Waveform-Preserving Blind Estimation of Multiple Independent Sources

Lang Tong, *Member, IEEE*, Yujiro Inouye, *Member, IEEE*, and Ruey-wen Liu, *Fellow, IEEE*

***Abstract*—The problem of blind estimation of source signals is to estimate the source signals without knowing the transmission channel. The so-called waveform-preserving blind estimation minimizes the difference between the signal waveforms and their estimates. In this paper, it is shown that the minimum-variance unbiased estimates can be obtained if and only if the transmission channel can be identified (blindly). We then show that the channel can be blindly identified if and only if there is no more than one Gaussian source. This condition suggests that waveform-preserving blind estimation can be achieved over a wide range of signal processing applications, including those cases when the source signals have identical non-Gaussian distributions. The constructive proof of the necessary and sufficient condition serves as a foundation for the development of waveform-preserving blind signal estimation algorithms. Examples are also presented to demonstrate the applications of the theoretical results.**

I. Introduction

VARIOUS signal processing applications involve estimating signals of interest from distorted and noise-contaminated observations. The so-called blind signal estimation is one that estimates unknown source signals from the observations without explicit knowledge of the transmission channel as shown in the schematic diagram in Fig. 1. Recently, blind signal estimation has attracted increasing research attention [1]–[14], [16] due to its wide range of applications from mobile radio and regenerative satellite communications to array signal processing.

The approaches to blind signal estimation have been categorized [1] into two groups: spatial-coherence exploitation techniques and property restoring techniques. The spatial-coherence exploitation techniques utilize the structural information of the transmission channel. In array signal processing, for example, the channel can be characterized by a Vandermonde matrix under a narrow-band assumption and a linear array structure. Well-known algorithms such as MUSIC [15] and ESPRIT [16] take advantage of such structural information to obtain the estimates of the parameters. Unfortunately, the structural information is not always available, nor always accurate in many practical circumstances. For example, when the sensor arrays have unknown sensor-gain and sensor-phase perturbation, or there is strong cochannel interference, the structural assumptions used in MUSIC and ESPRIT are no longer valid. As a consequence, the performance of these algorithms may suffer significant degradation [18], [19]. The property restoring techniques, on the other hand, capture the properties of the source signals while leaving the channel structure arbitrary. Signal properties such as low modules variation [2] and self-coherence [1] lead to blind signal estimation algorithms for communication applications. See also [17] where the physical nature of the signals is used. However, the extracted signals by property restoring techniques, though optimal with respect to certain known properties of the sources, may not possess other unknown but important features of the actual sources. For example, there is often no guarantee that the signals extracted via the property restoring techniques will have the same waveform as the actual source signals.

The objective of the waveform-preserving blind estimation is to minimize the difference between the waveforms of the extracted signals and those of the actual source signals. Several waveform-preserving algorithms have been developed for special situations [3]–[9]. Although these algorithms are successful under assumed conditions, they do have limitations. For example, blind identification using fourth-order moments (FOBI-type) proposed in [4], [10] fails when all sources are identically distributed. By using second-order property of signal processes, the algorithm referred to as AMUSE [10] can obtain waveform-preserving identification of Gaussian signals, but fails when the processes involved are white. In short, the waveform-preserving estimation algorithms proposed so far have not fully utilized the statistical information contained in the source signals. The important question that remains to be answered is the following: Can waveform-preserving blind estimation be achieved when all the statistical information of the source signals are used?

This paper presents a complete answer to the above question. For independent sources, a necessary and suf-

Manuscript received October 16, 1990; revised July 20, 1992. The associate editor coordinating the review of this paper and approving it for publication was Prof. S. Unnikrishna Pillai. This work was supported in part by the Office of Naval Research under Contract N00014-91-J-1461 and the Laboratory for Image and Signal Analysis, University of Notre Dame, Notre Dame, IN 46556.

L. Tong is with the Department of Electrical and Computer Engineering, West Virginia University, Morgantown, WV 26506.

Y. Inouye is with the Department of Systems Engineering, Osaka University, Toyonaka, Osaka 560, Japan.

R. Liu is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556.

IEEE Log Number 9208853.

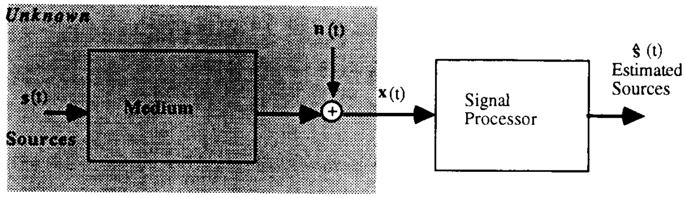


Fig. 1. A schematic diagram of the blind signal estimation.

ficient condition simply states that the waveform-preserving blind estimation can be achieved from the output statistics if and only if there is no more than one Gaussian source. This result suggests that the waveform-preserving blind signal estimation can be achieved over a wide range of practical applications. Another purpose of this paper is to establish the necessary theoretical results and the formulations which can be used for the development of waveform-preserving blind estimation algorithms [12].

This paper is organized as follows. Section II introduces notational conventions and some basic facts about higher order statistics; Section III establishes a fundamental relation between the blind signal estimation and the blind channel identification; Section IV presents the main theorem concerning the channel identifiability; and concluding remarks are given in Section V.

II. Preliminaries

A. Notations

For a matrix $A \in \mathbb{R}^{n \times m}$, $A_{(i,\cdot)}$ and $A_{(\cdot,j)}$ denote the ith row vector and jth column vector of matrix A, respectively. The symbol I_n is reserved for the identity matrix in $\mathbb{R}^{n \times n}$, and $\mathbf{O}_{m \times n}$ stands for the $m \times n$ zero matrix. The symbols $\mathbb{P}^n$ and $\mathbb{D}^n$ denote the set of permutation matrices and the set of nonsingular diagonal matrices in $\mathbb{R}^{n \times n}$, respectively. A sign matrix, denoted as $\mathring{I}$, is a diagonal matrix whose diagonal elements are either $+1$ or -1. The notation $R > 0$ means that R is positive definite. The set of positive integers is denoted as $\mathbb{Z}$ while the index set $\{1, 2, \cdots, n\}$ is denoted as $\mathbb{Z}_n$. The symbol $\mathbb{Z}_n^m$ denotes the Cartesian product of m $\mathbb{Z}_n$'s.

The following conventions are used for matrix and vector operations. The symbols, rank (A), tr (A), and A^t denote the rank, the trace and the transpose of a matrix A, respectively. For a positive definite matrix Λ, $\Lambda^{1/2}$ denotes the square root of Λ. The Moore–Penrose pseudoinverse of a matrix A is denoted by $A^\dagger$. The symbol diag $(a_1, a_2, \cdots, a_n)$ stands for a diagonal matrix whose (i, i)th element is a_i, while diag $(A_1, A_2, \cdots, A_n)$ denotes a block diagonal matrix with diagonal blocks $\{A_i\}$. For an n-dimensional vector $x \in \mathbb{R}^n$, diag (x) denotes the diagonal matrix whose (i, i)th diagonal entry is the ith component of x. The inverse operation of diag (D), denoted as diag^{-1} (D), maps a diagonal matrix D to a column vector whose ith component is the (i, i)th entry of D. The Hadamard product of two vectors x and y in $\mathbb{R}^n$, denoted as $x \blacklozenge y$, is a new vector $\mathbb{R}^n$ whose components are the products of respective products of x and y, i.e., $x \blacklozenge y = [x_1 y_1, x_2 y_2, \cdots, x_n y_n]^t$.

B. Higher Order Statistics

Higher order statistics have recently been applied to a wide range of signal processing applications that are difficult or impossible using second-order techniques. See a recent review by Mendel [20] and the references therein. As we shall see in this paper, by taking advantage of all the information contained in higher order statistics, one can estimate or reach the theoretical boundary of certain otherwise very difficult problems.

Given an n-dimensional random vector $x = [x_1, \cdots, x_n]^t$, the nth-order (joint) cumulant of n random variables $x_1, x_2, \cdots, x_n$, denoted as cum $(x_1, x_2, \cdots, x_n)$, is related to their moments as follows [22]:

$$\text{cum}(x_1, x_2, \cdots, x_n) = \sum (-1)^{(p-1)} \cdot (p-1)! E\left\{\prod_{j \in \nu_1} x_j\right\} E\left\{\prod_{j \in \nu_2} x_j\right\} \cdots E\left\{\prod_{j \in \nu_p} x_j\right\} \tag{2.1}$$

where $\nu_1, \nu_2, \cdots, \nu_p$ is a partition of the index set $\{1, 2, \cdots, n\}$ and the summation extends over all partitions. As simple examples, for zero-mean random variables x_1, x_2, x_3, and x_4, cum $(x_1, x_2) = E\{x_1 x_2\}$, cum $(x_1, x_2, x_3) = E\{x_1 x_2 x_3\}$, and cum $(x_1, x_2, x_3, x_4) = E\{x_1 x_2 x_3 x_4\} - (E\{x_1 x_2\} E\{x_3 x_4\} + E\{x_1 x_3\} E\{x_2 x_4\} + E\{x_1 x_4\} E\{x_2 x_3\})$. The nth-order cumulant of a random variable x, denoted as $\kappa_x^{(n)}$ for notational convenience, can be obtained from (2.1) by letting $x = x_1 = x_2 = \cdots = x_n$. The $(k+2)$th-order cumulants of an n-dimensional vector x are defined by a set of n^k cumulant matrices $C_v^{(k+2)}$'s in $\mathbb{R}^{n \times n}$ where $v = [v_1, v_2, \cdots, v_k] \in \mathbb{Z}_n^k$. In particular, the (p, q)th element of $C_v^{(k+2)} \in \mathbb{R}^{n \times n}$ is cum $(x_p, x_q, x_{v_1}, x_{v_2}, \cdots, x_{v_k})$. In other words, for every $v \in \mathbb{Z}_n^k$,

$$C_v^{(k+2)} = [\text{cum}(x_p, x_q, x_{v_1}, x_{v_2}, \cdots, x_{v_k})]_{p,q=1}^n. \tag{2.2}$$

The cumulant matrix defined above is an extension of the fourth-order cumulant matrix exploited by a number of researchers. See, for example, [21].

The following properties are useful in the later development and can be shown directly from the definition [22]:

P1) For any constants $a_1, a_2, \cdots, a_n$,

$$\text{cum}(a_1x_1, a_2x_2, \cdots, a_nx_n) = a_1a_2\cdots a_n \,\text{cum}(x_1, x_2, \cdots, x_n).$$

P2) If any nontrivial proper subgroup of the random variables x_i's is independent of the rest of the x_i's then $\text{cum}(x_1, x_2, \cdots, x_n) = 0$.

P3) If random variables x_i's are independent of random variables y_i's, then

$$\text{cum}(x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n) = \text{cum}(x_1, x_2, \cdots, x_n) + \text{cum}(y_1, y_2, \cdots, y_n).$$

P4) A random vector $\boldsymbol{x}$ is Gaussian if and only if all of its cumulants of order higher than 2 are equal to zero.

III. Blind Signal Estimation and Blind Channel Identification

A. The Basic Model and the Model Assumptions

In this paper, we consider a linear model given by

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t), \qquad t \in \mathbb{Z} \tag{3.1}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the observation vector, $\boldsymbol{s}(t) \in \mathbb{R}^m$ is the vector of (unknown) source signal, $\boldsymbol{n}(t) \in \mathbb{R}^n$ is the additive random noise vector, and $\boldsymbol{A} = [a_{ij}] \in \mathbb{R}^{n \times m}$ is the (unknown) parameter matrix which characterizes the channel. A typical environment that can be described by the above equation is an antenna array system excited by narrow-band signal $\boldsymbol{s}(t)$. The matrix $\boldsymbol{A}$ in (3.1) models the polarization and the direction-of-arrival-dependent sensor gains and phases along with mutual coupling effects. Although the presentation here deals with real cases, the complex cases follow easily.

Throughout this paper, $\boldsymbol{A}^0$ and $\boldsymbol{s}^0(\cdot)$ stand for the actual parameter matrix and the source signal vector process, respectively. The objective of the waveform-preserving blind estimation is to obtain, from the observation process $\boldsymbol{x}(\cdot)$, the best estimate of the source signal process $\boldsymbol{s}^0(\cdot)$ without knowing $\boldsymbol{A}^0$.

It is assumed that $\boldsymbol{A}^0$ and $\boldsymbol{s}^0(\cdot)$ satisfy the following assumptions:

AS1) $\boldsymbol{A}^0$ has full column rank.

AS2) $\boldsymbol{s}^0(\cdot)$ is an m-variate zero-mean stationary second-order ergodic process and $\boldsymbol{R}_{s^0} = \boldsymbol{E}(\boldsymbol{s}^0(\cdot)(\boldsymbol{s}^0(\cdot))') > 0$. For each t, random variables $\{s_i^0(t)\}_{i=1,m}$ are mutually independent.

AS3) Cumulants of all orders exist for all $s_i^0(t)$'s.

AS4) $\boldsymbol{n}(\cdot)$ is a zero-mean stationary Gaussian process with noise covariance $\boldsymbol{R}_n$. For each t, the random variables $\{n_i(t)\}_{i=1,n}$ are independent of $\{s_i^0(t)\}_{i=1,m}$.

B. Waveform-Preserving Blind Signal Estimation and Blind Channel Identification

Similar to the problem of the blind signal estimation, the blind channel identification is to identify $\boldsymbol{A}^0$ from $\boldsymbol{x}(t)$ without knowing the source signal $\boldsymbol{s}^0(t)$. We shall now establish a relation between the waveform-preserving blind estimation and the blind channel identification. Unlike the property restoring techniques where the identification of the image space of $\boldsymbol{A}^0$ may be sufficient for the blind estimation, to achieve the optimal waveform-preserving blind estimation, as we shall demonstrate in the following, requires that the column vectors of matrix $\boldsymbol{A}^0$ be identified up to scale factors and a permutation.

We first introduce a new notation of waveform-preserving processes. A vector process $\tilde{\boldsymbol{s}}(\cdot)$ is called a waveform-preserving process of $\boldsymbol{s}^0(\cdot)$ if there exist a $\boldsymbol{P} \in \mathbb{P}^m$ and a $\Lambda \in \mathbb{D}^m$ such that

$$\tilde{\boldsymbol{s}}(\cdot) = \boldsymbol{P}\Lambda\boldsymbol{s}^0(\cdot).$$

Here the nonsingular diagonal matrix Λ indicates the scale ambiguities in the source signal components, and the permutation matrix $\boldsymbol{P}$ indicates the arbitrary arrangement of the source signal components in the vector $\boldsymbol{s}^0(t)$. Note that these ambiguities in $\tilde{\boldsymbol{s}}^0(\cdot)$ are inherited from the nature of blind estimation in which neither the channel nor the source signal is completely specified. *Our objective is to find, from the observation process $\boldsymbol{x}(t)$, the best linear estimate of some $\tilde{\boldsymbol{s}}(t)$.*

The approach is to find a matrix $\boldsymbol{H}$ such that the estimation

$$\hat{\boldsymbol{s}}(t) = \boldsymbol{H}\boldsymbol{x}(t), \qquad t \in \mathbb{Z} \tag{3.2}$$

is the minimum-variance unbiased estimate of $\boldsymbol{P}\Lambda\boldsymbol{s}^0(\cdot)$ for some $\boldsymbol{P} \in \mathbb{P}^m$ and some $\Lambda \in \mathbb{D}^m$. Let $\boldsymbol{e}(t)$ be the estimated error defined by

$$\boldsymbol{e}(t) \equiv \hat{\boldsymbol{s}}(t) - \boldsymbol{P}\Lambda\boldsymbol{s}^0(t). \tag{3.3}$$

We then have

$$\boldsymbol{e}(t) = (\boldsymbol{H}\boldsymbol{A}^0 - \boldsymbol{P}\Lambda)\boldsymbol{s}^0(t) + \boldsymbol{H}\boldsymbol{n}(t). \tag{3.4}$$

Since $\boldsymbol{E}\{\boldsymbol{H}\boldsymbol{n}(t)\} = \boldsymbol{0}$, for a given sample $\boldsymbol{s}^0(\cdot)$, the condition on $\hat{\boldsymbol{s}}(\cdot)$ being an unbiased estimate of $\boldsymbol{s}^0(\cdot)$ requires $\boldsymbol{E}\{\boldsymbol{e}(t)|\boldsymbol{s}^0(t)\} = \boldsymbol{0}$, i.e.,

$$(\boldsymbol{H}\boldsymbol{A}^0 - \boldsymbol{P}\Lambda)\boldsymbol{s}^0(t) = \boldsymbol{0}, \qquad \text{for all } t.$$

Because $\boldsymbol{R}_{s^0} > 0$, the above equation implies

$$\boldsymbol{H}\boldsymbol{A}^0 = \boldsymbol{P}\Lambda. \tag{3.5}$$

Observe that if $\boldsymbol{A}^0$ is a square matrix, then

$$\boldsymbol{H}^{-1} = \boldsymbol{A}^0\Lambda^{-1}\boldsymbol{P}'.$$

This implies that the column vectors of $\boldsymbol{H}^{-1}$ identify the column vectors of $\boldsymbol{A}^0$ up to scale factors and a permutation. In general, define the time-average conditional mean-squares error J_{s^0} as

$$J_{s^0} = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} \boldsymbol{E}\{\boldsymbol{e}'(t)\boldsymbol{e}(t)|\boldsymbol{s}^0(t)\}. \tag{3.6}$$

From the ergodicity assumption (in AS2) and (3.2)–(3.4), we have

$$J_{s^0} = \text{tr}(\boldsymbol{H}\boldsymbol{R}_n\boldsymbol{H}') + \lim_{N\to\infty}\frac{1}{N}\text{tr}([(\boldsymbol{H}\boldsymbol{A}^0 - \boldsymbol{P}\Lambda)\boldsymbol{s}^0(t)] \cdot [(\boldsymbol{H}\boldsymbol{A}^0 - \boldsymbol{P}\Lambda)\boldsymbol{s}^0(t)]')$$

$$= \text{tr}\,(HR_nH') + \text{tr}\,((HA^0 - P\Lambda)R_{s^0}(HA^0 - P\Lambda)')$$
$$= E\{e'(t)e(t)\}. \tag{3.7}$$

Consequently, the minimum-variance unbiased estimate can be obtained from the solution of the following optimization problem:

$$\underset{H \in \mathbb{R}^{m\times n}}{\text{minimize}}\ J_{s^0} \tag{3.8a}$$

subject to the unbiasedness condition

$$E\{e(t)|s^0(t)\} = \mathbf{0}, \quad \text{for all } t. \tag{3.8b}$$

Combine (3.5) and (3.7), the above optimization is equivalent to

$$\underset{H \in R^{m\times n}}{\text{minimize}}\ \text{tr}\,(HR_nH') \tag{3.9a}$$

$$\text{subject to } HA^0 = P\Lambda. \tag{3.9b}$$

This (Gauss–Markov) constrained optimization problem has a unique solution [23] as follows:

$$H = (R_n^{-1/2}A^0\Lambda^{-1}P')^{\dagger}R_n^{-1/2}. \tag{3.10}$$

Hence

$$R_n^{1/2}(HR_n^{1/2})^{\dagger} = A^0\Lambda^{-1}P'. \tag{3.11}$$

Again the column vectors of A^0 are identified by the column vectors of $R_n^{1/2}(HR_n^{1/2})^{\dagger}$. Conversely, if the column vectors of A^0 can be identified up to scale factors and a permutation by $\tilde{A}$, i.e., $\tilde{A}$ is such that

$$\tilde{A} = A^0P\Lambda \tag{3.12}$$

for some $P \in \mathbb{P}^m$ and some $\Lambda \in \mathbb{D}^m$, then by letting

$$H = (R_n^{-1/2}\hat{A})^{\dagger}R_n^{-1/2}$$

the estimation $\hat{s}(\cdot) = Hx(\cdot)$ is the minimum-variance unbiased estimate of $\Lambda^{-1}P's^0(\cdot)$.

From the above discussion, one concludes that, for a given R_n, the source signals can be blindly estimated up to scale factors and a permutation with minimum estimation variance (for any sample $s^0(\cdot)$) if and only if the column vectors of A^0 can be blindly identified up to scale factors and a permutation. Therefore, from now on, we shall concentrate on the issue of the blind identification of the channel parameter matrix A^0.

C. Channel Identification Equations

We now derive some relations among the channel parameter matrix A^0, the statistical information of the source signals $s^0(t)$, and the statistical information of the observation process $x(t)$. The issue of uniqueness is then discussed.

Theorem 3.1: Under the assumption AS1–AS4, it is necessary that $(A^0, R_{s^0}, K_{s^0}^{(k+2)}, k \in \mathbb{Z})$ satisfies the following equations (by substituting $A = A^0$, $R_s = R_{s^0}$, and $K_s^{(k+2)} = K_{s^0}^{(k+2)}$):

$$R_x = AR_sA' + R_n \tag{3.13a}$$

$$C_v^{(k+2)} = AD_vK_s^{(k+2)}A', \quad v \in \mathbb{Z}_n^k, \quad k \in \mathbb{Z} \tag{3.13b}$$

where

$\{C_v^{(k+2)}, v \in \mathbb{Z}_n^k\}$ is the set of $(k+2)$th cumulants of the observation $x(t)$,

$$D_v = \text{diag}\,(A_{(v_1,\cdot)} \blacklozenge A_{(v_2,\cdot)} \blacklozenge \cdots \blacklozenge A_{(v_k,\cdot)}), \tag{3.13c}$$

$$K_s^{(k+2)} = \text{diag}\,(\kappa_{s_1}^{(k+2)}, \kappa_{s_2}^{(k+2)}, \cdots, \kappa_{s_m}^{(k+2)}). \tag{3.13d}$$

Proof: The parameter matrix A^0 and the source signal $s^0(t)$ satisfy the following linear equation (by substituting $A = A^0$ and $s(t) = s^0(t)$):

$$x(t) = As(t) + n(t). \tag{3.14}$$

Equation (3.13a) is immediate from AS2, AS4, and (3.14). For any $v = [v_1, v_2, \cdots, v_k] \in \mathbb{Z}_n^k$, considering the (p, q)th entry of $C_v^{(k+2)}$ along with AS2, AS4, P1, P3, P4, and (3.14), we have

$$\begin{aligned}\text{cum}\,&(x_p, x_q, x_{v_1}, x_{v_2}, \cdots, x_{v_k})\\ &= \text{cum}\left(\sum_{i=1}^m a_{pi}s_i + n_p, \sum_{i=1}^m a_{qi}s_i + n_q, \sum_{i=1}^m a_{v_1i}s_i\right.\\ &\qquad \left. + n_{v_1}, \sum_{i=1}^m a_{v_2i}s_i + n_{v_2}, \cdots, \sum_{i=1}^m a_{v_ki}s_i + n_{v_k}\right)\\ &= \sum_{i=1}^m a_{pi}a_{qi}\kappa_{s_i}^{(k+2)} \prod_{j=1}^k a_{v_ji}.\end{aligned}$$

Hence,

$$\begin{aligned}C_v^{(k+2)} &= A\,\text{diag}\,(A_{(v_1,\cdot)} \blacklozenge A_{(v_2,\cdot)} \blacklozenge\\ &\qquad \cdots \blacklozenge A_{(v_k,\cdot)})K_s^{(k+2)}A'\\ &= AD_vK_s^{(k+2)}A'.\end{aligned}$$

This proves (3.13b) for every positive integer k. ■

Note that the set of output cumulant matrices R_x, $C_v^{(k+2)}$ for all $v \in \mathbb{Z}_n^k$, $k \in \mathbb{Z}$, determines the distribution of $x(t)$. Therefore, it can be argued that (3.12) contains all the information about A^0 which can be obtained from the observation $x(t)$. Henceforth, the equations in (3.13) are called the identification equations of the blind identification problem. We also refer to (3.13a) and (3.13b) as the second-order and $(k+2)$th-order identification equations, respectively.

The important issue of uniqueness has to be investigated. We approach this problem by examining the solutions of the identification equation. We first define the "solution candidates" of the identification equation.

Definition 3.1: A matrix $A \in R^{n\times m}$ is said to be a solution candidate of (3.13) with respect to the set of output cumulants R_x and $C_v^{(k+2)}$'s if there exist a positive definite $R_s \in \mathbb{D}^n$, and diagonal matrices $K_s^{(k+2)}$, $k \in \mathbb{Z}$, such that (3.13) is satisfied. Let $\mathbb{A}$ denote the set of solution candidates of the identification equations in (3.13).

It is easily shown in the following theorem that the solution of the identification equation is not unique and each solution relates to a class of solutions in a special way.

Theorem 3.2: $\mathbb{A}$ is nonempty and is closed under post-multiplication of any permutation matrix and any nonsingular diagonal matrix, i.e.,

S1) $A \in \mathbb{A} \Rightarrow AP \in \mathbb{A}$ and $A\Lambda \in \mathbb{A}$ for any $P \in \mathbb{P}^m$ and any $\Delta \in \mathbb{D}^m$, or equivalently,

S2) $A \in \mathbb{A} \Rightarrow AP\Lambda \in \mathbb{A}$ and $A\Lambda P \in \mathbb{A}$ for any $P \in \mathbb{P}^m$ and any $\Lambda \in \mathbb{D}^m$.

IV. Channel Identifiability

In this section, we discuss the issue of the channel identifiability. Theorem 3.2 states that the set of identification equations (3.12) determines the channel parameter matrix, at best, up to a permutation matrix and a diagonal matrix. This fact motivates the following definition.

Definition 4.1: A^0 is said to be identifiable if $A \in \mathbb{A}$ implies that $A = A^0 P\Lambda$ for some $P \in \mathbb{P}^m$ and some $\Lambda \in \mathbb{D}^m$.

Note that the ambiguities introduced by a permutation matrix and a diagonal matrix in the above definition do not affect the waveform-preserving blind estimation, and the relation between A and A^0 defines an equivalent relation [10]. Therefore, A^0 is identifiable if and only if $\mathbb{A}$ contains a single equivalent class.

The main theorem of this paper concerns a necessary and sufficient condition for the parameter matrix A^0 to be identifiable.

Theorem 4.1 (Main Theorem): A^0 is identifiable if and only if all of the components of the source signal $s^0(t)$ are non-Gaussian except at most one of them is Gaussian.

Remark: i) The interesting (somewhat surprising) part of the main theorem is that even if all sources are identically distributed but non-Gaussian, they can still be identified. ii) In the formulation here, we do not use the temporal information of the observation process. Indeed, matrix A^0 may be identifiable even if the source signals are all Gaussian when second-order statistics of the observation process such as correlation in different times are used [10].

Proof of Necessity: It suffices to show that if $s_1^0(t)$ and $s_2^0(t)$ are Gaussian sources, then there exists a $A \in \mathbb{A}$ which cannot be written as $A^0 P\Lambda$ for some $P \in \mathbb{P}^m$, and some nonsingular $\Lambda \in \mathbb{D}^m$. Again, denote $R_{s^0} = \text{diag}(\sigma_1^0, \cdots, \sigma_m^0)$. Let

$$Q = \text{diag}\left(\frac{1}{\sqrt{2}}\begin{bmatrix}\sqrt{\sigma_1^0} & -\sqrt{\sigma_1^0}\\ \sqrt{\sigma_2^0} & \sqrt{\sigma_2^0}\end{bmatrix}, I_{m-2}\right).$$

Let $A = A^0 Q$, $R_s = \text{diag}(1, 1, \sigma_3^0, \cdots, \sigma_m^0)$, and $K_s^{(k+2)} = K_{s^0}^{(k+2)}$ for $k \in \mathbb{Z}$. Since all higher order cumulants of $s_1^0(t)$ and $s_2^0(t)$ are zero, it is straightforward to verify that $(A, R_s, K_s^{(k+2)}, k \in \mathbb{Z})$ satisfies (3.13). Hence $A \in \mathbb{A}$. On the other hand, $Q \neq P\Lambda$ for any $P \in \mathbb{P}^m$, and any $\Lambda \in \mathbb{D}^m$. This can be seen by realizing that the necessary condition for a matrix to be represented as the multiplication of a permutation matrix and a nonsingular diagonal matrix is that there is one and only one nonzero element in each of its column vectors and each of its row vectors. Therefore, there should be no more than one Gaussian source among the input source signals for A^0 to be identifiable.

To prove the sufficient part of the theorem, we need to consider cumulants of any order. As shown in [25], there exist random variables having zero third- and fourth-order cumulants but nonzero higher order cumulants. The proof of sufficiency of the main theorem is divided into two parts: 1) the case when there exists a $k > 0$ for which all the $(k + 2)$th-order cumulants of source signals are nonzero, and 2) the case when no so such k exists. The first case is of special interest and is presented by Theorem 4.1A.

Theorem 4.1A: A^0 is identifiable if there exists a $k > 0$, for which all the $(k + 2)$th-order cumulants of the source signals are nonzero.

The identifiability of A^0 is established by considering the following matrix equations having the same structure as the identification equations given by (3.13):

$$AD_iA^t = B_i, \quad i = 1, 2, \cdots, h \tag{4.1}$$

where h is an arbitrary integer, $A \in \mathbb{R}^{n\times m}$, and D_i's are diagonal. The D_i's, can be represented by a single matrix $M \in \mathbb{R}^{m\times h}$ in a more compact form as follows:

$$M = [\text{diag}^{-1}(D_1), \text{diag}^{-1}(D_2), \cdots, \text{diag}^{-1}(D_h)]. \tag{4.2}$$

Conversely, diagonal matrices D_i's can be retrieved from M as follows:

$$D_i = \text{diag}(M_{(\cdot,i)}), \quad \text{for } i = 1, 2, \cdots, h. \tag{4.3}$$

The following lemmas are used to prove Theorem 4.1A. Their proofs are given in the Appendix.

Lemma 4.1: Let A and $B \in R^{n\times m}$ have full column rank, and let Λ_1, Λ_2, Λ_3, and Λ_4 be diagonal matrices. Assume that $\Lambda_1 > 0$ (and hence $\Lambda_2 > 0$), and that $\Lambda_1^{-1}\Lambda_3$ (and hence $\Lambda_2^{-1}\Lambda_4$) has m distinct diagonal elements. If

$$A\Lambda_1A^t = B\Lambda_2B^t \tag{4.4}$$

$$A\Lambda_3A^t = B\Lambda_4B^t \tag{4.5}$$

then

$$A = BP\Lambda \tag{4.6}$$

for some $P \in \mathbb{P}^m$ and some $\Lambda \in \mathbb{D}^m$.

Lemma 4.2: Let the pair $\{A, M\}$ satisfy (4.1). If rank (A) = rank $(M) = m$, then another pair $\{\tilde{A}, \tilde{M}\}$ is also a solution of (4.1) if and only if there exist a $P \in \mathbb{P}^m$ and a $\Lambda \in \mathbb{D}^m$ such that

$$\tilde{A} = AP\Lambda \tag{4.7}$$

and $\tilde{M}$ satisfies the following equation:

$$\tilde{D}_i = \Lambda^{-1}P^tD_iP\Lambda^{-1}, \quad i = 1, 2, \cdots, h \tag{4.8}$$

where $\tilde{M}$ and $\tilde{D}_i$'s are related by (4.2) and (4.3).

Lemma 4.2 provides the essential ground for the proof of Theorem 4.1A. Comparing matrix equations (4.1) with the kth-order identification equations in (3.13b), Lemma 4.2 can be applied directly if we can construct a matrix

M_k from the diagonal matrices $D_v K_s^{(k+2)}$'s in the same way as in (4.2), and show that rank $(M_k) = m$. This can be accomplished by relating D_v's to the Kronecker product of A's. In particular, the construction of M_k can be carried out as follows:

Let $f_k: \mathbb{Z}_n^k \to \mathbb{Z}_{n^k}$ be defined as

$$j = f_k(v) = (v_k - 1)n^{k+1} + (v_{k-1} - 1)n^{k-2} + \cdots + (v_2 - 1)n + v_1. \quad (4.9)$$

It can be shown that f_k defined by (4.9) is bijective, and whose inverse function is denoted by f_k^{-1}. Now for any $v = (v_1, v_2, \cdots, v_k) \in \mathbb{Z}_n^k$, let $j = f_k(v)$. Then define $g_j^k \in \mathbb{R}^{m \times 1}$ by

$$g_j^k = (A_{(v_1, \cdot)} \blacklozenge A_{(v_2, \cdot)} \blacklozenge \cdots \blacklozenge A_{(v_k, \cdot)})' = \text{diag}^{-1}(D_v). \quad (4.10)$$

The set of g_j^k's (or D_v's) can be represent by a single matrix $G_k \in \mathbb{R}^{m \times n^k}$ as

$$G_k = [g_1^k, g_2^k, \cdots, g_{n^k}^k]. \quad (4.11)$$

Conversely, the diagonal matrix D_v can be retrieved from G_k by

$$D_v = \text{diag}(G_{k(\cdot, j)}) \quad (4.12)$$

where $j = f_k(v)$.

As a simple illustration, consider a 3×3 matrix $A = [a_{ij}]$ with $k = 2$. Matrix G_k is constructed as follows:

$$j = 1; f_2^{-1}(1) = (1, 1); \quad g_1^2 = (A_{(1,\cdot)} \blacklozenge A_{(1,\cdot)})' = [a_{11}^2 \;\; a_{12}^2 \;\; a_{13}^2]'$$

$$j = 2; f_2^{-1}(2) = (2, 1); \quad g_2^2 = (A_{(2,\cdot)} \blacklozenge A_{(1,\cdot)})' = [a_{21}a_{11} \;\; a_{22}a_{12} \;\; a_{23}a_{13}]'$$

$$\cdots$$

$$j = 3^2; f_2^{-1}(9) = (3, 3,); \quad g_9^2 = (A_{(3,\cdot)} \blacklozenge A_{(3,\cdot)})' = [a_{31}^2 \;\; a_{32}^2 \;\; a_{33}^2]'.$$

From (4.11), we have

$$G_2 = \begin{bmatrix} a_{11}^2 & a_{21}a_{11} & a_{31}a_{11} & a_{11}a_{21} & a_{21}^2 & a_{31}a_{21} & a_{11}a_{31} & a_{21}a_{31} & a_{31}^2 \\ a_{12}^2 & a_{22}a_{12} & a_{32}a_{12} & a_{12}a_{22} & a_{22}^2 & a_{32}a_{22} & a_{12}a_{32} & a_{22}a_{32} & a_{32}^2 \\ a_{13}^2 & a_{23}a_{13} & a_{33}a_{13} & a_{13}a_{23} & a_{23}^2 & a_{33}a_{23} & a_{13}a_{33} & a_{23}a_{33} & a_{33}^2 \end{bmatrix}. \quad (4.13)$$

If we let

$$M_k = K_s^{(k+2)} G_k \quad (4.14)$$

then it is easy to show that the diagonal matrices $D_v K_s^{(k+2)}$'s given in (3.12b) are related to M_k according to (4.3). Observe that all elements in G_2 are elements of the Kronecker product $A \otimes A$, which is defined for any $A \in \mathbb{R}^{n \times m}$ as

$$A \otimes A = \begin{bmatrix} a_{11}A & a_{12}A & \cdots & a_{1m}A \\ a_{21}A & a_{22}A & \cdots & a_{2m}A \\ & \cdots & & \\ a_{n1}A & a_{n2}A & \cdots & a_{2m}A \end{bmatrix}.$$

For the definition and properties of Kronecker product, see [24, ch. 12]. With a simple comparison between the rows of G_2 and the columns of $A \otimes A$, one finds that all the rows of matrix G_2 in (4.13) are columns of $A \otimes A$. In fact, this is true for G_k in general. This property is used to prove the following lemma.

Lemma 4.3: If a matrix $A \in \mathbb{R}^{n \times m}$ is of full column rank, i.e., rank $(A) = m$, then rank $(G_k) = m$ for all $k \in \mathbb{Z}$.

Now we are ready to prove Theorem 4.1A.

Proof of Theorem 4.1A: Let the hypothesis be satisfied for some $k > 0$. Then $K_{s^0}^{(k+2)}$ is nonsingular. Let G_k^0 be the matrix constructed from A^0 according to (4.10) and (4.11). It follows from Lemma 4.3 that rank $(G_k^0) = m$, and hence rank $(K_{s^0}^{(k+2)} G_k^0) = m$. Now, let another pair (A, s) satisfies the $(k + 2)$th-order identification equations, and let $K_s^{(k+2)}$ be the $(k + 2)$th-order cumulant matrix of s. Applying Lemma 4.2 to $(A^0, K_{s^0}^{(k+2)} G_k^0)$ and $(A, K_s^{(k+2)} G_k)$, we have

$$A = A^0 P \Lambda$$

for some $P \in \mathbb{P}^m$ and some nonsingular $\Lambda \in \mathbb{D}^m$. ■

Theorem 4.1A implies that if, for example, all the third-order cumulants are nonzero, then A is identifiable. The proof of the general case requires Lemma 4.4.

Lemma 4.4: Let

$$K_{s^0}^{(k+2)} = \text{diag}(\kappa_{s_1^0}^{(k+2)}, \cdots, \kappa_{s_m^0}^{(k+2)}) \neq 0$$

for some $k > 0$. Without loss of generality, let $\kappa_{s_i^0}^{(k+2)} \neq 0$ for $i = 1, 2, \cdots, r$ where $1 \le r < m$, and $\kappa_{s_i^0}^{(k+2)} = 0$ for $r + 1 \le i \le m$. Then for any $A \in \mathbb{A}$, there exists a nonsingular $\Lambda \in \mathbb{D}^r$ such that

$$[A_{(\cdot, j_1)}, \cdots A_{(\cdot, j_r)}] = [A_{(\cdot, 1)}^0, \cdots A_{(\cdot, r)}^0]\Lambda \quad (4.15)$$

for some $(j_1, \cdots, j_r) \in \mathbb{Z}_m^r$.

Lemma 4.4 provides a partial identification result in the sense that if the kth-order cumulant of a source signal $s_i^0(t)$ is nonzero, then the ith column-vector of A^0 is identified up to a scale factor by the kth-order identification equations. We are now in the position to prove Theorem 4.1.

Proof of Theorem 4.1: Without loss of generality, assume that $s_1^0(t), \cdots, s_{m-1}^0(t)$ are non-Gaussian and $s_m^0(t)$ may or may not be Gaussian. From P4 in Section II, there exist $k_1, \cdots, k_{m-1}$ such that $\kappa_{s_i^0}^{(k+2)} \neq 0$, for $i = 1, \cdots, m-1$. Then, for any $A \in \mathcal{A}$, by Lemma 4.4, there exist $\{j_1, \cdots, j_{m-1}\}$ and a $\Lambda = \text{diag}(\lambda_1, \cdots, \lambda_{m-1}) \in \mathbb{D}^{m-1}$ such that

$$[A_{(\cdot, j_1)}, \cdots A_{(\cdot, j_{m-1})}] = [A^0_{(\cdot, 1)}, \cdots A^0_{(\cdot, r-1)}]\Lambda. \tag{4.16}$$

Hence all the columns, except the last one, of A^0 can be identified. Now consider the identifiability of the last column of A^0. Let (A, s) satisfy (3.12) and let $R_s = \text{diag}(\sigma_1, \cdots \sigma_m)$. Denoting $R_{s^0} = \text{diag}(\sigma_1^0, \cdots, \sigma_m^0)$, we then have, from (3.12a),

$$\sum_{i=1}^{m-1} \sigma_i^0 A^0_{(\cdot, i)}(A^0_{(\cdot, i)})' + \sigma_m^0 A^0_{(\cdot, m)}(A^0_{(\cdot, m)})' = \sum_{i=1}^{m} \sigma_i A_{(\cdot, i)} A'_{(\cdot, i)}. \tag{4.17}$$

Using (4.16), we have

$$\sum_{i=1}^{m-1} (\sigma_i^0 - \lambda_i^2 \sigma_{j_i}) A^0_{(\cdot, i)}(A^0_{(\cdot, i)})' + \sigma_m^0 A^0_{(\cdot, m)}(A^0_{(\cdot, m)})' = \sigma_{j_m} A_{(\cdot, j_m)} A'_{(\cdot, j_m)} \tag{4.18}$$

or in a matrix form

$$A^0 \,\text{diag}\,(\sigma_1^0 - \lambda_1^2 \sigma_{j_1}, \sigma_2^0 - \lambda_2^2 \sigma_{j_2}, \cdots, \sigma_{m-1}^0 - \lambda_{m-1}^2 \sigma_{j_{m-1}}, \sigma_m^0)(A^0)' = \sigma_{j_m} A_{(\cdot, j_m)} A'_{(\cdot, j_m)} \tag{4.19}$$

where j_m is the integer which fulfills the equation $\{j_1, \cdots, j_m\} = \{1, \cdots, m\}$. Hence, the matrix on the left-hand side of (4.19) must have rank 1. Since A^0 is of full column rank, and $\sigma_m^0 > 0$, (4.19) implies $\sigma_i^0 = \lambda_i^2 \sigma_{j_i}$, for $i = 1, 2, \cdots, m-1$. Then

$$\sigma_m^0 A^0_{(\cdot, m)}(A^0_{(\cdot, m)})' = \sigma_{j_m} A_{(\cdot, j_m)} A'_{(\cdot, j_m)}. \tag{4.20}$$

Therefore,

$$[A_{(\cdot, j_1)}, \cdots, A_{(\cdot, j_{m-1})}, A_{(\cdot, j_m)}] = A^0 \,\text{diag}\left[\lambda_1, \cdots, \lambda_{m-1}, \left(\frac{\sigma_m^0}{\sigma_{j_m}}\right)^{1/2}\right]$$

or

$$AP = A^0 \,\text{diag}\left[\lambda_1, \cdots, \lambda_{m-1}, \left(\frac{\sigma_m^0}{\sigma_{j_m}}\right)^{1/2}\right].$$

Hence $A = A^0 P' \Lambda$, where $P \in \mathcal{P}^m$, and $\Lambda = P\,\text{diag}\,[\lambda_1, \cdots, \lambda_{m-1}, (\sigma_m^0/\sigma_{j_m})^{1/2}]P' \in \mathcal{D}^m$. ■

V. An Algorithm and Examples

Two examples are presented in this section to illustrate the necessary and sufficient condition shown earlier. The algorithm presented here is somewhat restricted and the general version can be found in [12].

A. An Identifiable Case: Solving the Third-Order Identification Equations

We consider here a case where the sources are identically distributed. For simplicity, we consider the noiseless case and assume $R_s = I$. The second- and third-order identification equation is given by (3.12)

$$R_x = AA' \tag{5.1a}$$

$$C_v^{(3)} = \kappa A D_v A', \qquad v \in \mathcal{Z}_n^1 \tag{5.1b}$$

where

$$D_v = \text{diag}\,(a_{v1}, \cdots, a_{vm})$$

and κ is the third-order cumulant of the source. Solving the above two equations is equivalent to the problem of simultaneous diagonalizing R_x and $C_v^{(3)}$ by congruence transformations. The solution to this problem is known and can be found in [26, pp. 469–471]. An algorithm for finding a solution $\hat{A}$ that differs from the actual parameter matrix A by postmultiplication of some permutation and nonsingular diagonal matrices is given as follows:

An Algorithm for Identifying the Channel Parameter Matrix A

i) Compute the Schur decomposition of R_x: $R_x = [U_s U_n]\,\text{diag}\,(D^2, 0)[U_s U_n]'$.

ii) Compute $T - D^{-1}U_s$.

iii) Choose a set of real number $\{\alpha_v\}$ such that

$$C = T\left(\sum_{v=1}^{n} \alpha_v C_v^{(3)}\right) T'$$

has distinct eigenvalues.

iv) Compute the Schur decomposition of C:

$$C = Q \Sigma Q'.$$

iv) $\hat{A} = U_s D Q$.

Remark: The FOBI algorithm proposed by Cardoso [4] is equivalent to applying the simultaneous diagonalization algorithm to R_x and a fourth-order moment matrix. It fails when the fourth-order moments of the sources are identical.

We apply the above algorithm to a case involving three sources and four sensors with a channel parameter matrix

$$A = \begin{bmatrix} 0.2665 & 0.5421 & 0.3997 \\ 0.1039 & 0.5674 & 0.5076 \\ 0.9582 & 0.3029 & 0.2274 \\ 0.0110 & 0.5407 & 0.7286 \end{bmatrix}.$$

The three sources were drawn from the same one-sided exponential distribution. Clearly, the FOBI-type proposed in [4], [8] fails in this case. Applying the above algorithm, 70 000 samples are used to estimate the fourth-order cumulant matrices. At SNR = 20 dB, Table I shows

TABLE I
THE ESTIMATION MEAN AND ROOT-MEAN-SQUARE ERROR. SNR = 20 dB, 70 000 SAMPLES, 100 MONTE CARLO RUNS

Estimation Mean			Root-Mean-Square Error		
0.2639	0.5536	0.3679	0.0106	0.0287	0.0864
0.1014	0.5732	0.4853	0.0121	0.0121	0.0685
0.9589	0.2953	0.2320	0.0047	0.0277	0.0648
0.0106	0.5234	0.7475	0.0131	0.0499	0.0477

the root-mean-square error and estimation mean estimated in 100 Monte Carlo runs.

B. Unidentifiable Case: Two Gaussian Sources

The necessary condition is manifested in a simple example of two Gaussian sources. Let $s^0(t) = [(s_1^0(t) \;\; s_2^0(t)]'$ be a vector of two independent Gaussian source signals. Without loss of generality, let $R_{s^0} = I_2$. We then have

$$x(t) = A^0 s^0(t) + n(t).$$

Now for any orthogonal matrix $Q \in \mathbb{R}^{2\times 2}$, the above equation can also be written as

$$x(t) = As(t) + n(t)$$

where $A = A^0 Q'$ and $s(t) = Qs^0(t)$. Since $s^0(t)$ is a vector of two independent Gaussian signals, so is $s(t)$. This implies that the observation $x(t)$ can not specify the waveforms of the source signals uniquely.

VI. CONCLUSIONS

In this paper, we have shown that i) the optimal blind signal estimation can be obtained if and only if the channel parameter matrix can be identified up to postmultiplications of diagonal and permutation matrices; ii) the channel parameter matrix can be identified up to postmultiplications of diagonal and permutation matrices if and only if there is no more than one Gaussian source. In addition, the problem of channel identification is shown, from an algebraic point of view, to be equivalent to simultaneously diagonalizing the observation cumulant matrices. An identification algorithm is presented along with simulation examples.

APPENDIX

Proof of Lemma 4.1: Let us first show that if

$$AA' = BB' \tag{A.1}$$

then there exists an orthogonal matrix Q such that

$$A = BQ. \tag{A.2}$$

This can be shown by letting $Q = B'A(A'A)^{-1}$. Under (A.1), $Q'Q = I$. Now applying this fact to (4.4), we have

$$A\Lambda_1^{1/2} = B\Lambda_2^{1/2}Q \tag{A.3}$$

for some orthogonal matrix Q. Substituting A from (A.3) into (4.2), one obtains

$$B\Lambda_2^{1/2}Q\Lambda_1^{-1}\Lambda_3 Q'\Lambda_2^{1/2}B' = B\Lambda_4 B'.$$

Since B has full column rank, we have

$$Q\Lambda_1^{-1}\Lambda_3 Q' = \Lambda_2^{-1}\Lambda_4. \tag{A.4}$$

The left-hand side of (A.4) is a Schur decomposition [26] of the right-hand side of (A.4). Hence the diagonal elements of $\Lambda_1^{-1}\Lambda_3$ and $\Lambda_2^{-1}\Lambda_4$ are the same. Since these diagonal elements are distinct by hypothesis, the matrix Q must have the form

$$Q = P\mathring{I}$$

where $P \in \mathbb{P}^m$, and $\mathring{I} \in \mathbb{D}^m$ is a sign matrix. Substituting Q into (A.3), we have

$$A = B\Lambda_2^{1/2}P\mathring{I}\Lambda_1^{-1/2}.$$

Let $\Lambda = P'\Lambda_2^{1/2}P\mathring{I}\Lambda_1^{-1/2}$. We then have a $\Lambda \in \mathbb{D}^m$ and a $P \in \mathbb{P}^m$ such that

$$A = BP\Lambda$$

which is (4.6) ∎

Proof of Lemma 4.2: The if part is verified by substituting (4.7) and (4.8) into (4.1). For the only if part, let $(\tilde{A}, \tilde{M})$ also satisfies (4.1). We then have

$$AD_iA' = \tilde{A}\tilde{D}_i\tilde{A}', \qquad i = 1, 2, \cdots, h. \tag{A.5}$$

Since rank $(M) = m$, there exist $(m \times h)$-vectors α and γ such that

$$M\alpha = [1, \cdots, 1]; \quad M\gamma = [1, 2, \cdots, m]. \tag{A.6}$$

Consequently, from (4.3), we obtain $\Sigma_{i=1}^h \alpha_i D_i = I$ and $\Sigma_{i=1}^h \gamma_i D_i = \text{diag}(1, 2, \cdots, m)$. Denoting

$$\tilde{D}_\alpha \equiv \sum_{i=1}^h \alpha_i \tilde{D}_i \tag{A.7}$$

$$\tilde{D}_\gamma \equiv \sum_{i=1}^h \gamma_i \tilde{D}_i \tag{A.8}$$

then it follows from (A.5) that

$$\tilde{A}\tilde{D}_\alpha\tilde{A}' = AA' \tag{A.9}$$

$$\tilde{A}\tilde{D}_\gamma\tilde{A}' = A\,\text{diag}(1, 2, \cdots, m)A'. \tag{A.10}$$

Applying Lemma 4.1 to (A.9) and (A.10), we have (4.7). Finally, substituting (4.7) into (A.5), we obtain (4.8). ∎

Proof of Lemma 4.3: According to the construction of G_k defined by (4.11), it can be shown that the ith row of matrix G_k is the

$$\left(\frac{(i-1)(m^k-1)}{m-1} + 1\right)\text{th column}$$

of

$$A^{[k]} \equiv A \otimes A \otimes \cdots \otimes A \in \mathbb{R}^{n^k \times m^k}.$$

Since rank $(A^{[k]}) = (\text{rank}\,(A))^k = m^k$ ([24, p. 410]), i.e., $A^{[k]}$ is of full column rank, and G_k consists of m columns of $A^{[k]}$, it holds true that rank $(G_k) = m$. ■

Proof of Lemma 4.4: For any $A \in \mathbb{A}$, there exist a positive definite $R_s \in \mathbb{D}^n$ and a set of diagonal matrices $\{K_s^{(k+2)}, k \in \mathbb{Z}\}$ such that

$$A^0 R_{s0}(A^0)^t = AR_sA^t \tag{A.11}$$

$$A^0 D_v^0 K_{s0}^{(k+2)}(A^0)^t = AD_v K_s^{(k+2)} A^t \qquad \text{for all } v \in \mathbb{Z}_m^k. \tag{A.12}$$

Since A^0 is of full column rank, from (A.11), A is of full column rank. In addition, by Lemma 4.3, G_k^0 and G_k are of full row rank, where $G_k^0 \in \mathbb{R}^{m \times n^k}$ and $G_k \in \mathbb{R}^{m \times n^k}$ are constructed, from A^0 and A, respectively, according to (4.11). Therefore, there exists an $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha^{n^k}]^t \in \mathbb{R}^{n^k}$ such that $G_k^0 \alpha = \text{Col.}\,(1, 1, \cdots, 1)$. Similar to the procedure used in the proof of Lemma 4.2, we obtain from (A.12) that

$$A^0 K_{s0}^{(k+2)}(A^0)^t = ADK_s^{(k+2)}A^t \tag{A.13}$$

where

$$D = \sum_{v \in \mathbb{Z}_n^k} \alpha_{f(v)} D_v.$$

Then (A.13) implies that rank $(K_s^{(k+2)}) \geq$ rank $(K_{s0}^{(k+2)}) = r$. Similarly, by choosing a $\beta = [\beta_1, \beta_2, \cdots, \beta_{n^k}]^t \in \mathbb{R}^{n^k}$ such that $G_k \beta = I_m$, we have

$$A^0 D^0 K_{s0}^{(k+2)}(A^0)^t = AK_s^{(k+2)}A^t \tag{A.14}$$

where

$$D^0 = \sum_{v \in \mathbb{Z}_n^k} \beta_{f(v)} D_{v_j}^0.$$

Equation (A.14) implies that rank $(K_s^{(k+2)}) \leq$ rank $(K_{s0}^{(k+2)}) = r$. Therefore, rank $(K_s^{(k+2)}) = r$, i.e., there are exactly r nonzero diagonal elements of $K_s^{(k+2)}$. Let these nonzero diagonal elements be indexed by (j_i, j_i), $i = 1, 2, \cdots, r$. Let a permutation matrix P be such that the first r diagonal elements of $P^t K_s^{(k+2)} P$ are nonzero. In particular, we choose P such that the (j_i, j_i)th (nonzero) entry of $K_s^{(k+2)}$ are placed at the (i, i)th entry of $P^t K_s^{(k+2)} P$. From (A.12), we then have

$$A^0 D_v^0 K_{s0}^{(k+2)}(A^0)^t = APP^t D_v K_s^{(k+2)} PP^t A^t \qquad \text{for all } v \in \mathbb{Z}_m^k. \tag{A.15}$$

Since the last $(m - r)$ diagonal elements of both $D_v^0 K_{s0}^{(k+2)}$ and $P^t D_v K_s^{(k+2)} P$ are zero for all $v \in \mathbb{Z}_m^k$, (A.15) can be simplified as

$$\overline{A}^0 \overline{D}_i^0 (\overline{A}^0)^t = \overline{A}\,\overline{D}_i \overline{A}^t, \qquad i = 1, 2, \cdots, n^k \tag{A.16}$$

where $\overline{A}^0$ and $\overline{A}$ are the submatrices that consist of the first r column vectors of A^0 and AP, respectively, $\overline{D}_i^0$ and $\overline{D}_i$ are the rth-order leading block matrix (the submatrix consisting of the first r columns and the first r rows) of $D_v^0 K_{s0}^{(k+2)}$ and $P^t D_v K_s^{(k+2)} P (v = f_k^{-1}(i))$, respectively. Specifically,

$$\overline{A}^0 = [A^0_{(\cdot,1)}, \cdots A^0_{(\cdot,r)}] \tag{A.17}$$

$$\overline{A} = [A_{(\cdot,j_1)}, \cdots A_{(\cdot,j_r)}]. \tag{A.18}$$

Now define $\overline{M}_k \in \mathbb{R}^{r \times n^k}$ as

$$\overline{M}_k^0 = [\text{diag}^{-1}(\overline{D}_1^0), \text{diag}^{-1}(\overline{D}_2^0), \cdots, \text{diag}^{-1}(\overline{D}_{n^k}^0)]. \tag{A.19}$$

Since M_k^0 is of full row rank, and $\overline{M}_k^0$ consists of the first r rows of M_k^0, we have rank $(\overline{M}_k^0) = r$. By applying Lemma 4.2 to the equations in (A.16), there exist a $P' \in \mathbb{P}^r$ and a $\Lambda' \in \mathbb{D}^r$, such that

$$\overline{A} = \overline{A}^0 P' \Lambda' = \overline{A}^0 P' \Lambda' (P')^t P'$$
$$= \overline{A}^0 \Lambda P'$$

where $\Lambda = P' \Lambda' (P')^t \in \mathbb{D}^r$. With (A.17) and (A.18), we have

$$[A_{(\cdot,j_1)}, \cdots A_{(\cdot,j_r)}](P')^t = [A^0_{(\cdot,1)}, \cdots A^0_{(\cdot,r)}]\Lambda.$$

Equation (4.15) is then proven because $(P')^t$ is a permutation matrix that simply rearranges the columns of $[A_{(\cdot,j_1)}, \cdots A_{(\cdot,j_r)}]$. ■

Acknowledgment

The authors would like to thank Dr. S.U. Pillai and the reviewers for their comments and suggestions.

References

[1] B. G. Agee, S. V. Schell, and W. A. Gardner, "Spectral self-coherence restoral: A new approach to blind adaptive signal extraction to blind using antenna arrays," *Proc. IEEE*, vol. 78, pp. 753–767, Apr. 1990.

[2] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modules signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 459–472, Apr. 1983.

[3] J. Herault and C. Jutten, "Space or time adaptive signal processing by neural network models," in *Proc. AIP Conf.*, Snowbird, UT, 1986, pp. 206–211.

[4] J. Cardoso, "Source separation using higher order moments," in *Proc. IEEE ICASSP*, vol. 4, Glasgow, Scotland, May, 1989, pp. 2109–2112.

[5] P. Comon, "Separation of stochastic processes," in *Proc. Workshop Higher-Order Spectral Anal.*, 1989, pp. 174–179.

[6] P. Comon, "Independent component analysis," in *Proc. Workshop Higher-Order Spectral Anal.*, Grenoble, France, July 1991.

[7] P. Ruiz and J. L. Lacoume, "Extraction of independent sources from correlated inputs: A solution based on cumulants," in *Proc. Workshop Higher-Order Spectral Anal.*, 1989, pp. 146–151.

[8] L. Tong, V. C. Soon, Y. F. Huang, and R. Liu, "Multiple source separation in noise," presented at the 27th Annu. Allerton Conf. Commun., Cont., Computing, Urbana, IL, Sept. 27–29, 1989.

[9] L. Tong, Y. Inouye, and R. Liu, "Eigenstructure-based-identification of independent signals," in *Proc. IEEE ICASSP*, May 1991, pp. 3329–3332.

[10] L. Tong, V. Soon, Y. F. Huang, and R. Liu, "Indeterminacy and identifiability of blind identification," *IEEE Trans. Circuits Syst.*, pp. 499–509, May 1991.

[11] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, to be published.

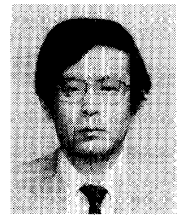
[12] L. Tong, "Blind channel identification and blind signal estimation," Ph.D. dissertation, Univ. of Notre Dame, 1990.

[13] V. C. Soon, L. Tong, Y. F. Huang, and R. Liu, "A wide-band blind

identification approach to speech acquisition using a microphone array," in *IEEE Proc. ICASSP-92*, vol. I, 1992, pp. 293-296.
[14] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization of multipath channels: A time domain approach," in *Proc. 25th Asilomar Conf.*, Oct. 1991.
[15] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 276-280, Mar. 1986.
[16] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameter via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 984-993, July 1989.
[17] Y. Bar-Ness, J. W. Carlin, and M. L. Steinberger, "Bootstrapping adaptive crosspol canceler for satellite communication," in *Proc. ICC-82*, June 1982, paper 4F.2.
[18] B. Friedlander, "A sensitive analysis of the MUSIC algorithm," in *Proc. IEEE ICASSP*, Glasgow, Scotland, May 1989.
[19] V. Soon and Y. F. Huang, "An analysis of ESPRIT under random sensor uncertainties," presented at the 28th Annu. Allerton Conf. Syst., Contr., Comput., Urbana-Champaign, IL., Oct. 1990.
[20] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, vol. 79, no. 3, pp. 278-305, Mar. 1991.
[21] R. Pan and C. L. Nikias, "Harmonic decomposition method in the cumulant domain," in *Proc. ICASSP'88*, Apr. 1988, pp. 2356-2359.
[22] M. Rosenblatt, *Stationary Sequences and Random Fields*. Boston: Birkhauser, 1985.
[23] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1968.
[24] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*. New York: Academic, 1985.
[25] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics: Distribution Theory*, vol. 1, fifth ed. London: Charles, 1987.
[26] G. H. Golub and C. Van Loan, *Matrix Computations*, second ed. Baltimore, MD: Johns Hopkins University Press, 1989.

Lang Tong (S'88-M'91) received the B.E. degree in control engineering from Tsinghua University, Beijing, China, in 1985, and the M.S. and Ph.D. degrees in electrical engineering in 1987 and 1990, respectively, from the University of Notre Dame, Notre Dame, IN.

After his position as a postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University, he joined the Department of Electrical and Computer Engineering, West Virginia University, Morgantown, WV, where he is now an Assistant Professor. He was also a Visiting Assistant at Stanford University in the summer of 1992. His research interests include signal processing, communications, and systems theory.



Yujiro Inouye (M'76) received the B.E., M.E., and D.Eng. degrees in control engineering from Osaka University, Osaka, in 1966, 1968, and 1971, respectively.

From 1971 to 1992, he was with the Department of Control Engineering, Faculty of Engineering Science, Osaka University. Currently he is an Associate Professor at the Department of Systems Engineering, Faculty of Engineering Science, Osaka University. His research interests are in signal processing, neural networks, system identification, and control engineering.

Dr. Inouye is a member of the International Neural Network Society, the Institute of Electronics, Information and Communication Engineers in Japan, the Society of Instrument and Control Engineers in Japan, and the Institute of Systems, Control, and Information Engineers in Japan.

Ruey-wen Liu (S'53-M'59-SM'78-F'91) received the B.S., M.S., and Ph.D. degrees in electrical engineering in 1954, 1955, and 1960, respectively, from the University of Illinois, Urbana.

Since 1960 he has been with the University of Notre Dame, and is now Frank M. Freimann Chair Professor of Electrical Engineering. Currently, he is on leave at the Hong Kong University of Science and Technology.

Dr. Liu was a member of the Administrative Committee (1987-1989) and is currently a Vice President of the IEEE Circuits and Systems Society. He has served as the Editor-in-Chief (1989-1991) of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, and has been on the editorial boards of several other journals. He is a member of the Overseas Advisory Committee for the *IEICE Transactions of Japan* (1991-1992). He received an Honorary Professorship from the Fudan University of Shanghai, China, in 1986, and from the Chinese Academy of Science of Beijing, China in 1988. Nonlinear circuits and systems has been his main interest. His other interests include analog fault diagnosis and feedback system design, and, recently, neural networks and blind signal processing.