Identification of multichannel MA parameters using higher-order statistics[1]

Lang Tong

Department of Electrical and Systems Engineering, University of Connecticut, Storrs, CT 06269, USA

Received 1 March 1995; revised 15 April 1996

Abstract

The identification of multichannel moving average (MA) parameter matrices $\{H(k)\}$ using fourth-order output cumulants is considered. By analyzing the eigenstructures of the cumulant matrices, it is shown that the MA parameters matrices can be identified uniquely up to a post-multiplication of monomial matrices if $H(0)$ does not have columns that are pairwise colinear and $H = [H^T(0),\ldots,H^T(L)]^T$ has full column rank. The constructive proof of this condition suggests a possible closed-form identification algorithm.

Zusammenfassung

Es wird die Identifikation der Parametermatrizen $\{H(k)\}$ von mehrkanaligen MA Modellen mit hilfe von Ausgangskumulanten vierter Ordnung untersucht. Mit einer Analyse der Eigenstruktur der Kumulantenmatrizen wird gezeigt, daß die MA Parametermatrizen bis auf eine Nachmultiplikation mit monomischen Matrizen eindeutig identifiziert werden können, falls $\{H(0)\}$ keine Spalten besitzt, die paarweise kollinear sind, und $H = [H^T(0),\ldots,H^T(L)]^T$ vollen Spaltenrang besitzt. Der konstruktive Beweis dieser Bedingung legt einen möglichen Identifkationsalgorithmus in geschlossener Form nahe.

Résumé

Ce papier s'intéresse à l'identification de matrices $H(k)$ à paramètres à moyenne mobile (MA) multicanal par des cumulants de sortie d'ordre 4. En analysant les structures propres des matrices des cumulants, on montre que les matrices à paramètres MA peuvent être identifiées de manière unique à une post-multiplication de matrices monomiales près, si $H(0)$ n'a pas de colonnes qui sont colinéaires deux à deux et si H est de plein rang. La preuve par construction de cette condition suggère un algorithme possible d'identification en boucle fermée.

1. Introduction

Much attention has been given to the problem of identifying systems using only the system output signals. Often referred to as blind system identification problems, their solutions have a wide range of applications in signal processing and communications. One such example is the blind identification of communication channels when channel probing signals are not available.

Considered in this paper is the problem of identifying parameters of the multiple-input multiple-output (MIMO) moving average (MA) process

$$x(t) = \sum_{k=0}^{L} H(k)s(t-k) + n(t), \tag{1}$$

using the fourth-order cumulants of $x(\cdot)$. By identification we mean finding $H(k)$ from the fourth-order statistics of $x(t)$ up to a post-multiplication of monomial matrix.[2]

The choice of using fourth-order cumulants is natural. First, the second-order statistics of $x(t)$ are insufficient for the identification of $\{H(k)\}$ because, for arbitrary orthogonal matrices Q, systems $\{H(k)Q\}$ and $\{H(k)\}$ have the same second-order statistics. One exception is the case when the system has single-input and multiple-outputs (SIMO) [18, 17, 9]. Second, many practical signals have symmetrical distributions whose third-order cumulants vanish. Third, the estimation of cumulants of an order higher than four becomes difficult in practice.

The basic question of MA parameter identification using fourth-order cumulants can be stated as follows: *given the complete fourth-order cumulants of the vector process* $x(\cdot)$, *to what extent can* $\{H(k)\}$ *be determined?* Perhaps the best answer in current literature is given in [6, 7] where Giannakis, Inouye and Mendel made two important contributions to the multichannel MA parameter identification problem:

(1) *Identifiability*: Under the condition that both $H(0)$ and $H(L)$ have full column ranks, it was shown that $\{H(k)\}$ is uniquely determined by the third/fourth-order cumulant matrices up to a post-multiplication of a constant monomial matrix. In other words, if both $\{H(k)\}$ and $\{H'(k)\}$ satisfy the same set of identification equations that relate the output cumulants to the MA parameters, then $H(k) = H'(k)P$ for all k, where P is a monomial matrix.

(2) *Identification*: Under the condition that both $H(0)$ and $H(L)$ have full column ranks, the identification of $\{H(0)\}$ depends on the identification of $H(0)$, and there is a closed-form relation between $H(k)$ and $H(0)$ through third/fourth-order cumulants.

To identify $H(0)$, an eigenstructure-based estimation scheme that ensures global convergence in finite steps was proposed in [13]. Later, a closed-form solution to the same set of identification equations was developed in [15, 16]. When $L = 0$, stronger results have been obtained [14]. As far as the identifiability is concerned, the solution seems to be quite satisfactory when $H(0)$ and $H(L)$ both have full column ranks.

The weakness of existing approaches come from the key assumption that *both* $H(0)$ *and* $H(L)$ have full column ranks. This assumption, shown unnecessary in this paper, is restrictive and sometimes unrealistic in practice. For example, if there are more sources than receivers, neither $H(0)$ nor $H(L)$ will have full rank. Requiring $H(L)$ to have full column rank also implies that all the single-input multiple-output sub-systems between each source and the output vector processes must have the *same* order. The limitation of existing approaches is evident when they are applied to the following simple case:

$$x(t) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} s(t) + \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 \end{bmatrix} s(t-1) + \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \end{bmatrix} s(t-2) + n(t). \tag{2}$$

Clearly the two rank conditions are not satisfied.

A more general identifiability condition is presented in this paper. It is shown that if $[H^T(0), \ldots, H^T(L)]^T$ has full column rank and the columns of $H^T(0)$ are not pairwise colinear, then MA parameters are uniquely determined by the output fourth-order cumulants. A simple verification shows that the MA process in (2) can be identified by the fourth-order cumulants. The constructive proof of this new sufficient condition suggests a possible closed-form identification approach, which is demonstrated in Section 5 using the MA process (2) as an example.

---

[2] A monomial matrix [5] is a product of a permutation matrix and a nonsingular diagonal matrix.

2. Model formulation

2.1. Notations

A few notations are necessary.

$E\{\cdot\}$ expectation operator;

$(\cdot)^{\mathrm{T}}$ transpose;

$(\cdot)^{\dagger}$ Moore–Penrose generalized inverse;

$a \parallel b$ colinear vectors: two vectors a and b are colinear if there exists an $\alpha \neq 0$ such that $a = \alpha b$;

$a \nparallel b$ a and b are not colinear;

$A \odot B$ Given an $n \times m$ matrix $A = [a_{ij}]$ and a $p \times m$ matrix $B = [b_{ij}]$,

$$
A \odot B \triangleq
\begin{bmatrix}
a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1m}b_{1m} \\
\vdots & \vdots & \vdots & \vdots \\
a_{n1}b_{11} & a_{n2}b_{12} & \cdots & a_{nm}b_{1m} \\
a_{11}b_{21} & a_{12}b_{22} & \cdots & a_{1m}b_{2m} \\
\vdots & \vdots & \vdots & \vdots \\
a_{n1}b_{21} & a_{n2}b_{22} & \cdots & a_{nm}b_{2m} \\
\vdots & \vdots & \vdots & \vdots \\
a_{11}b_{p1} & a_{12}b_{p2} & \cdots & a_{1m}b_{pm} \\
\vdots & \vdots & \vdots & \vdots \\
a_{n1}b_{p1} & a_{n2}b_{p2} & \cdots & a_{nm}b_{pm}
\end{bmatrix};
$$

$\mathrm{diag}(x)$ Given $x = [x_1, \ldots, x_m]$ or $x = [x_1, \ldots, x_m]^{\mathrm{T}}$, $\mathrm{diag}(x) = \mathrm{diag}(x_1, \ldots, x_m)$;

$\pi_{ij}(A \odot B)$ Given an $n \times m$ matrix $A = [a_{ij}]$ and $p \times m$ matrix $B = [b_{ij}]$, $\pi_{ij}(A \odot B) = [a_{i1}b_{j1}, \ldots, a_{im}b_{jm}]$.

2.2. Model and assumptions

We consider here the m-input n-output discrete-time finite-impulse response system

$$
x(t) = \sum_{k=0}^{L} H(k)s(t - k) + n(t), \tag{3}
$$

where $x(t)$ is the n-dimensional real output vector; $s(t)$ is the m-dimensional real input vector; $H(k)$ is the impulse response of the system and $n(t)$ is the n-dimensional additive noise vector.

The model Eq. (3) can be rewritten as

$$
y(t) = \mathcal{H}z(t) + w(t), \tag{4}
$$

where

$$
\mathcal{H} =
\begin{bmatrix}
H(0) & H(1) & \cdots & H(L) & & \\
 & H(0) & \cdots & H(L-1) & H(L) & \\
 & & \ddots & & & \ddots \\
 & & & H(0) & H(1) & \cdots & H(L)
\end{bmatrix}, \tag{5}
$$

$$
y(t) = \begin{bmatrix} x(t) \\ \vdots \\ x(t-L) \end{bmatrix}, \qquad
w(t) = \begin{bmatrix} n(t) \\ \vdots \\ n(t-L) \end{bmatrix}, \qquad
z(t) = \begin{bmatrix} s(t) \\ \vdots \\ s(t-2L) \end{bmatrix}. \tag{6}
$$

The block Toeplitz structure of $\mathscr{H}$ is exploited in our approach. It is noted here that if $\mathscr{H}$ has a full column rank, a closed-form identification of H using cumulant matrices of arbitrary order has been obtained. See [14].

We assume in this paper the following:

A1: The independent sources $\{s_i(t)\}$ are i.i.d. with unknown nonzero fourth-order cumulants. Without loss of generality, it is assumed that

$$J_i \stackrel{\triangle}{=} E(s_i^4(t)) - 3E(s_i^2(t)) = \pm 1, \quad \forall i. \tag{7}$$

A2: The noise process $n(\cdot)$, independent of $s(\cdot)$, is Gaussian.

A3: Both m and L are assumed to be known.

A4: Matrix $H = [H^{\mathrm{T}}(0), \ldots, H^{\mathrm{T}}(L)]^{\mathrm{T}}$ has full column rank.

A5: Column vectors of $H(0)$ are not pairwise colinear.

A5′: Matrix $H(0)$ does not have zero columns.

Remark. Assumptions A1 and A2 are fairly standard in the identification of MA processes when higher-order statistics are used. Setting all source cumulants to ± 1 only simplifies the notation. It should be noted that the identification of MA parameter matrices can only be made up to a monomial matrix. Assumption A4 is a weaker condition than that required by existing techniques [6, 14] where both $H(0)$ and $H(L)$ are assumed to have full column rank. Assumption A5 is restrictive whereas Assumption A5′ is made without loss of generality. For example, if the first column of $H(0)$ is zero, but the first column of $H(1)$ is nonzero, defining $s_1'(t) = s_1(t-1)$, an equivalent system satisfying A5′ is obtained.

3. Identification equations

Cumulant-based identification equations of arbitrary order were obtained by Swami et al. in [11, 12] using Kronecker products. Although compact and complete, the Kronecker-product form of identification equations involves matrices of high dimensions. This presents a formidable task of solving these equations. In addition, it is difficult to exploit certain symmetrical properties of cumulants. Following the approach given in [6, 8], a set of cumulant matrices with the same dimension as the covariance matrix of $x(t)$ is considered. The advantage of using these cumulant matrices is that their symmetrical properties can be exploited more conveniently, as demonstrated in our approach.

3.1. Cumulant matrices

Consider n random variables $y_1, \ldots, y_n$. A kth-order cumulant, when it exists, is denoted by

$$\mathrm{cum}(\overbrace{y_1, \ldots, y_1}^{q_1 \text{ times}}, \ldots, \overbrace{y_n, \ldots, y_n}^{q_n \text{ times}}), \tag{8}$$

where $q_1 + \cdots + q_n = k$. For most commonly used third-order and fourth-order cumulants of zero-mean random variables $y_1, \ldots, y_n$, relations between cumulants and moments are given by

$$\mathrm{cum}(y_i, y_j, y_k) = E(y_i y_j y_k),$$

$$\mathrm{cum}(y_i, y_j, y_k, y_l) = E(y_i y_j y_k y_l) - E(y_i y_j)E(y_k y_l) - E(y_i y_k)E(y_j y_l) - E(y_i y_l)E(y_j y_k). \tag{9}$$

Similarly, the fourth-order cumulants of vector y can be organized to form a set of n^2 cumulant matrices. They are related to the moments by

$$C_{ij}(y) = E(yy^T y_i y_j) - E(yy^T)E(y_i y_j) - E(yy_i)E(y^T y_j) - E(yy_j)E(y^T y_i), \quad i, j = 1, 2, \ldots, n. \tag{10}$$

In the sequel, the argument y is dropped since only the output cumulants are involved.

3.2. Identification equations

First presented is the identification equation that relates the output cumulants $\{C_{ij}\}$ with the MA parameters $\{H(k)\}$.

Theorem 1 (Identification equations). *Consider the* MA *model equation* (1) *and the equivalent linear structure* (4) *under assumptions* A1 *and* A2. *The complete set of fourth-order cumulant matrices of vector* $y(t)$ *contains all the fourth-order cumulants of the process* $x(\cdot)$. *The cumulant matrices* $\{C_{ij}\}$ *of* $y(t)$ *are related to* $\mathcal{H}$ *by*

$$C_{ij} = \mathcal{H} \operatorname{diag}(\pi_{ij}(\mathcal{H} \odot \mathcal{H})) \underbrace{\operatorname{diag}(J, \ldots, J)}_{(2L+1) \text{ blocks}} \mathcal{H}^T, \quad 1 \leqslant i, j \leqslant nL, \tag{11}$$

where $J = \operatorname{diag}(J_1, \ldots, J_m)$.

Proof. Since the source $s(t)$ is i.i.d. and $x(\cdot)$ is an MA process with finite order L, all the cumulants involving $x(t)$ and its delays are contained in the cumulant matrices of vector $y(t)$. For the linear structure in (4), the relation between $(K+2)$th-order cumulant matrices of $y(t)$ and $\mathcal{H}$ is given in [14]. Presented here is a special case for $k = 2$. $\quad\square$

4. Identifiability: The main result

4.1. Key lemmas

Recognizing the special strucutre of identification equations in (11), we consider first a simpler problem: *Given two* $N \times N$ *matrices* F_1, F_2 *satisfying*

$$F_i = A \operatorname{diag}(d_i) A^T, \quad i = 1, 2, \tag{12}$$

where A *is an* $n \times m$ *matrix with full column rank and* d_i *are* $1 \times m$ *dimensional vectors, identify* A *from* F_i. Clearly A cannot be uniquely determined by $\{F_i\}$. Again, it is sufficient if A can be determined up to a post-multiplication of a monomial matrix. This problem is related to the problem of simultaneous diagonalization of matrices by congruence, a problem investigated extensively by mathematicians [1–4]. The following lemma provides a closed-form solution to this problem. This result was given in a slightly different form in [2] and plays a key role in the derivation of the well-known ESPRIT algorithm [10] in array signal processing.

Lemma 1. *Consider*

$$F_i = A \operatorname{diag}(d_i) A^T, \quad i = 1, 2, \tag{13}$$

where A *is an* $n \times m$ *matrix with full column rank and* d_i *are* $1 \times m$ *dimensional vectors. Let*

$$D = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} d_{11} & \cdots & d_{1m} \\ d_{21} & \cdots & d_{2m} \end{bmatrix}. \tag{14}$$

If (i) D *does not have zero columns*; (ii) *columns of* D *are not pairwise colinear, then there exists a monomial matrix* P *such that*

$$A = V[u_1, \ldots, u_m]^{-\mathrm{T}} P, \tag{15}$$

where the m columns of V are the first m left singular vectors of

$$G = [F_1, F_2], \tag{16}$$

and u_i is the ith generalized eigenvector of $\{V^{\mathrm{T}} F_1 V, V^{\mathrm{T}} F_2 V\}$.

Proof. See Appendix A. □

Remarks. The above lemma gives a closed-form identification of A from F_i. Since both V and u_i in (15) are functions of only F_i, the above lemma implies that for any $\{A', d'_i\}$ satisfying (13), there exists a monomial matrix P such that

$$A' = AP. \tag{17}$$

The next two lemmas are important when Lemma 1 is applied to obtain the main result of this paper.

Lemma 2. *If $A = [a_1, \ldots, a_m]$ is an $N \times m$ matrix that does not have* (i) *all zero columns,* (ii) *colinear columns, then there exists a $2 \times N$ matrix C such that $B = CA$ is a $2 \times m$ matrix that does not have* (i) *zero elements,* (ii) *colinear columns.*

Proof. See Appendix A. □

Lemma 3. *Let $A = [a_{ij}]$ be an $N \times m$ matrix without colinear columns and zero columns. Let $B = [b_{ij}]$ be a $p \times m$ matrix without all zero columns. Then $A \odot B$ does not have* (i) *colinear columns;* (ii) *zero columns.*

Proof. See Appendix A. □

4.2. The main theorem

The main result of this paper is given in the following identifiability theorem whose proof leads to an identification algorithm.

Theorem 2 (identifiability). *Under A1–A5, if two sets of MA and cumulant parameters $\{H(k), J\}$ and $\{H'(k), J'\}$ satisfy the same indentification equation given in* (11), *then there exists a monomial matrix P such that*

$$H(k) = H'(k)P, \quad \forall k. \tag{18}$$

Proof. To better illustrate this idea, consider the case when $L = 2$ where, from (5),

$$\mathcal{H} = \begin{bmatrix} H(0) & H(1) & H(2) & & \\ & H(0) & H(1) & H(2) & \\ & & \underbrace{H(0)}_{H} & H(1) & H(2) \\ & & & & \end{bmatrix}. \tag{19}$$

The generalization to arbitrary L is straightforward. Two cases of $H(2)$ are treated separately:

Case I. Matrix $H(2)$ does not have zero columns.

Consider C_{ij} for $1 \leqslant i \leqslant n$, $2n + 1 \leqslant j \leqslant 3n$. From (11), we have

$$C_{ij} = \underbrace{\begin{bmatrix} H(2) \\ H(1) \\ H(0) \end{bmatrix}}_{H} \mathrm{diag}(\pi_{ij}(H \odot H))J\underbrace{(H^{\mathrm{T}}(2)\,H^{\mathrm{T}}(1)\,H^{\mathrm{T}}(0))}_{H^{\mathrm{T}}}, \tag{20}$$

$$= H\,\mathrm{diag}(\pi_{ij}(H \odot H))JH^{\mathrm{t}}, \quad 1 \leqslant i \leqslant n, 2n + 1 \leqslant j \leqslant 3n. \tag{21}$$

Note that J is a diagonal matrix with diagonal entrices ± 1. If for some $1 \leqslant i, k \leqslant n$ and $2n + 1 \leqslant j, l \leqslant 3n$, matrix

$$\Pi = \begin{bmatrix} \pi_{ij}(H \odot H) \\ \pi_{kl}(H \odot H) \end{bmatrix}$$

satisfies the condition of Lemma 1, i.e., (i) Π does not have colinear columns, (ii) Π does not have zero columns, the theorem is proved, and a closed-form identification algorithm is obtained by applying Lemma 1 to $F_1 = C_{ij}$ and $F_2 = C_{kl}$. Unfortunately, A1–A5 do not ensure that such i, j, k, l exist.

Note that when $1 \leqslant i \leqslant n$, $2n + 1 \leqslant j \leqslant 3n$, $\pi_{ij}(H \odot H)$ are selected from $H(0) \odot H(2)$. With A5 and the fact that $H(2)$ does not have zero columns in this case, matrix $H(2) \odot H(0)$ does not have colinear columns and zero columns by Lemma 3. Next by Lemma 2, there is a $2 \times n^2$ matrix

$$G = \begin{bmatrix} g_{11} & \cdots & g_{1n^2} \\ g_{21} & \cdots & g_{2n^2} \end{bmatrix}, \tag{22}$$

such that

$$B = G(H(2) \odot H(0)) = \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ b_{21} & \cdots & b_{2m} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{23}$$

does not have colinear columns and $b_{ij} \neq 0$ for all i and j. Define

$$F_k = \sum_{\substack{1 \leqslant i \leqslant n \\ (2n-1) \leqslant l \leqslant 3n}} g_{k,(j-2n+(i-1)n)}C_{ij}, \quad k = 1, 2. \tag{24}$$

From (21)–(24),

$$F_i = H\,\mathrm{diag}(b_i)JH^{\mathrm{T}}, \quad i = 1, 2. \tag{25}$$

Note again that J is a diagonal matrix with entries ± 1. Applying Lemma 1 to (25), H can be obtained (up to a post multiplication of a monomial) by (15).

Case II. $H(2) = [H_l(2) \vdots \mathbf{0}]$.

Without loss of generality, consider the case when the first m_l columns of $H(2)$ are nonzero columns and the last $m_\tau = m - m_l$ columns of $H(2)$ are all zero. Matrix $H(2)$ is partitioned by $H(2) = [H_l(2) \vdots \mathbf{0}]$, and correspondingly, H, J and $\mathscr{H}$ by

$$H = (H_l \vdots H_r) = \begin{bmatrix} H_l(2) & \vdots & \mathbf{0} \\ H_l(1) & \vdots & H_r(1) \\ H_l(0) & \vdots & H_r(0) \end{bmatrix}, \qquad J = \mathrm{diag}(J_l, J_r), \tag{26}$$

$$
\mathcal{H} = \begin{bmatrix} H_l(0) \vdots H_r(0) & H_l(1) \vdots H_r(1) & H_l(2) \vdots 0 & \\ & H_l(0) \vdots H_r(0) & H_l(1) \vdots H_r(1) & H_l(2) \vdots 0 & \\ & & \underbrace{H_l(0) \vdots H_r(0)}_{H_l \quad H_r} & H_l(1) \vdots H_r(1) & H_l(2) \vdots 0 \end{bmatrix}.
$$
(27)

As in Case I, if the same set of cumulant matrices C_{ij} for $1 \leqslant i \leqslant n$, $2n+1 \leqslant j \leqslant (2+1)n$ is used, from (11),

$$
C_{ij} = H_l \operatorname{diag}(\pi_{ij}(H_l \odot H_l)) J_l H_l^{\mathsf{T}}.
$$
(28)

Following the same argument in Case I, H_l can be identified by $\hat{H}_l$ from $C_{i,j}$ using the generalized eigendecomposition up to a post-multiplication of a monomial matrix, Specifically,

$$
H_l = \hat{H}_l P \Lambda
$$
(29)

for some unknown permutation matrix P and nonsingular diagonal matrix Λ.

While P, Λ and J_l are unknown, it can be shown, by substituting (29) into (28), that

$$
E_l \stackrel{\triangle}{=} P \Lambda^4 J_l P = \hat{H}_l^{\dagger} C_{ij} (\hat{H}_l^{\mathsf{T}})^{\dagger} \operatorname{diag}^{-1}(\pi_{ij}(\hat{H}_l \odot \hat{H}_l))
$$
(30)

is known. In other words, since $\hat{H}_l$ is computed from C_{ij}, E_l can also be computed from C_{ij}.

To identify H_r, it is necessary to use other cumulant matrices. Consider next the set of cumulant matrices C_{ij} for $n+1 \leqslant i \leqslant 2n$, $2n+1 \leqslant j \leqslant 3n$, and denote

$$
K = \begin{bmatrix} H_l(2) & 0 & \\ H_l(1) & H_r(1) & H_l(2) \\ \underbrace{H_l(0)}_{H_l} & \underbrace{H_r(0)}_{H_r} & \underbrace{H_l(1)}_{\bar{H}_l} \end{bmatrix} = [H_l \quad H_r \quad \bar{H}_l].
$$
(31)

From (11), we have

$$
C_{ij} = K \operatorname{diag}(\pi_{ij}(K \odot K)) \operatorname{diag}(J_l, J_r, J_l) K^{\mathsf{T}}
$$
(32)

$$
= H_r \operatorname{diag}(\pi_{ij}(H_r \odot H_r)) J_r (H_r)^{\mathsf{T}} + H_l \operatorname{diag}(\pi_{ij}(H_l \odot H_l)) J_l (H_l)^{\mathsf{T}}
$$

$$
+ \bar{H}_l \operatorname{diag}(\pi_{ij}(\bar{H}_l \odot \bar{H}_l)) J_l \bar{H}_l^{\mathsf{T}}.
$$
(33)

Note that H_l has already been identified by $\hat{H}_l$. Substituting (29) into the above equation, it is easy to verify that

$$
H_l \operatorname{diag}(\pi_{ij}(H_l \odot H_l)) J_l (H_l)^{\mathsf{T}} = \hat{H}_l P \Lambda \operatorname{diag}(\pi_{ij}(\hat{H}_l P \Lambda \odot \hat{H}_l P \Lambda)) \Lambda P^{\mathsf{T}} \hat{H}_l^{\mathsf{T}}
$$

$$
= \hat{H}_l \underbrace{P \Lambda^4 J_l P^{\mathsf{T}}}_{E_l} \operatorname{diag}(\pi_{ij}(\hat{H}_l \odot \hat{H}_l)) \hat{H}_l^{\mathsf{T}},
$$
(34)

$$
\bar{H}_l \operatorname{diag}(\pi_{ij}(\bar{H}_l \odot \bar{H}_l)) J_l \bar{H}_l^{\mathsf{T}} = \hat{\bar{H}}_l E_l \operatorname{diag}(\pi_{ij} \hat{\bar{H}}_l \odot \hat{\bar{H}}_l)) \hat{\bar{H}}_l^{\mathsf{T}},
$$
(35)

where $\hat{\hat{H}}_l = [0^T, \hat{H}^T(2), \hat{H}^T(1)]^T$ is also known. Therefore, the last two terms of (33) involving H_l can be obtained from $\hat{H}_l$ which has already been obtained. Let

$$C'_{ij} \triangleq C_{ij} - \hat{H}_l E_l \operatorname{diag}(\pi_{ij}(\hat{H}_l \odot \hat{H}_l))\hat{H}_l^T - \hat{\hat{H}}_l E_l \operatorname{diag}(\pi_{ij}(\hat{\hat{H}}_l \odot \hat{\hat{H}}_l))\hat{\hat{H}}_l^T \tag{36}$$

$$= H_r \operatorname{diag}(\pi_{ij}(H_r \odot H_r))J_r H_r^T. \tag{37}$$

We now have a new set of identification equations involving H_r of the same form as in (11). Two observations can be made about (37):

(1) By A4, H_r has full column rank.

(2) By A5, $H_r(0)$ does not have colinear columns.

Consequently, we have the same identification problem as above, and again, two cases of $H_r(1)$ are considered.

Case II (a). $H_r(1)$ *does not have zero columns.*

In this case, the same argument as in Case I is used and both $H_r(1)$ and $H_r(0)$ are identified from C'_{ij}. The complete solution is obtained.

Case II (b). $H_r(1) = [H_{rl}(1) \vdots 0]$.

Again without loss of generality, we assume that the last m_{rr} columns of $H_r(1)$ are zero. We partition $H_r(1)$ by $H_r(1) = [H_{rl}(1) \vdots 0]$, J_r by $J_r = \operatorname{diag}(J_{rl}, J_{rr})$, and accordingly,

$$H_r = \begin{bmatrix} 0 \\ H_r(1) \\ H_r(0) \end{bmatrix} = \begin{bmatrix} 0 & \vdots & 0 \\ H_{rl}(1) & \vdots & 0 \\ \underbrace{H_{rl}(0)}_{H_{rl}} & \vdots & \underbrace{H_{rr}(0)}_{H_{rr}} \end{bmatrix} = [H_{rl} \vdots H_{rr}]. \tag{38}$$

Using the same argument presented in Case I, H_{rl} can be identified from the generalized eigenvalue decomposition of C'_{ij}. Therefore, the only parameter that has not been identified is $H_{rr}(0)$.

It is necessary to use the remaining cumulant matrices for the identification of $H_{rr}(0)$ consisting of the last m_{rr} columns of $H(0)$. Note that $H_{rr}(0)$ is the only remaining unidentified parameter matrix; all other parameters have been identified up to a monomial matrix. To use the last sets of cumulant matrices that have not been used previously, let

$$X = \begin{bmatrix} H_l(0) \vdots H_{rl}(0) \vdots H_{rr}(0) & H_l(1) \vdots H_{rl}(1) \vdots 0 & H_l(2) \\ & H_l(0) \vdots H_{rl}(0) \vdots H_{rr}(0) & H_l(1) \\ & & H_l(0) \end{bmatrix}. \tag{39}$$

Obtained from (11) with $1 \leqslant i, j \leqslant n$,

$$C_{ij} = X \operatorname{diag}(\pi_{ij}(X \odot X)) \operatorname{diag}(J, J, J_l) X^T$$

$$= \begin{pmatrix} H_{rr}(0) \\ 0 \end{pmatrix} \operatorname{diag}(\pi_{ij}(H_{rr}^T(0) \odot H_{rr}(0)))J_{rr}(H_{rr}(0)\ 0) + \text{identified parameters}. \tag{40}$$

Since $H_{rr}(0)$ has full column rank, $H_{rr}(0)$ can now be identified using (40). Thus the proof is now complete.
□

5. A numerical example

To illustrate the identification algorithm outlined in the proof of Theorem 2, we consider the example given in (2). We assume here that the fourth-order cumulant matrices of the observation process $x(t)$ are known. This example is the Case II in the proof, and

$$
\mathcal{H} = \begin{bmatrix}
1 & 1 & 1 & 1 & \vdots & 0 & 1 & 1 & 2 & \vdots & 1 & 1 & 0 & 0 & \vdots & & & & \\
1 & 2 & 3 & 4 & \vdots & 1 & 1 & 0 & 0 & \vdots & 1 & 2 & 0 & 0 & \vdots & & & & \\
& & & & & 1 & 1 & 1 & 1 & \vdots & 0 & 1 & 1 & 2 & \vdots & 1 & 1 & 0 & 0 & \vdots \\
& & & & & 1 & 2 & 3 & 4 & \vdots & 1 & 1 & 0 & 0 & \vdots & 1 & 2 & 0 & 0 & \vdots \\
& & & & & & & & & & 1 & 1 & 1 & 1 & \vdots & 0 & 1 & 1 & 2 & \vdots & 1 & 1 & 0 & 0 \\
& & & & & & & & & & 1 & 2 & 3 & 4 & \vdots & 1 & 1 & 0 & 0 & \vdots & 1 & 2 & 0 & 0
\end{bmatrix},
$$

$$
\tag{41}
$$

$$
H = \begin{bmatrix}
1 & 1 & \vdots & 0 & 0 \\
1 & 2 & \vdots & 0 & 0 \\
0 & 1 & \vdots & 1 & 0 \\
1 & 1 & \vdots & 0 & 0 \\
1 & 1 & \vdots & 1 & 1 \\
1 & 2 & \vdots & 3 & 4
\end{bmatrix}, \quad
H_l = \begin{bmatrix}
1 & 1 \\
1 & 2 \\
0 & 1 \\
1 & 1 \\
1 & 1 \\
1 & 2
\end{bmatrix}, \quad
H_r = \begin{bmatrix}
0 & 0 \\
0 & 0 \\
1 & 2 \\
0 & 0 \\
1 & 1 \\
3 & 4
\end{bmatrix}, \quad
\bar{H}_l = \begin{bmatrix}
0 & 0 \\
0 & 0 \\
1 & 1 \\
1 & 2 \\
0 & 1 \\
1 & 1
\end{bmatrix}. \tag{42}
$$

It is clear that A4 is satisfied since H has full column rank. Assume in this example that (unknown) $J = \text{diag}(1, 1, -1, -1)$.

To identify $\{H(k)\}$, or equivalently, H from the cumulant matrices of $y(t)$, we follow the Case II in the proof of the main theorem.

Step I. Compute C_{15} and C_{16}.

From (11),

$$
C_{15} = \begin{bmatrix}
1 & 1 \\
1 & 2 \\
0 & 1 \\
1 & 1 \\
1 & 1 \\
1 & 2
\end{bmatrix}
\begin{bmatrix}
1 & 1 & 0 & 1 & 1 & 1 \\
1 & 2 & 1 & 1 & 1 & 2
\end{bmatrix}
= \begin{bmatrix}
2 & 3 & 1 & 2 & 2 & 3 \\
3 & 5 & 2 & 3 & 3 & 5 \\
1 & 2 & 1 & 1 & 1 & 2 \\
2 & 3 & 1 & 2 & 2 & 3 \\
2 & 3 & 1 & 2 & 2 & 3 \\
3 & 5 & 2 & 3 & 3 & 5
\end{bmatrix}, \tag{43}
$$

$$
C_{16} = \begin{bmatrix}
1 & 1 \\
1 & 2 \\
0 & 1 \\
1 & 1 \\
1 & 1 \\
1 & 2
\end{bmatrix}
\begin{bmatrix}
1 & \\
& 2
\end{bmatrix}
\begin{bmatrix}
1 & 1 & 0 & 1 & 1 & 1 \\
1 & 2 & 1 & 1 & 1 & 2
\end{bmatrix}
= \begin{bmatrix}
3 & 5 & 2 & 3 & 3 & 5 \\
5 & 9 & 4 & 5 & 5 & 9 \\
2 & 4 & 2 & 2 & 2 & 4 \\
3 & 5 & 2 & 3 & 3 & 5 \\
3 & 5 & 2 & 3 & 3 & 5 \\
5 & 9 & 4 & 5 & 5 & 9
\end{bmatrix}. \tag{44}
$$

Step II. Apply Lemma 1 to find $\hat{H}_l$.
Here $F_1 = C_{15}$ and $F_2 = C_{16}$. Computing the two left eigenvectors of $[F_1, F_2]$ to form matrix

$$
V = \begin{bmatrix} 0.3243 & -0.4093 \\ 0.5606 & 0.2221 \\ 0.2363 & 0.6314 \\ 0.3243 & -0.4093 \\ 0.3243 & -0.4093 \\ 0.5606 & 0.2221 \end{bmatrix}
\tag{45}
$$

Compute next the generalized eigenvectors $\{u_1, u_2\}$ of $\{V^T F_1 V, V^T F_2 V\}$, and form $\hat{H}_l$ by

$$
\hat{H}_l = V[u_1 u_2]^{-T} = \begin{bmatrix} 0.6742 & -1.0445 \\ 1.3484 & -1.0445 \\ 0.6742 & -0.0000 \\ 0.6742 & -1.0445 \\ 0.6742 & -1.0445 \\ 1.3484 & -1.0445 \end{bmatrix}.
\tag{46}
$$

It is easy to verify that $\hat{H}_l$ differs from H_l by a monomial matrix.
Step III. Compute C_{35} and C_{36}.
Again from (11),

$$
C_{35} = \begin{bmatrix} 1 & 2 & 1 & 1 & 1 & 2 \\ 2 & 4 & 2 & 2 & 2 & 4 \\ 1 & 2 & -7 & 3 & -3 & -16 \\ 1 & 2 & 3 & 5 & 3 & 4 \\ 1 & 2 & -3 & 3 & -1 & -8 \\ 2 & 4 & -16 & 4 & -8 & -36 \end{bmatrix}, \quad
C_{36} = \begin{bmatrix} 2 & 4 & 2 & 2 & 2 & 4 \\ 4 & 8 & 4 & 4 & 4 & 8 \\ 2 & 4 & -31 & 5 & -16 & -67 \\ 2 & 4 & 5 & 7 & 4 & 7 \\ 2 & 4 & -16 & 4 & -8 & -36 \\ 4 & 8 & -67 & 7 & -36 & -145 \end{bmatrix}.
\tag{47}
$$

Step IV. Compute C'_{35} and C'_{36}.
To compute C'_{35} and C'_{36} from (36), we need to compute E_l from (30) using $\hat{H}_l$ and C_{15} (or C_{16}):

$$
E_l = \hat{H}_l^\dagger C_{ij} (\hat{H}_l^T)^\dagger \operatorname{diag}^{-1}(\pi_{ij}(\hat{H}_l \odot \hat{H}_l)) = \operatorname{diag}(4.84, 0.8403).
\tag{48}
$$

Substituting E_l and $\hat{H}_l$ into (36), we have

$$
C'_{35} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -9 & 0 & -5 & -19 \\ 0 & 0 & -5 & 0 & -3 & -11 \\ 0 & 0 & -19 & 0 & -11 & -41 \end{bmatrix}, \quad
C'_{36} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -35 & 0 & -19 & -73 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -19 & 0 & -11 & -41 \\ 0 & 0 & -73 & 0 & -41 & -155 \end{bmatrix}.
\tag{49}
$$

Step V. Compute $\hat{H}_r$ from C'_{35} and C'_{36}.

Applying Lemma 1 to C'_{35} and C'_{36}, we have

$$\hat{H}_r = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -2.7080 & -1.8708 \\ 0 & 0 \\ -1.3540 & -1.8708 \\ -5.4160 & -5.6125 \end{bmatrix}. \tag{50}$$

Finally, we have the identified parameter matrices

$$\hat{H}(0) = \begin{bmatrix} 0.6742 & -1.0445 & -1.3540 & -1.8708 \\ 1.3484 & -1.0445 & -5.4160 & -5.6125 \end{bmatrix},$$

$$\hat{H}(1) = \begin{bmatrix} 0.6742 & -0.0000 & -2.7080 & -1.8708 \\ 0.6742 & -1.0445 & 0.0000 & -0.0000 \end{bmatrix}, \tag{51}$$

$$\hat{H}(2) = \begin{bmatrix} 0.6742 & -1.0445 & 0.0000 & -0.0000 \\ 1.3484 & -1.0445 & 0.0000 & -0.0000 \end{bmatrix}.$$

Comparing with the 'true' parameter matrices $\{H(k)\}$, we have

$$H(k) = \hat{H}(k)P, \tag{52}$$

where P is a generalized permutation matrix:

$$P = \begin{bmatrix} 0 & 1.4832 & 0 & 0 \\ -0.9574 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.7385 \\ 0 & 0 & -0.5345 & 0 \end{bmatrix}. \tag{53}$$

6. Conclusion

Two contributions have been made in this paper. The first one is the establishment of a sufficient condition of identifying multichannel MA parameters that improves the existing result. The generalization obtained in this paper enables the application of HOS techniques to a much wider range of applications. The second contribution of this paper is the constructive proof of the sufficient condition. This proof suggests a possible closed-form identification algorithm although a number of tests must be designed to distinguish various possibilities. Such a test can be obtained, in theory, from the generalized eigenvalues of matrices F_i in (25). For example, we have Case II if and only if F_i has a rank less than m for all possible linear combinations of C_{ij}'s. Other sub-cases can be determined similarly. Such a test may not be practical when the cumulant matrices are estimated. This and other implementation and performance issues of algorithm are under investigation.

Appendix A. Proofs of lemmas

Proof of Lemma 1. From (13), (16), we have

$$[F_1, F_2] = A[\text{diag}(d_1)A^T, \text{diag}(d_2)A^T]. \tag{A.1}$$

If V consists of the first m left singular vectors of $[F_1, F_2]$, then the columns of V span the range space of A, and there is a nonsingular matrix T such that

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

Therefore

$$V^T F_i V = T \text{diag}(d_i) T^T. \tag{A.3}$$

Let u and λ be the generalized eigenvector and eigenvalue of $\{V^T F_1 V, V^T F_2 V\}$, respectively, then

$$T(\text{diag}(d_1 - \lambda d_2)) T^T u = 0. \tag{A.4}$$

Under the assumption that columns of D are not (pairwise) colinear, the generalized eigenvalues are all distinct. Since T is full rank, there is *one and only one* i such that $b_{1i} = \lambda b_{2i}$. Consequently, for some $\delta_i \neq 0$,

$$T^T u = \delta_i e_i,$$

where e_i is a vector with nonzero elements at the ith entry and zero elsewhere. This is true for every generalized eigenvalue–eigenvector pair.

$$T^T[u_1, \ldots, u_m] = P, \tag{A.5}$$

where P is a monomial matrix. We have (15) by solving T from (A.5) and substituting it into (A.2). $\square$

Proof of Lemma 2. It is first shown that if $A = [a_1, \ldots, a_m]$ is an $n \times m$ matrix without colinear columns, then there exists an $(n-1) \times n$ matrix C such that $B = CA$ is an $(n-1) \times m$ matrix that does not have (i) zero columns, (ii) colinear columns. Let

$$A = \begin{bmatrix} X \\ y \end{bmatrix},$$

where $X = [x_1, \ldots, x_m]$ is an $(n-1) \times m$ matrix with colinear columns and $y = [y_1, \ldots, y_m]$ is a $1 \times m$ vector. If $x_i \parallel x_j$, then

$$x_i + \varepsilon \begin{bmatrix} 0 \\ y_i \end{bmatrix} \nparallel x_j + \varepsilon \begin{bmatrix} 0 \\ y_j \end{bmatrix}, \quad \forall \varepsilon \neq 0.$$

On the other hand, if $x_i \nparallel x_j$, then there exists a $\delta_{ij} > 0$ such that

$$x_i + \varepsilon \begin{bmatrix} 0 \\ y_i \end{bmatrix} \nparallel x_j + \varepsilon \begin{bmatrix} 0 \\ y_j \end{bmatrix}, \quad \forall |\varepsilon| \leq \delta_{ij}.$$

Therefore, for any $\varepsilon \in (0, \min\{\delta_{ij}\})$

$$C_\varepsilon = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 & \varepsilon \end{bmatrix},$$

$C_\varepsilon a_i \not\Vert C_\varepsilon a_j$ for all i and j. Next, let

$$\Phi = \bigcup_i \{\varepsilon \in (0, \min\{\delta_{ij}\}):\ C_\varepsilon a_i = \mathbf{0}\}.$$

Since Φ is set of discrete points, the set $(0, \min\{\delta_{ij}\}) - \Phi$ is non-empty. Hence for any $\varepsilon \in (0, \min\{\delta_{ij}\}) - \Phi$, and all i and j, $C_\varepsilon a_i \neq \mathbf{0}$ and $C_\varepsilon a_i \not\Vert C_\varepsilon a_j$. Consequently, there exists a $2 \times n$ matrix C such that

$$B = \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ b_{21} & \cdots & b_{2m} \end{bmatrix} = CA$$

does not have (i) colinear columns, (ii) all zero columns.
To complete the proof, define

$$\Gamma = \bigcup_i \{(\alpha, \beta):\ \alpha b_{1i} + \beta b_{2i} = 0\}.$$

Then for any $d_1 \not\Vert d_2 \in R^2 - \Gamma$, matrix

$$B' = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} B$$

does not have (i) zero elements, (ii) colinear columns. $\square$

Proof of Lemma 3. By the definition of $A \odot B$, if both A and B do not have zero columns, then $A \odot B$ does not have zero columns. It is sufficient to show that $A \odot B$ does not have colinear columns. By contradiction, if the first and the second columns of $A \odot B$ are colinear, then $b_{i1} b_{i2} = 0$, for all i. Since B does not have all zero columns, without loss of generality, let $b_{11} \neq 0$ and $b_{12} \neq 0$. Hence $b_{12} = 0$ and $b_{21} = 0$. In such a case, the first and the second columns of $A \odot B$ cannot be colinear. $\square$

References

[1] F.V. Atkinson, *Multiparameter Eigenvalue Problems*, Academic Press, New York, NY, 1972, Vol. 1.

[2] Y.H. Au-Yeung, "A necessary and sufficient condition for the simultaneous diagonalization of two Hermitian matrices and its application", *Glasgow Math. J.*, Vol. 11, December 1970, pp. 81–83.

[3] R.I. Becker, "Necessary and sufficient conditions for the simultaneous diagonalization of two quadratic forms", *Linear Algebra Appl.*, Vol. 30, 1980, pp. 129–139.

[4] P. Binding, "Simultaneous diagonalization of several Hermitian matrices", *SIAM J. Matrix Anal. Appl.*, Vol. 11, No. 4, 1990.

[5] G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, Macmillan, New York, 1941.

[6] G. Giannakis, Y. Inouye and J. Mendel, "Cumulant based identification of multichannel moving-average models", *IEEE Trans. Automat. Control*, Vol. 34, No. 7, July 1989, pp. 783–787.

[7] Y. Inouye, G.B. Giannakis and J. Mendel, "Cumulant based parameter estimation of multichannel moving-average processes", in: *Proc. IEEE Internat. Conf. Acoust. Speech Signal Process.*, 1988, pp. 729–732.

[8] Y. Inouye, L. Tong and R. Liu, "Identification and estimation of multichannel non-Gaussian moving-average processes using fourth-order cumulants", in: *Proc. 28th Allerton Conf. on Communication, Control and Computing*, October 1990.

[9] E. Moulines, P. Duhamel, J.F. Cardoso and S. Mayrargue, "Subspace-methods for the blind identification of multichannel FIR filters", *IEEE Trans. Signal Process.*, Vol. SP-43, No. 2, February 1995, pp. 516–525.

[10] A. Paulraj, R. Roy and T. Kailath, "Estimation of signal parameter via rotational invariance techniques – ESPRIT", in: *Proc. 19th Asilomar Conf. on Circuits, Systems and Comp.*, Asilomar, CA, November 1985.

[11] A. Swami, G.B. Giannakis and J. Mendel, "A unified approach to modeling multichannel ARMA processes", in: *Proc. Internat. Conf. Acoust. Speech Signal Process.-89*, Vol. 4, Glasgow, Scotland, May 1989, pp. 2182–2185.

[12] A. Swami, G.B. Giannakis and S. Shamsunder, "Multichannel ARMA processes", *IEEE Trans. Signal Process.*, Vol. 42, No. 4, April 1994, pp. 898–913.

[13] L. Tong, Y. Inouye and R. Liu, "A finite-step global convergence algorithm for the parameter estimation of multichannel MA processes", *IEEE Trans. Signal Process.*, October 1992, pp. 2547–2558.

[14] L. Tong, Y. Inouye and R. Liu, "Waveform preserving blind estimation of multiple independent sources", *IEEE Trans. Signal Process.*, Vol. 41, No. 7, July 1993.

[15] L. Tong and R. Liu, "A new eigenstructure-based parameter estimation of multichannel moving average processes", in: *Proc. Internat. Conf. Acoust. Speech Signal Process.-92*, Vol. 5, San Francisco, CA, March 1992, pp. V249–V252.

[16] L. Tong and R. Liu, "A closed-form identification of multichannel moving average processes by ESPRIT", *Circuits, Systems and Signal Process.*, 1995, In press.

[17] L. Tong, G. Xu, B. Hassibi and T. Kailath, "Blind identification and equalization of multipath channels: A frequency domain approach", *IEEE Trans. Inform. Theory*, Vol. 41, No. 1, January 1995, pp. 329–334.

[18] L. Tong, G. Xu and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach", *IEEE Trans. Inform. Theory*, Vol. 40, No. 2, March 1994.