions for Partial Identifiability

The identifiability condition in the previous section assumes that all the source signals are to be separated. This section generalizes the results of the previous section to when only certain sources are to be separated. Furthermore, explicit formulae are given for determining the separation matrix.

We first assume that the source signals have distinct cycle frequencies. Let $\mathbf{b}_i$ be an $n \times 1$ vector. Analogously to Theorem 1, it can be shown that the scalar random variable $z_i(t) = \mathbf{b}_i^\star \mathbf{x}(t)$ is an estimate of $s_i(t)$ [that is, $\mathbf{b}_i^\star \mathbf{A}\mathbf{s}(t) = \alpha s_i(t)$ for some

unit-norm scalar α] if and only if the following two conditions hold:

$$\left\langle |z_i(t)|^2 \sum_{j\neq i} e^{\mathrm{J}\beta_j t} \right\rangle = 0 \tag{10}$$

$$\left\langle |z_i(t)|^2 e^{\mathrm{J}\beta_i t} \right\rangle = 1. \tag{11}$$

This leads to the following theorem for separating a single source signal when each source has a distinct cycle frequency.

Theorem 3: Define the vector $\mathbf{b}_i = \mathbf{R}^{-H/2}\mathbf{c}_i$, where $\mathbf{R}^{-H/2}$ is the conjugate transpose of $\mathbf{R}^{-1/2}$, and $\mathbf{R}^{-1/2}$ is the pseudo-inverse of an $n \times m$ square root of $\mathbf{R} \stackrel{\text{def}}{=} \langle \mathbf{x}(t)\mathbf{x}^\star(t) \sum_{j=1}^{m} e^{\mathrm{J}\beta_j t} \rangle$. The vector $\mathbf{c}_i$ is the least eigenvector of $\mathbf{R}^{-1/2}\tilde{\mathbf{R}}_i\mathbf{R}^{-H/2}$, where $\tilde{\mathbf{R}}_i \stackrel{\text{def}}{=} \langle \mathbf{x}(t)\mathbf{x}^\star(t) \sum_{j\neq i} e^{\mathrm{J}\beta_j t} \rangle$. Then, $\mathbf{b}_i$ separates source i, that is, $z_i(t) = \mathbf{b}_i^\star \mathbf{x}(t)$ is an estimate of $s_i(t)$.

Proof: We seek a solution to (10) and (11). Observe that (10) and (11) are equivalent to

$$\mathbf{b}_i^\star \mathbf{R} \mathbf{b}_i = 1 \quad \text{and} \quad \mathbf{b}_i^\star \tilde{\mathbf{R}}_i \mathbf{b}_i = 0.$$

Therefore, if we define the $m \times 1$ vector $\mathbf{c}_i \stackrel{\text{def}}{=} \mathbf{R}^{H/2}\mathbf{b}_i$, the problem becomes

$$\|\mathbf{c}_i\| = 1 \quad \text{and} \quad \mathbf{c}_i^\star \mathbf{R}^{-1/2}\tilde{\mathbf{R}}_i\mathbf{R}^{-H/2}\mathbf{c}_i = 0$$

whose solution is given by the least eigenvector of $\mathbf{R}^{-1/2}\tilde{\mathbf{R}}_i\mathbf{R}^{-H/2}$. □

We now consider the general case where we permit the source signals to have common cycle frequencies. The following notation is used. Assume there are d distinct cycle frequencies $\beta_1, \cdots, \beta_d$ and for each i, there are precisely d_i source signals with cycle frequency β_i. (Clearly, $m = d_1 + \cdots + d_d$.) We write $\mathbf{s}(t) = [\mathbf{s}_1^T(t), \cdots, \mathbf{s}_d^T(t)]^T$, where each vector $\mathbf{s}_i(t)$ contains the d_i source signals with cycle frequency β_i. Similarly, we partition the mixing matrix as $\mathbf{A} = [\mathbf{A}_1, \cdots, \mathbf{A}_d]$.

The following result is an extension of Theorem 2 and can be proved in a similar fashion. Let $\mathbf{z}_i(t) = \mathbf{B}_i^H \mathbf{x}(t)$ be a $d_i \times 1$ random vector satisfying

$$\left\langle \mathbf{z}_i(t)\mathbf{z}_i^\star(t) \sum_{\beta_j \neq \beta_i} e^{\mathrm{J}\beta_j t} \right\rangle = \mathbf{0} \tag{12}$$

$$\left\langle \mathbf{z}_i(t)\mathbf{z}_i^\star(t) e^{\mathrm{J}\beta_i t} \right\rangle = \mathbf{I} \tag{13}$$

$$\left\langle \mathbf{z}_i(t+k)\mathbf{z}_i^\star(t) e^{\mathrm{J}\beta_i t} \right\rangle \quad \text{is diagonal} \tag{14}$$

for $k = \tau_1, \cdots, \tau_K$. Then, $\mathbf{z}_i(t)$ is an estimate of $\mathbf{s}_i(t)$ [that is, $\mathbf{B}_i^H \mathbf{A}\mathbf{s}(t) = \mathbf{P}\mathbf{\Lambda}\mathbf{s}_i(t)$, where $\mathbf{P}$ is a permutation matrix, and $\mathbf{\Lambda}$ a nonsingular diagonal matrix]. This leads to the following generalization of Theorem 3 for separating sources with a common cycle frequency.

Theorem 4: Define the matrix $\mathbf{B}_i = \mathbf{R}^{-H/2}\mathbf{U}_i\mathbf{V}_i$, where $\mathbf{R}^{-1/2}$ is defined as in Theorem 3, $\mathbf{U}_i$ is an $m \times d_i$ matrix whose column vectors form an orthogonal basis of $\mathrm{Ker}(\mathbf{R}^{-1/2}\tilde{\mathbf{R}}_i\mathbf{R}^{-H/2})$, where $\tilde{\mathbf{R}}_i \stackrel{\text{def}}{=} \langle \mathbf{x}(t)\mathbf{x}^\star(t) \sum_{\beta_j \neq \beta_i} e^{\mathrm{J}\beta_j t} \rangle$, and $\mathbf{V}_i$ is a $d_i \times d_i$ unitary matrix that jointly diagonalizes[3] the matrices $\mathbf{M}_i(k)$ for $k = \tau_1, \cdots, \tau_K$, where $\mathbf{M}_i(k) \stackrel{\text{def}}{=} \mathbf{U}_i^H \mathbf{R}^{-1/2}\mathbf{R}_i(k)\mathbf{R}^{-H/2}\mathbf{U}_i$, and $\mathbf{R}_i(k) \stackrel{\text{def}}{=} \langle \mathbf{x}(t+k)\mathbf{x}^\star(t) e^{\mathrm{J}\beta_i t} \rangle$. Then, $\mathbf{B}_i$ separates out the source signals with common cycle frequency β_i, that is, $\mathbf{z}_i(t) = \mathbf{B}_i^H \mathbf{x}(t)$ is an estimate of $\mathbf{s}_i(t)$.

Proof: We seek to solve (12)–(14). Observe that (12) and (13) are equivalent to

$$\mathbf{B}_i^H \mathbf{R} \mathbf{B}_i = \mathbf{I} \quad \text{and} \quad \mathbf{B}_i^H \tilde{\mathbf{R}}_i \mathbf{B}_i = 0.$$

Therefore, if we define the $m \times d_i$ matrix $\mathbf{U}_i \stackrel{\text{def}}{=} \mathbf{R}^{H/2}\mathbf{B}_i$, the problem becomes

$$\mathbf{U}_i^H \mathbf{U}_i = \mathbf{I} \quad \text{and} \quad \mathbf{U}_i^H \mathbf{R}^{-1/2}\tilde{\mathbf{R}}_i\mathbf{R}^{-H/2}\mathbf{U}_i = 0.$$

This shows that $\mathbf{U}_i$ can be taken to be any orthogonal basis of $\mathrm{Ker}(\mathbf{R}^{-1/2}\tilde{\mathbf{R}}_i\mathbf{R}^{-H/2})$. $\mathbf{V}_i$ should be chosen to satisfy (14). In other words, $\mathbf{V}_i$ is a unitary matrix that joint diagonalizes $\mathbf{M}_i(k)$ [i.e., such that $\mathbf{V}_i\mathbf{M}_i(k)\mathbf{V}_i^H$ are diagonal] for $k = \tau_1, \cdots, \tau_K$. □

Remarks:

1) The number of sources m can be estimated as the number of nonzero eigenvalues of $\mathbf{R}$, e.g., [11]. Similarly, the number of sources d_i with cycle frequency β_i can be estimated as the dimension of the kernel of $\mathbf{R}^{-1/2}\tilde{\mathbf{R}}_i\mathbf{R}^{-H/2}$.
2) In the case where (2) is not satisfied, we replace (11) by $|\langle |z_i(t)|^2 e^{\mathrm{J}\beta_i t} \rangle| > 0$. As a consequence, the constraints $\mathbf{b}_i^\star \mathbf{R}\mathbf{b}_i = 1$ and $\mathbf{B}_i^H \mathbf{R}\mathbf{B}_i = \mathbf{I}$ are replaced[4] by $\mathbf{b}_i^\star \mathbf{U}_s\mathbf{U}_s^H\mathbf{b}_i = 1$ and $\mathbf{B}_i^H \mathbf{U}_s\mathbf{U}_s^H\mathbf{B}_i = \mathbf{I}$, respectively, where $\mathbf{U}_s$ denotes the $n \times m$ matrix of the m principal left singular eigenvectors of $\mathbf{R}$. Using the new constraints, the separating vector $\mathbf{b}_i$ and the separating matrix $\mathbf{B}_i$ are obtained by replacing in their respective expressions $\mathbf{R}^{1/2}$ by $\mathbf{U}_s$, i.e., $\mathbf{b}_i = \mathbf{U}_s\mathbf{c}_i$ and $\mathbf{B}_i = \mathbf{U}_s\mathbf{U}_i\mathbf{V}_i$.

V. Implementation

The algorithmic implementations of Theorems 3 and 4 can be obtained easily from the theorems themselves. This section derives an iterative optimization algorithm based on Theorem 2. An implementation of Theorem 1 follows readily from this because Theorem 1 is essentially Theorem 2 with $K = 0$.

Based on Theorem 2, we define the following contrast function [12]:

$$G(\mathbf{z}) \stackrel{\text{def}}{=} \sum_{k=\tau_0}^{\tau_K} \sum_{1\leq i<j\leq m} \left[|r_{ij}(k) + r_{ji}(k)|^2 + |r_{ij}(k) - r_{ji}(k)|^2 \right] + \sum_{i=1}^{m} |r_{ii}(0) - 1|^2 \tag{15}$$

where $\tau_0 = 0$. Note that other contrast functions are possible; $G(\mathbf{z})$ turns out to be a convenient choice. The separation criterion of Theorem 2 takes the form

$$G(\mathbf{z}(t)) = 0 \Rightarrow \mathbf{B} \text{ is a separating matrix} \tag{16}$$

[3]An efficient joint diagonalization algorithm is presented in [6].

[4]The proof is similar to that of Theorems 3 and 4 and thus is omitted for simplicity.

where $\mathbf{z}(t) = \mathbf{B}\mathbf{x}(t)$. The following method of solving $G(\mathbf{z}(t)) = 0$ was inspired by [7]. It is a block algorithm based on the natural gradient technique [10]. Solutions are obtained iteratively in the form

$$\mathbf{B}^{(p+1)} = \left(\mathbf{I} + \boldsymbol{\epsilon}^{(p)}\right)\mathbf{B}^{(p)} \tag{17}$$

$$\mathbf{z}^{(p+1)}(t) = \left(\mathbf{I} + \boldsymbol{\epsilon}^{(p)}\right)\mathbf{z}^{(p)}(t). \tag{18}$$

At iteration p, the matrix $\boldsymbol{\epsilon}^{(p)} = [\epsilon_{ij}^{(p)}]_{1\le i,\, j\le m}$ is determined from a local linearization of $G(\mathbf{B}\mathbf{x}(t))$. It is an approximate Newton technique with the benefit that $\boldsymbol{\epsilon}^{(p)}$ is simple to compute (no Hessian inversion) under the additional assumption that $\mathbf{B}^{(p)}$ is close to a separating matrix. The derivation of $\boldsymbol{\epsilon}^{(p)}$ is now given.

At the pth iteration, we approximate $r_{ij}^{(p)}(\tau) \stackrel{\text{def}}{=} \langle z_i^{(p)}(t+\tau) z_j^{*(p)}(t) e^{\mathrm{J}\beta_i t}\rangle$ by its sample estimate

$$r_{ij}^{(p)}(k) \approx \frac{1}{T-k}\sum_{t=1}^{T-k} z_i^{(p)}(t+k) z_j^{*(p)}(t) e^{\mathrm{J}\beta_i t} \tag{19}$$

where T is the number of observations. When $\beta_i = \beta_j$, by using (18), we have

$$\begin{aligned} r_{ij}^{(p+1)}(k) &\stackrel{\text{def}}{=} \left\langle z_i^{(p+1)}(t+k) z_j^{(p+1)}(t)^* e^{\mathrm{J}\beta_i t}\right\rangle \\ &= \left\langle \left(z_i^{(p)}(t+k) + \sum_{q=1}^{m} \epsilon_{iq}^{(p)} z_q^{(p)}(t+k)\right) \right. \\ &\qquad \left. \cdot \left(z_j^{(p)^*}(t) + \sum_{l=1}^{m} \epsilon_{il}^{(p)*} z_l^{(p)*}(t)\right) e^{\mathrm{J}\beta_i t}\right\rangle. \end{aligned}$$

By assumption, $\mathbf{B}^{(p)}$ is close to being a separating matrix. This implies that the following terms are negligible: $|\epsilon_{ij}^{(p)}| \ll 1$, $|\langle z_i^{(p)}(t+k) z_l^{(p)*}(t) e^{\mathrm{J}\beta_i t}\rangle| \ll 1$ for $l \neq i$, and $|\langle z_q^{(p)}(t+k) z_j^{(p)*}(t) e^{\mathrm{J}\beta_i t}\rangle| \ll 1$ for $q \neq j$. A first-order approximation of $r_{ij}^{(p+1)}(k)$ is thus given by

$$r_{ij}^{(p+1)}(k) \approx r_{ij}^{(p)}(k) + \epsilon_{ji}^{*(p)} r_{ii}^{(p)}(k) + \epsilon_{ij}^{(p)} r_{jj}^{(p)}(k). \tag{20}$$

When $i \neq j$, $\epsilon_{ij}^{(p)}$ is chosen to be the solution of the following least squares (LS) minimization problem obtained by substituting (20) into (15)

$$\min \left\| \left[\mathbf{r}_{jj}^{(p)}, \mathbf{r}_{ii}^{(p)}\right]\mathbf{E}_{ij}^{(p)} + \left[\frac{1}{2}\left(\mathbf{r}_{ij}^{(p)} + \mathbf{r}_{ji}^{(p)}\right), \frac{1}{2\mathrm{J}}\left(\mathbf{r}_{ij}^{(p)} - \mathbf{r}_{ji}^{(p)}\right)\right] \right\|^2$$

where

$$\mathbf{E}_{ij}^{(p)} \stackrel{\text{def}}{=} \begin{bmatrix} \Re e\left(\epsilon_{ij}^{(p)}\right) & \Im m\left(\epsilon_{ij}^{(p)}\right) \\ \Re e\left(\epsilon_{ji}^{(p)}\right) & -\Im m\left(\epsilon_{ji}^{(p)}\right) \end{bmatrix} \tag{21}$$

$$\mathbf{r}_{ij}^{(p)} = \left[r_{ij}^{(p)}(0), r_{ij}^{(p)}(\tau_1), \cdots, r_{ij}^{(p)}(\tau_K)\right]^T. \tag{22}$$

Here, $\Re e(x)$ and $\Im m(x)$ denote the real and imaginary parts of the complex variable x. The solution of the LS minimization problem is

$$\begin{aligned} \mathbf{E}_{ij}^{(p)} &= -\left[\mathbf{r}_{jj}^{(p)}, \mathbf{r}_{ii}^{(p)}\right]^{\#}\left[\frac{1}{2}\left(\mathbf{r}_{ij}^{(p)} + \mathbf{r}_{ji}^{(p)}\right), \frac{1}{2\mathrm{J}}\left(\mathbf{r}_{ij}^{(p)} - \mathbf{r}_{ji}^{(p)}\right)\right] \\ &= \left(\left[\mathbf{r}_{jj}^{(p)}, \mathbf{r}_{ii}^{(p)}\right]^H \left[\mathbf{r}_{jj}^{(p)}, \mathbf{r}_{ii}^{(p)}\right]\right)^{-1} \left[\mathbf{r}_{jj}^{(p)}, \mathbf{r}_{ii}^{(p)}\right]^H \\ &\qquad \cdot \left[\frac{1}{2}\left(\mathbf{r}_{ij}^{(p)} + \mathbf{r}_{ji}^{(p)}\right), \frac{1}{2\mathrm{J}}\left(\mathbf{r}_{ij}^{(p)} - \mathbf{r}_{ji}^{(p)}\right)\right] \end{aligned} \tag{23}$$

where the superscript $\#$ denotes the pseudo-inverse of a matrix. Similarly, for $i = j$, we obtain

$$\epsilon_{ii}^{(p)} = \frac{1 - r_{ii}^{(p)}(0)}{2 r_{ii}^{(p)}(0)}. \tag{24}$$

When $\beta_i \neq \beta_j$, we can further simplify (20) by using the fact that $|\langle z_j^{(p)}(t+k) z_j^{(p)*}(t) e^{\mathrm{J}\beta_i t}\rangle| \ll 1$. The approximation becomes $r_{ij}^{(p+1)}(k) \approx r_{ij}^{(p)}(k) + \epsilon_{ji}^{*(p)} r_{ii}^{(p)}(k)$. Substituting this into (15) leads to

$$\epsilon_{ij}^{(p)} = -\left(\frac{\mathbf{r}_{jj}^{(p)\star}\mathbf{r}_{ji}^{(p)}}{\left\|\mathbf{r}_{jj}^{(p)}\right\|^2}\right)^*. \tag{25}$$

In summary, the iterative algorithm for Theorem 2 is as follows. Use (19) to compute $\mathbf{r}_{ij}^{(p)}$ defined in (22). Compute the diagonal elements of the matrix $\boldsymbol{\epsilon}^{(p)}$ by (24) and the off-diagonal elements by (23) if $\beta_i = \beta_j$ or by (25) if $\beta_i \neq \beta_j$. Finally, update $\mathbf{B}$ and $\mathbf{z}(t)$ by (17) and (18).

The iterative algorithm for Theorem 1 is the same as for Theorem 2 with the simplification that the off-diagonal elements are computed by

$$\epsilon_{ij}^{(p)} = -\left(\frac{r_{ji}^{(p)}(0)}{r_{jj}^{(p)}(0)}\right)^*. \tag{26}$$

Remarks:

1) Adaptive versions of the above algorithms for Theorems 1 and 2 can be derived following the approaches presented in [8], [9], and [13].
2) Similarly to the previous algorithms, we can generalize the iterative algorithm to the case where (3) is not satisfied. In that case, we replace in (8) and (9) $r_{ii}(0) = 1$ by $|r_{ii}(0)| = 1$, which leads to the same updating equations (17)–(26) except that (24) is replaced by

$$\epsilon_{ii}^{(p)} = \frac{1 - \left|r_{ii}^{(p)}(0)\right|}{2\left|r_{ii}^{(p)}(0)\right|}.$$

VI. Numerical Simulations

This section presents simulation results for the four algorithms derived from Theorems 1 to 4. These algorithms are, respectively, called ATH1, ATH2, ATH3, and ATH4 for conve-

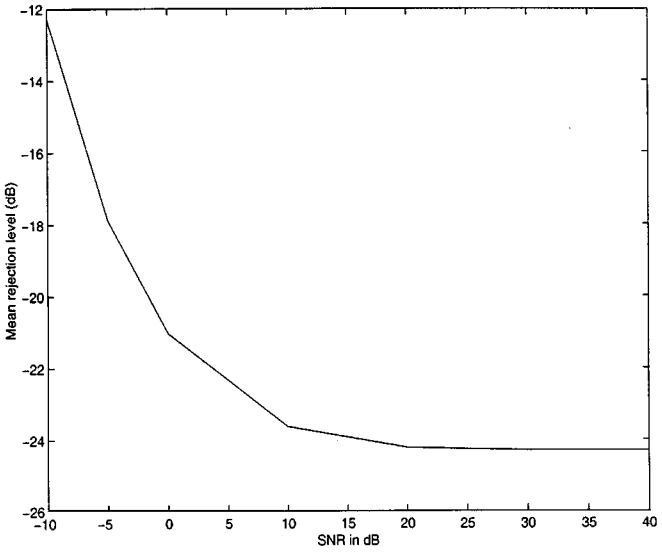

Fig. 1. Performance of ATH1 versus SNR.

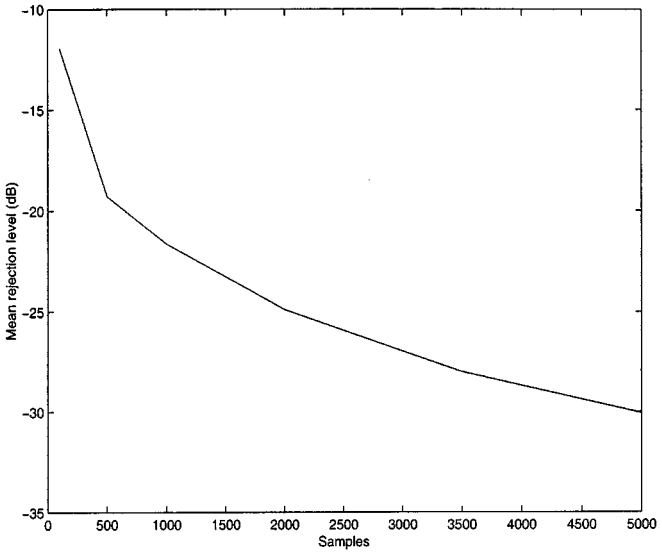

Fig. 2. Performance of ATH1 versus sample size (SNR $= 0$ dB).

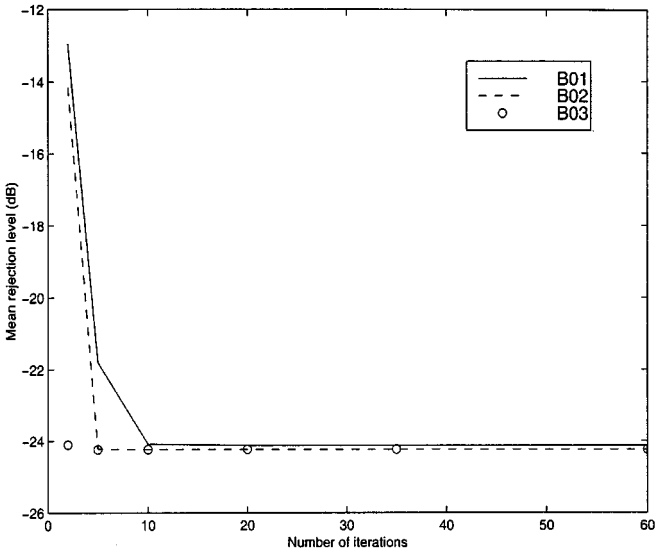

Fig. 3. Performance of ATH1 versus iterations for different initiations of $\mathbf{B}$, where SNR $= 25$ dB.

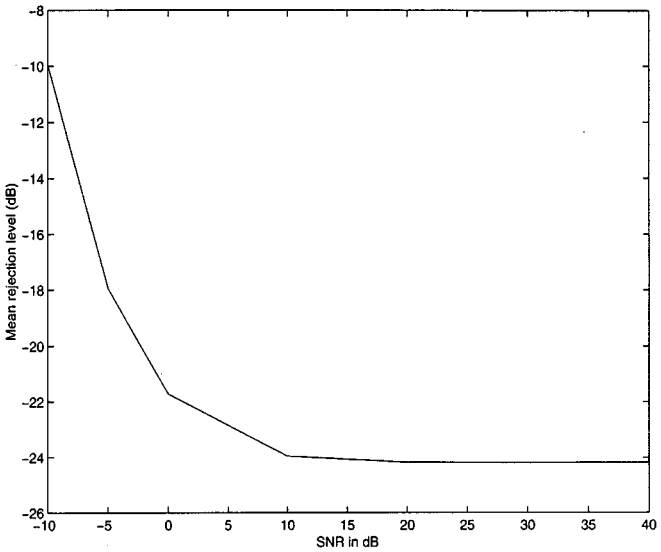

Fig. 4. Performance of ATH3 versus SNR.

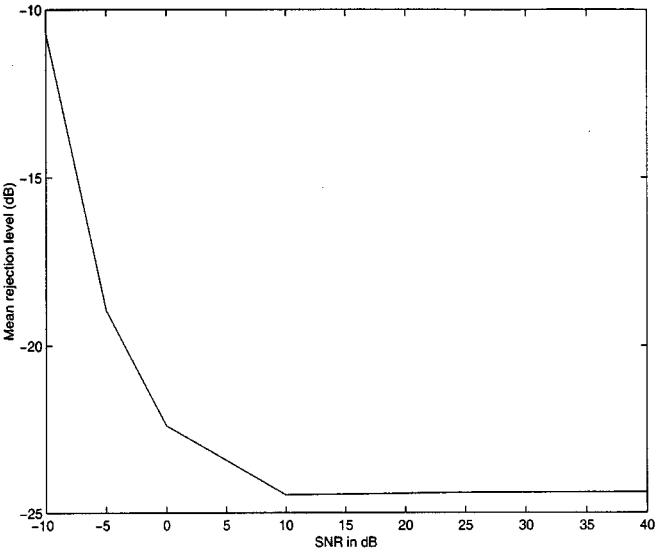

Fig. 5. Performance of ATH2 versus SNR.

nience. The performance of each algorithm is measured by its "mean rejection level" performance index [6] defined as

$$\mathcal{I}\,\mathrm{perf} \stackrel{\mathrm{def}}{=} \sum_{q \neq p} \frac{E\,|(\mathbf{BA})_{pq}|^2}{E\,|(\mathbf{BA})_{pp}|^2}.$$

It is estimated by averaging 100 independent trials. Each simulation is based on the following model. A five-element ($n = 5$) uniform linear array with half wavelength sensor spacing receives two signals ($m = 2$) in the presence of stationary complex temporally white but spatially colored noise. The two signals are first-order autoregressive Gaussian processes [with coefficients $a_1 = 0.95\exp(\jmath 0.5)$ and $a_2 = 0.5\exp(\jmath 0.7)$] modulated by sinusoids $\cos\alpha_1$ and $\cos\alpha_2$, respectively. The sources are thus cyclostationary with cycle frequencies $2\alpha_1$ and $2\alpha_2$ (see [1]). The sources arrive from the directions $\phi_1 = 10$ and $\phi_2 = 30°$. The snapshot size is $T = 1000$ samples. The signal-to-noise ratio (SNR) is defined as SNR $= -10\log_{10}\sigma^2$,

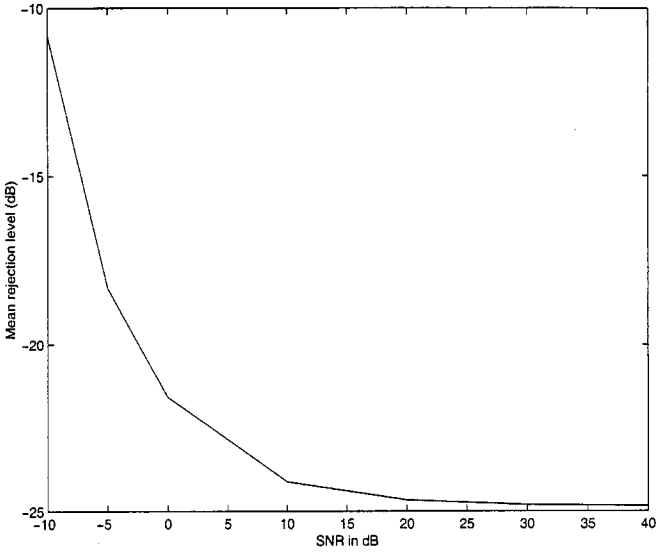

Fig. 6. Performance of ATH4 versus SNR.

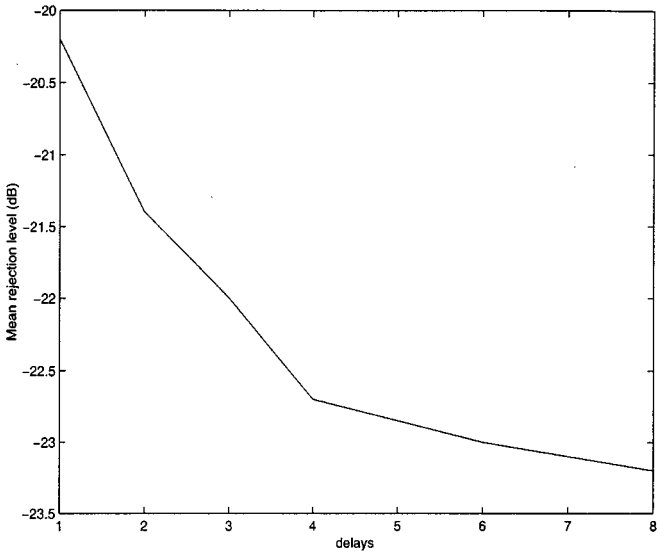

Fig. 7. Performance of ATH2 versus delays (lags), where SNR $= 0$ dB.

where σ^2 is the noise variance. The noise covariance is assumed to be of the form $\mathbf{R}_n = \sqrt{n}\sigma^2\mathbf{Q}\mathbf{Q}^H/\|\mathbf{Q}\|^2$, where $\mathbf{Q}$ is given by $\mathbf{Q}_{ij} = 0.9^{|i-j|}$.

Example 1: The cycle frequencies of the two sources are $\beta_1 = 2\alpha_1 = 0.6$ and $\beta_2 = 2\alpha_2 = 1.4$. Figs. 1–3 show the mean rejection levels of ATH1 against, respectively, the SNR, the sample size, and the number of iterations. It is clear from Fig. 1 that ATH1 performs well at moderate and high SNRs. Fig. 2 shows that as is to be expected, the greater the number of samples, the smaller the rejection level. Each of the three traces in Fig. 3 corresponds to a different initialization scheme for the separating matrix $\mathbf{B}$. The solid line represents the case when $\mathbf{B}^{(0)}$ is a random matrix. The dashed line corresponds to $\mathbf{B}^{(0)} = \mathbf{\Lambda}_s^{-(1/2)}\mathbf{U}_s^H$, where $\mathbf{U}_s$ and $\mathbf{\Lambda}_s$ are the eigenvector matrix and diagonal eigenvalue matrix of the autocorrelation matrix of the array output. The circles correspond to the choice $\mathbf{B}^{(0)} = \mathbf{A}^{\#} + \mathbf{\Delta}$, where $\mathbf{\Delta}$ denotes a small perturbation matrix.[5] The figure shows the robustness of ATH1 to different $\mathbf{B}^{(0)}$.

The performance of ATH3 against SNR is shown in Fig. 4. Simulation results of ATH3 are similar to those of ATH1.

Example 2: The cycle frequencies of the two sources are $\beta_1 = \beta_2 = 0.6$. Figs. 5 and 6 show the mean rejection levels of ATH2 and ATH4 versus SNR. The number of time lags used was $K = 4$. Both ATH2 and ATH4 achieve good separation performance for moderate to high SNRs. Fig. 7 shows the performance gain caused by increasing the number of lags K. We have found experimentally that the gain in performance is most notable in difficult environments such as poor SNR, small spectral difference, ill-conditioned mixture matrix, etc.

VII. Conclusion

This paper studied the blind source separation (BSS) problem with the assumption that the source signals are cyclostationary. Identifiability and separability criteria based on second-order cyclostationary statistics (SOCS) alone were derived. The identifiability condition was used to define an appropriate contrast function. An iterative algorithm (ATH2) was derived to minimize this contrast function. This algorithm separates the sources even when they do not have distinct cycle frequencies. If the cycle frequencies are distinct, then ATH2 simplifies to ATH1. Because these algorithms separate all the source signals they may be inefficient if only a small number of sources are of interest. A noniterative algorithm (ATH4) is derived to separate only those sources of a particular cycle frequency. When all source signals have distinct cycle frequencies, ATH4 simplifies to ATH3. Simulation results showed the performance of these BSS algorithms.

References

[1] W. Gardner, "Exploitation of the spectral redundancy in cyclostationary signals," *IEEE Signal Processing Mag.*, vol. 8, pp. 14–37, 1991.
[2] Y. C. Liang, A. R. Leyman, and B. H. Soong, "Blind source separation using second order cyclic statistics," in *Proc. SPAWC*, Paris, France, 1997, pp. 57–60.
[3] S. V. Schell and W. A. Gardner, "Maximum likelihood and common factor analysis-based blind adaptive spatial filtering for cyclostationary signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Minneapolis, MN, Apr. 1993, pp. IV 292–295.
[4] B. G. Agee, S. V. Schell, and W. A. Gardner, "Spectral self-coherence restoral: A new approach to blind adaptive signal extraction using antenna arrays," *Proc. IEEE*, vol. 78, pp. 753–766, Apr. 1990.
[5] L. Castedo and A. R. Figueiras-Vidal, "An adaptive beamforming technique based on cyclostationary signal properties," *IEEE Trans. Signal Processing*, vol. 43, pp. 1637–1650, July 1995.
[6] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, and E. Moulines, "Blind source separation using second order statistics," *IEEE Trans. Signal Processing*, vol. 45, pp. 434–444, Feb. 1997.
[7] D. T. Pham and P. Garrat, "Blind separation of mixture of independent sources through a quasimaximum likelihood approach," *IEEE Trans. Signal Processing*, vol. 45, pp. 1712–1725, July 1997.
[8] J. F. Cardoso, A. Belouchrani, and B. Laheld, "A new composite criterion for adaptive and iterative blind source separation," in *Proc. ICASSP*, 1994, pp. 273–276.
[9] J. F. Cardoso and B. Laheld, "Equivariant adaptive source separation," *IEEE Trans. Signal Processing*, vol. 44, pp. 3017–3030, Dec. 1996.
[10] S. Amari and S. Douglas, "Why natural gradient," in *Proc. ICASSP*, vol. II, 1998, pp. 1213–1216.
[11] Q. Cheng and Y. Hua, "Detection of cisoids using least square error function," *IEEE Trans. Signal Processing*, vol. 45, pp. 1584–1590, July 1997.

[5]This initialization is used to check the convergence of our algorithm.

[12] P. Comon, "Independent component analysis—A new concept," *Signal Process.*, vol. 36, pp. 287–314, 1994.
[13] Y. Xiang, K. Abed-Meraim, and Y. Hua, "Adaptive blind source separation by second order statistics and natural gradient," in *Proc. ICASSP*, 1999.
[14] A. V. Dandawate and G. B. Giannakis, "Asymptotic theory of mixed time averages and kth-order cyclic moment and cumulant statistics," *IEEE Trans. Inform. Theory*, vol. 41, pp. 216–232, Jan. 1995.

Karim Abed-Meraim was born in Algiers, Algeria, in 1967. He received the M.S. degree from Ecole Polytechnique, Paris, France, in 1990 and the Ph.D. degree from Ecole Nationale Supérieure des Télécommunications (ENST), Paris, in 1995 in the field of signal processing and communications.

He currently is Associate Professor at the Signal and Image Processing Department of ENST. His research interests are in statistical signal processing and include system identification, multiuser detection, space-time coding, (blind) array processing, and performance analysis.

Yong Xiang received the B.Eng. and M.Eng. degrees from the University of Electronic Science and Technology of China, Chengdu, in 1983 and 1989, respectively. He is currently pursuing the Ph.D. degree at the University of Melbourne, Parkville, Australia.

He was with the Southwest Institute of Electronic Equipment of China, Chengdu, from 1983 to 1986. In 1989, he joined the University of Electronic Science and Technology of China, where he was a Lecturer from 1989 to 1991 and an Associate Professor from 1992 to 1997. He is currently a Senior Communications Engineer with Bandspeed Pty. Ltd., Melbourne.

Jonathan H. Manton received the B.Sc. degree in mathematics and B.Eng. degree in electrical engineering in 1995 and the Ph.D. degree in 1998, all from the University of Melbourne, Parkville, Australia.

In 2000, he became a holder of the prestigious Research Fellow Award by the Australian Research Council, which he chose to take up at the University of Melbourne. His current research interests include precoding for wireless communications, rank constrained filtering, and optimization on Riemannian manifolds and Lie groups.

Yingbo Hua (SM'92) was born in China in 1960. He received the B.E. degree from Nanjing Institute of Technology, China, in 1982, and M.S. and Ph.D. degrees (with GPA 4.00/4.00) from Syracuse University (SU), Syracuse, NY, in 1983 and 1988, respectively.

He worked as Teaching Assistant, Research Assistant, Summer Lecturer, and Research Fellow at SU. In July 1989, he joined the University of Melbourne, Parkville, Australia, where he was promoted from Lecturer to Senior Lecturer in 1992 and to Associate Professor and Reader in 1995. He was on Visiting Faculty with the Hong Kong University of Science and Technology from 1999 to 2000. He is now a Professor of electrical engineering with the University of California, Riverside, CA. He is an author/coauthor of over 80 journal articles and book chapters and 130 conference papers in the fundamental areas of estimation, detection, system identification, and fast algorithms, with applications in communications, remote sensing, and medical data analysis. He is a coeditor of the book *Signal Processing Advances in Wireless Communications* (Englewood Cliffs, NJ: Prentice-Hall, 2000).

Dr. Hua was a winner of Chinese Government Scholarship for Overseas Study from 1983 to 1984 and a Syracuse University Graduate Fellowship from 1985 to 1986). He served as Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1994 to 1997 and as Associate Editor of the IEEE SIGNAL PROCESSING LETTERS since 1998. He is an Invited Reviewer for over 16 international journals. He is an Invited Speaker, Session Chair, and Organizer of many international conferences. He has been an Elected Member of the IEEE Signal Processing Society's Technical Committee for Sensor Array and Multichannel Signal Processing since 1998 and was an Elected Member of the Underwater Acoustic Signal Processing Technical Committee from 1997 to 1998.