System Identification Using a Linear Combination of Cumulant Slices

José A. R. Fonollosa, *Member, IEEE,* and Josep Vidal, *Student Member, IEEE*

Abstract—**In this paper we develop a new linear approach to identify the parameters of a moving average (MA) model from the statistics of the output. First, we show that, under some constraints, the impulse response of the system can be expressed as a linear combination of cumulant slices. Then, this result is used to obtain a new well-conditioned linear method to estimate the MA parameters of a non-Gaussian process. The proposed method presents several important differences with existing linear approaches. The linear combination of slices used to compute the MA parameters can be constructed from different sets of cumulants of different orders, providing a general framework where all the statistics can be combined. Furthermore, it is not necessary to use second-order statistics (the autocorrelation slice), and therefore the proposed algorithm still provides consistent estimates in the presence of colored Gaussian noise. Another advantage of the method is that while most linear methods developed so far give totally erroneous estimates if the order is overestimated, the proposed approach does not require a previous estimation of the filter order. The simulation results confirm the good numerical conditioning of the algorithm and the improvement in performance with respect to existing methods.**

I. Introduction

SPECTRAL analysis based on higher order statistics has received great attention in recent years. The developed tools allow dealing with problems where either nonlinearities, non-Gaussianity, or nonminimum phase systems are present, and they are of great value in diverse fields such as radar, sonar, array processing, blind equalization, time-delay estimation, image and speech processing, and seismology [1], [2].

This paper presents a new approach to identify a (possibly) nonminimum phase linear system driven by i.i.d. non-Gaussian noise from just output measurements. This problem can only be solved if higher order statistics are used since, as it is well known, second-order statistics are phase-blind, i.e., they contain only magnitude information. The developed method can also be applied when the measurements are contaminated with additive colored Gaussian noise or i.i.d. noise.

Although only systems with a finite impulse response (FIR) are considered, the method we propose can be applied to the identification of the moving average (MA) parameters of an autoregressive moving average (ARMA) model using the AR compensated process (or the AR compensated cumulants). It can also be used to identify both the AR and MA parts of a noncausal ARMA model using the "double MA algorithm" [2].

Apart from the theoretically important, but impractical, closed-form solutions, the available methods for the identification of FIR systems can be classified in two categories: linear algebra solutions and optimization solutions [2]. Optimization solutions usually provide better estimates than linear methods but they require complex nonlinear optimization algorithms. Another disadvantage of the optimization-based methods is that they may converge to a local minimum. Hence, even if the accuracy of these methods is necessary for our application, it is a good idea to initialize the nonlinear iterative algorithm with a good linear algebra solution.

The developed algorithm is linear but it presents several interesting differences with respect to existing linear algebra solutions. These new characteristics can be summarized as follows.

1) The impulse response is computed as a linear combination of cumulants slices. The weights of this linear combination are the solution of an underdetermined linear system which is always well conditioned if SVD techniques are used.

2) Different sets of cumulant slices of different orders can be considered. Although it is not necessary, all the cumulants of orders 4, 3 and 2 (autocorrelation) can be combined. The methods developed by Giannakis and Mendel (GM) [6] and Tugnait (T) [8] require the use of the autocorrelation function while in our algorithm this is just an option. Consequently, it can provide consistent estimates in the presence of colored Gaussian noise of unknown power spectral density. No linear approaches published so far provide a general framework to combine all the statistics.

3) Existing linear methods for estimating the MA parameters as the $C(q, k)$ [5], GM, and T methods require an exact knowledge of the MA order and give totally erroneous estimates if the order is overestimated. The proposed algorithm does not need a previous estimation of the filter order or any other parameters as required by the bicepstral method [7].

The following section presents the basic results that allow expressing the impulse response or MA parameters

Manuscript received December 10, 1991; revised August 7, 1992. The associate editor coordinating the review of this paper and approving it for publication was Dr. Steven Elgar.

The authors are with Universitat Politècnica de Catalunya, E.T.S.E. Telecomunicació, Barcelona 08080, Spain.

IEEE Log Number 9208857.

as a linear combination of cumulant slices. Section III discusses the application of these results to the estimation of the parameters of a MA process and proposes a simple algorithm. Finally, Section IV presents some simulations results comparing the new method with existing linear approaches.

II. Cumulant Slices

In [4] it was shown that for an AR model, there is always a linear combination of 1-D cumulant slices (w-slice) that gives the impulse response of the system. This w-slice is used in [9] to develop an adaptive algorithm for estimating the AR parameters. In this paper this result is extended to MA processes and a general procedure for computing the weights of the linear combination is given.

Consider a MA (q) process $x(n)$

$$x(n) = \sum_{k=0}^{q} b(k)u(n-k) \qquad (b(0) := 1,\ b(q) \neq 0) \tag{1}$$

where $u(n)$ is a driving i.i.d. non-Gaussian sequence with $E\{u(n)\} = 0$ and mth order cumulants γ_m. The Brillinger–Rosenblatt summation formula relates the mth order cumulant of $x(n)$ to the impulse response $b(n)$ as follows:

$$C_{m,x}(i_1, i_2, \cdots, i_{m-1}) = \gamma_m \sum_{n=0}^{q} \prod_{k=0}^{m-1} b(n+i_k) \qquad i_0 = 0,\quad m \geq 2 \tag{2}$$

Our goal, the FIR system identification problem, is to recover the right side terms $b(n)$ from the left size terms $C_{m,x}(\,)$.

If we fix the indexes i_2 to i_{m-1} the resulting 1-D cumulant slice can be expressed as the cross correlation of $b(n)$ and $b(n; i_2, \cdots, i_{m-1})$

$$C_{m,x}(i, i_2, \cdots, i_{m-1}) = \sum_{n=0}^{q} b(n+i)b(n; i_2, \cdots, i_{m-1}) \tag{3}$$

where the causal sequence $b(n; i_2, \cdots, i_{m-1})$ is defined as

$$b(n; i_2, \cdots, i_{m-1}) = \gamma_m b(n) \prod_{k=2}^{m-1} b(n+i_k). \tag{4}$$

From (3) and (4) its clear that any linear combination of slices

$$C_w(i) = w_2 C_{2,x}(i) + \sum_{j=-q}^{q} w_3(j)C_{3,x}(i,j) + \sum_{j=-q}^{q}\sum_{k=-q}^{j} w_4(j,k)C_{4,x}(i,j,k) + \cdots \tag{5}$$

can also be expressed as the cross-correlation of $b(n)$ and $g_w(n)$

$$C_w(i) = \sum_{n=0}^{q} b(n+i)g_w(n) \tag{6}$$

where $g_w(n)$ is the following causal sequence:

$$g_w(n) = w_2 b(n) + \sum_{j=-q}^{q} w_3(j)b(n;j) + \sum_{j=-q}^{q}\sum_{k=-q}^{j} w_4(j,k)b(n;j,k) + \cdots. \tag{7}$$

Equation (6) shows that, for a MA model, any w-slice can be expressed as the cross correlation of two finite causal sequences $b(n)$ and $g_w(n)$. The idea behind the developed FIR system identification method is to choose the weights that give

$$g_w(n) = \delta(n) = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases}$$

since then $C_w(i)$ will be equal to $b(i)$.

Of course, we cannot use (7) to choose the weights because we do not know the sequences $b(i; \cdots)$, but we can use the following results and properties of the w-slices.

Theorem 1: If a w-slice $C_w(i)$ is causal, then $C_w(i) = C_w(0)b(i)$.

Proof: This theorem and other important properties of the cumulant slices can be proven using basic linear algebra results if we interpret the 1-D slices as vectors:

$$\boldsymbol{C}_{m,x}(i_2, \cdots) = (C_{m,x}(-q, i_2, \cdots), \cdots, C_{m,x}(0, i_2, \cdots), \cdots, C_{m,x}(q, i_2, \cdots))'$$

$$\boldsymbol{C}_w = (C_w(-q), \cdots, C_w(0), \cdots, C_w(q))'.$$

Then, (5) can be expressed in matrix notation as

$$\boldsymbol{C}_w = \boldsymbol{S}\boldsymbol{w} \tag{8}$$

where the matrix of cumulants $\boldsymbol{S}$ and the vector $\boldsymbol{w}$ are defined as

$$\boldsymbol{S} = \begin{pmatrix} C_{2,x}(-q) & C_{3,x}(-q,-q) & \cdots & C_{3,x}(-q,q) & C_{4,x}(-q,-q,-q) & \cdots & C_{4,x}(-q,q,q) & \cdots \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \\ C_{2,x}(0) & C_{3,x}(0,-q) & \cdots & C_{3,x}(0,q) & C_{4,x}(0,-q,-q) & \cdots & C_{4,x}(0,q,q) & \cdots \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \\ C_{2,x}(q) & C_{3,x}(q,-q) & \cdots & C_{3,x}(q,q) & C_{4,x}(q,-q,-q) & \cdots & C_{4,x}(q,q,q) & \cdots \end{pmatrix}$$

$$\boldsymbol{w} = (w_2 \quad w_3(-q) \cdots w_3(q) \quad w_4(-q,-q) \cdots w_4(q,q) \cdots)'.$$

Equation (7) can also be rewritten as

$$\boldsymbol{C}_w = g_w(0)\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ b(q-1) \\ b(q) \end{bmatrix} + g_w(1)\begin{bmatrix} 0 \\ \vdots \\ 1 \\ b(1) \\ \vdots \\ b(q) \\ 0 \end{bmatrix} + \cdots + g_w(q)\begin{bmatrix} 1 \\ \vdots \\ b(q-1) \\ b(q) \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

$$= \sum_{n=0}^{q} g_w(n)\boldsymbol{b}_n. \tag{9}$$

This equation clearly shows that any slice or w-slice can be expressed as the linear combination of the $q + 1$ coefficient vectors $\boldsymbol{b}_n$. Observe that these vectors are linearly independent. From (9) it is now straightforward to see that if a w-slice is causal, then $g_w(q) = g_w(q-1) = \cdots = g_w(1) = 0$, that is, $C_w(i)$ is proportional to $\boldsymbol{b}_0$.

If $\gamma_m \neq 0$ the slice $\boldsymbol{C}_{m,x}(q, 0, \cdots, 0)$ is a simple example of a causal w-slice since

$$C_{m,x}(i, q, \cdots, 0) = \begin{cases} 0 & i < 0 \\ \gamma_m b(q) \neq 0 & i = 0. \end{cases}$$

This example is also a proof of the existence of causal w-slices.

Two particular cases of this theorem will be of interest in the rest of the paper.

Corollary 1: If a w-slice $C_w(i)$ is causal and $C_w(0) = 1$, then $C_w(i) = b(i)$.

Corollary 2: If a w-slice $C_w(i)$ is equal to zero for $i \leq 0$, then $C_w(i)$ is also zero for $i > 0$.

The matrix notation used in (8) and (9) allows us to obtain other useful properties of the slices subspace.

Theorem 2: The rank of $\boldsymbol{S}$, i.e., the dimension of the slices subspace, is equal to r, the number of nonzero coefficients.

Proof: First, it is shown that the rank cannot be greater than r. Then, we find r linearly independent slices to complete the proof.

From (9) it is clear that the rank is less than or equal to the number of coefficients $q + 1$. Furthermore, if a coefficient $b(n)$ is equal to zero, then from (4) and (7) we can conclude that $g_w(n)$ will be also equal to zero for any w-slice, i.e., the number of terms in the summation (9) can be reduced from $q + 1$ to r:

$$\boldsymbol{C}_w = \sum_{n, b(n) \neq 0} g_w(n)\boldsymbol{b}_n. \tag{10}$$

Therefore, the upper bound of the rank is also reduced to r. We show now that r is also a lower bound.

Consider the slices $\boldsymbol{C}_{m,x}(i, 0, \cdots, 0)$ whose same-numbered coefficients $b(i)$ are nonzero

$$\boldsymbol{C}_{m,x}(i, 0, \cdots) = (0, \cdots, 0, C_{m,x}(-q + i, i, \cdots), \cdots, C_{m,x}(0, i, \cdots), \cdots, C_{m,x}(q, i, \cdots))^t \quad (i = 0, \cdots, q; b(i) \neq 0). \tag{11}$$

If $\gamma_m \neq 0$ then these r slices are linearly independent since they have a strictly increasing number of zero terms in the first rows

$$C_{m,x}(n, i, \cdots, 0) = \begin{cases} 0 & n < -q + i \\ \gamma_m b(i) b(q) \neq 0 & n = -q + i. \end{cases}$$

This result completes the proof and it also shows that any w-slice can be expressed as a linear combination of the above set of slices, i.e., they are a basis for the slices subspace.

Consider the submatrix $\boldsymbol{S}_u$ formed with the upper $q + 1$ rows of $\boldsymbol{S}$

$$\boldsymbol{S}_u = \begin{pmatrix} C_{2,x}(-q) & \cdots & C_{3,x}(-q, j) & \cdots & C_{4,x}(-q, j, k) & \cdots \\ \vdots & & \vdots & & \vdots & \\ C_{2,x}(0) & \cdots & C_{3,x}(0, j) & \cdots & C_{4,x}(0, j, k) & \end{pmatrix}. \tag{12}$$

With minor modifications the proof of Theorem 2 is also valid to prove that the rank of $\boldsymbol{S}_u$ is r. This is not an unexpected result since Corollary 2 tells us that $\boldsymbol{S}_u$ has the same nullspace as $\boldsymbol{S}$. Observe that we can also construct a basis for this subspace with the first $q + 1$ elements of the slices (11).

So far we have assumed that the matrix $\boldsymbol{S}$ was formed by all the cumulant slices of any order. This is the most general assumption since any linear combination of slices is considered. Hence, it is clear that Theorem 1 is still valid if only a finite number of slices (columns) is considered. The rank of the corresponding finite matrix $\boldsymbol{S}$ will also be r if a basis as the set of slices given by (11) is included.

III. Algorithm

In Section II several properties of the w-slices and slices matrix or subspace have been presented. In this section we will study the application of these theoretical results in the development of FIR system identification methods based on cumulants.

Consider the MA (q) process given by (1). Corollary 1 indicates that the coefficients $b(i)$ of the FIR filter can be recovered from the cumulants if we find a set of weights that gives a causal w-slice with $C_w(0) = 1$. The resulting system of equations can be condensed in a single matrix equation

$$\boldsymbol{S}\boldsymbol{w} = \boldsymbol{b}_0 \tag{13}$$

where the unknowns are the vector $\boldsymbol{w}$ (weights), and the last q elements of $\boldsymbol{b}_0$ (coefficients). In general, the number of unknowns is greater than the number of equations but, provided that the true statistics are used to construct $\boldsymbol{S}$, Corollary 1 assures that the solution is unique for the coefficients.

In a practical system identification problem the cumulants have to be estimated. In this case the rank of $\boldsymbol{S}$ will not be exactly r and the solution of (13) for both $\boldsymbol{w}$ and the coefficients will not be unique. If the covariances of the cumulant estimates were available, we could think in developing an optimum criterion to choose one solution for the coefficients. Even in this case the resulting algorithm might be quite involved. We will discuss here a new simple approach that does not require any knowledge of the cumulant covariances.

As we have just mentioned, in general, the solution for $\boldsymbol{w}$ will not be unique. As a first approach we could think in using only one nonzero weight. In fact, in Section II it was shown that the slices $\boldsymbol{C}_{m,x}(q, 0, \cdots, 0)$ were causal, hence, we can use only one of these slices in the linear combination to obtain

$$\boldsymbol{b}_0 = \boldsymbol{S}\boldsymbol{w} = w_m(q, 0, \cdots, 0)\boldsymbol{C}_{m,x}(q, 0, \cdots, 0)$$

$$= \frac{1}{C_{m,x}(0, q, 0, \cdots, 0)} \boldsymbol{C}_{m,x}(q, 0, \cdots, 0).$$

The resulting equation is nothing else but the $C(q, k)$ method [5]. This is an important theoretical result. Nevertheless, in practice, the $C(q, k)$ method does not provide good estimates since only one slice is considered. Instabilities are also likely to appear as the estimated cumulant $C_{m,x}(0, q, 0, \cdots, 0)$ may be close to zero.

The proposed approach to combine all the slices in the estimation is to choose the vector $\boldsymbol{w}$ with the minimum norm. This is a simple and usual solution to an underdetermined system equation and, if SVD is used, the solution will always be well conditioned. The results of the simulations confirm the good behavior of this method. In fact, this solution would be the optimum solution if we did not consider the error in the estimation of the weights and the cumulants estimates were independent and had all the same variance. The resulting algorithm is the following:

W-Slice (WS) Algorithm: The matrix equation (13) is solved in two steps.

S1) Computation of the minimum-norm weights that give a causal w-slice with $C_w(0) = 1$.

$$\boldsymbol{S}_u \boldsymbol{w} = (0, \cdots, 0, 1)^t = \boldsymbol{1}$$

$$\boldsymbol{w}_m = \boldsymbol{S}_u^{\#}\boldsymbol{1} \tag{14}$$

where $\boldsymbol{S}_u^{\#}$ denotes the pseudoinverse of $\boldsymbol{S}_u$.

S2) Computation of the coefficients as $b(i) = C_w(i)$ or in matrix notation

$$\boldsymbol{b}_0 = \boldsymbol{S}\boldsymbol{w}_m = \boldsymbol{S}\boldsymbol{S}_u^{\#}\boldsymbol{1}. \tag{15}$$

Observe that Theorem 1 is still valid if $b(q)$ is zero. Hence, if SVD is used to obtain the minimum-norm weights we do not require an exact knowledge of the order q or the number of nonzero coefficients r. The $C(q, k)$ method is also a proof that we do not require the complete basis given by (11) to obtain a consistent estimator. If the order q is known we just need to include the slice $\boldsymbol{C}_{m,x}(q, 0, \cdots, 0)$ in $\boldsymbol{S}$ to assure the consistency of the WS estimation. If we only have an upper bound $(q_{\max})$ and a lower bound $(q_{\min})$ of q then the required minimum set of slices for the WS algorithm is

$$\boldsymbol{C}_{m,x}(j, 0, \cdots, 0) \qquad q_{\min} \le j \le q_{\max}. \tag{16}$$

Observe it is not necessary to include the $\boldsymbol{C}_{m,x}(0, 0, \cdots, 0)$ slice to obtain a consistent estimator. Hence, we can still use this approach in the presence of non-Gaussian i.i.d. noise. Of course, any kind of Gaussian noise is not a problem either, since the slice $\boldsymbol{C}_{2,x}$ is not necessary in any case.

In the WS algorithm the matrix $\boldsymbol{S}_u$ may be decomposed using SVD to obtain the pseudoinverse. Another approach is to apply SVD directly to the complete matrix $\boldsymbol{S}$. We know from Theorem 2 that, when the true statistics are used, the rank of $\boldsymbol{S}$ is less than or equal to $q + 1$. When the estimated cumulants are used in $\boldsymbol{S}$ the rank is, in general, greater. Since this increase in rank is due to the noise in the estimates we can perform a previous rank reduction of $\boldsymbol{S}$ using SVD. The singular values of $\boldsymbol{S}$ can also be used to estimate the number r of nonzero coefficients.

Additional information as the variances of the estimated cumulants may be used to improve the performance of the algorithm. Combining weighted slices with the same variance seems a plausible choice in this case.

TABLE I
EXAMPLE 1, 1000 MONTE CARLO RUNS, $N = 400/1000$

Parameter	True Value	$C(q, k)$	GMT1	GMT2	WS
$b(1)$	−0.80	0.81 ± 0.14	−0.81 ± 0.14	−0.79 ± 0.06	−0.80 ± 0.03
$b(1)$	−1.25	−1.27 ± 0.23	−1.27 ± 0.23	−1.23 ± 0.11	−1.23 ± 0.11
$b(1)$	−1.40	−1.43 ± 0.22	−1.46 ± 0.26	−1.41 ± 0.29	−1.41 ± 0.16
$b(2)$	0.98	1.00 ± 0.24	0.98 ± 0.08	0.97 ± 0.11	0.99 ± 0.17
$b(1)$	−1.13	−1.15 ± 0.18	−1.19 ± 0.26	−1.12 ± 0.16	−1.13 ± 0.09
$b(2)$	0.60	0.61 ± 0.20	0.59 ± 0.06	0.59 ± 0.11	0.61 ± 0.07

IV. EXAMPLES

The objective of the simulations is to compare the performance of the proposed algorithm with existing approaches. We are also interested in checking the good behavior of the algorithm when the system order is overestimated or the observations are noisy.

As stated in the introduction, several methods have been reported in the literature for the identification of FIR systems. The GM method [6] is one of the best known and it has received the attention of authors such as Tugnait [8], [10], [11], and Porat and Friedlander [3] who have studied its performance and have proposed various modifications to the original approach. We present the results obtained with two modified versions of the GM method that will be denoted as GMT1 and GMT2. Details can be found in the Appendix. The results of the $C(q, k)$ method [5] are also included.

Another linear approach that shares various characteristics with the WS approach is the bicepstral method developed by Pan and Nikias [7]. As in the WS method, this cepstra-based method requires neither order determination nor the use of second-order statistics. Nevertheless, the bicepstral method is, in some sense, a nonparametric approach and it does not work when the zeros are close to the unit circle. In the simulations, we have restricted ourselves to the clearly parametric approaches mentioned in the previous paragraph. All of them need a prior estimation of the MA order, but no other restrictions on the MA coefficients are required.

Except in Example 5, for the WS algorithm the $2Q + 1$ third-order cumulant slices $\boldsymbol{C}_{3,x}(j)$ $(j = -Q, \cdots, Q)$ have been used to form a $2Q + 1$ by $2Q + 1$ statistics matrix $\boldsymbol{S}$.

Example 1: Porat and Friedlander studied in [3] the performance of the GM approach for different systems. In this example we consider the same MA (1) and MA (2) processes. The input is an i.i.d. (one-sided) exponentially distributed random sequence. No noise is added to the output samples used to estimate the parameters. For the WS algorithm Q is equal to the true order q. The results of a Monte Carlo simulation with 1000 runs are shown in Table I (mean ± standard deviation). The record length was 400 for the MA (1) models and 1000 for the MA (2) models.

Example 2: The first MA (1) process of Example 1 is used now to test the algorithms with noisy observations. The Monte Carlo runs were performed in the same conditions as in Example 1, but white Gaussian noise was added to the output signal. The signal-to-noise ratio was 10 dB. As expected, the variance of the estimation increased for all the algorithms. The estimation given by the GMT2 algorithm is also clearly biased since the expected value of the measured $C_2(0)$ is affected by the Gaussian noise (Table II).

Example 3: In this case we have considered the MA (5) process used in [10], [11], where several methods of MA parameter estimation were compared. Our results for the $C(q, k)$, GMT1, and GMT2 algorithms are in agreement with those presented in [10], [11] for a record length of 5120 samples. Table III shows that the proposed WS method outperforms all the other linear methods in both bias and standard deviation.

Example 4: The same MA (5) process is used now to test the WS algorithm when the order is overestimated ($Q = 7$). Since none of the other algorithms work at all in this case, only the WS algorithm results are presented in Table IV. The Monte Carlo runs were performed in the same conditions of Example 3. It can be observed that all the coefficients, including $b(6)$ and $b(7)$, are correctly estimated. Furthermore, the variance of the estimates does not increase substantially in respect to the case $Q = 5$.

Example 5: The purpose of this example is to illustrate the effect of combining different sets of cumulant slices of different orders. The system in Examples 3 and 4 is used. Three different sets of slices are considered:

S_2: $\boldsymbol{C}_{2,x}$ (1 slice)

S_3: $\boldsymbol{C}_{3,x}(j)$ $(j = -Q, \cdots, Q)$ $(2Q + 1$ slices)

S_4: $\boldsymbol{C}_{4,x}(j, 0)$ $(j = -Q, \cdots, Q)$ $(2Q + 1$ slices)

When two or more sets of different orders are used to form the statistics matrix $\boldsymbol{S}$, the normalization of each set has an important influence on the accuracy of the estimates. The study in depth of this normalization is outside the scope of this paper. As commented at the end of Section III, combining weighted slices with equal variance seems a good choice, but it requires information about the covariances of the sample cumulants. In our simulations, the signal was first normalized to obtain $C_{2,x}(0) = 1$. In addition to this normalization, the fourth-order sample cumulants were multiplied by $1/6$.

As in the previous examples, an i.i.d. exponentially distributed sequence was used as the input of the system.

TABLE II
EXAMPLE 2, 1000 MONTE CARLO RUNS, $N = 400$, SNR = 10 dB

Parameter	True Value	$C(q, k)$	GMT1	GMT2	WS
$b(1)$	−0.80	−0.82 ± 0.17	−0.82 ± 0.17	−0.76 ± 0.08	−0.80 ± 0.07

TABLE III
EXAMPLE 3, 256 MONTE CARLO RUNS, $N = 5120$, $q = 5$

Parameter	True Value	$C(q, k)$	GMT1	GMT2	WS
$b(1)$	0.100	1.59 ± 21.2	1.68 ± 2.10	0.01 ± 0.54	0.14 ± 0.31
$b(2)$	−1.870	−5.04 ± 47.7	−1.66 ± 1.79	−1.33 ± 0.51	−1.98 ± 0.56
$b(3)$	3.020	6.68 ± 77.9	1.14 ± 0.97	2.05 ± 0.70	3.05 ± 0.63
$b(4)$	−1.435	−2.76 ± 40.7	−0.47 ± 0.67	−0.72 ± 0.60	−1.43 ± 0.33
$b(5)$	0.490	0.50 ± 18.5	0.16 ± 0.26	0.22 ± 0.29	0.47 ± 0.18

TABLE IV
EXAMPLE 4, 256 MONTE CARLO RUNS, $N = 5120$

Parameter	True Value	WS ($Q = 5$)	WS ($Q = 7$)
$b(1)$	0.100	0.14 ± 0.31	0.09 ± 0.31
$b(2)$	−1.870	−1.98 ± 0.56	−1.84 ± 0.56
$b(3)$	3.020	3.05 ± 0.63	2.91 ± 0.67
$b(4)$	−1.435	−1.43 ± 0.33	−1.41 ± 0.39
$b(5)$	0.490	0.47 ± 0.18	0.46 ± 0.21
$b(6)$	0.000	—	0.00 ± 0.17
$b(7)$	0.000	—	0.00 ± 0.13

TABLE V
EXAMPLE 5, 256 MONTE CARLO RUNS, $N = 5120$, $Q = 5$

Parameter	True Value	WS (S_3)	WS ($S_3 + S_4$)	WS ($S_2 + S_3$)	WS ($S_2 + S_3 + S_4$)
$b(1)$	0.100	0.14 ± 0.31	0.09 ± 0.26	0.15 ± 0.30	−0.05 ± 0.25
$b(2)$	−1.870	−1.98 ± 0.56	−1.79 ± 0.51	−1.98 ± 0.56	−1.70 ± 0.47
$b(3)$	3.020	3.05 ± 0.63	2.92 ± 0.58	3.04 ± 0.63	2.68 ± 0.56
$b(4)$	−1.435	−1.43 ± 0.33	−1.32 ± 0.28	−1.43 ± 0.33	−1.30 ± 0.27
$b(5)$	0.490	0.47 ± 0.18	0.47 ± 0.15	0.47 ± 0.17	0.43 ± 0.13

For this distribution, fourth-order sample cumulants have a high variance. Nevertheless, the variance of the estimated parameters decreased when the slices S_2 and S_4 were added to the set S_3 (Table V).

The Computational Complexity of the Algorithms

Since different set of slices can be considered, the WS method provides flexibility also in the computational cost of its implementation. This gives some degree of freedom to choose between complexity and performance. For a large number of data points ($N > 1000$), the calculation of the sample cumulants represents the main computational load of cumulant-based methods. Therefore, it is sufficient in practice to count the number of different sample cumulants required.

In Examples 1–4, we considered all the $2Q + 1$ third-order slices of the support region. With this set of slices, the WS method requires $(Q + 1)(Q + 2)/2$ third-order cumulants. For comparison purposes, the number of cumulants and autocorrelations used by other algorithms are given in Table VI. Although with this election the number of cumulants is of order Q^2, the complexity of the WS algorithm is not much greater than the complexity of the GMT algorithms, even for a MA order as large as $q = Q = 7$.

V. CONCLUSIONS

This paper studied the properties of the cumulant slices of a MA process. We showed that all the slices lie in a subspace of dimension equal to the number of nonzero MA coefficients. The properties of the slices' subspace were used to develop a new method for the identification of FIR systems from output measurements. The proposed algorithm can be based on different sets of cumulant slices and allows obtaining consistent estimates when colored Gaussian noise and/or i.i.d. noise is present in the measurements. The simulations also showed that it gives competitive results in bias and variance, and that it has a good behavior even if the order is overestimated.

TABLE VI
NUMBER OF SAMPLE STATISTICS REQUIRED BY THE ALGORITHMS

MA Order	$C(q, k)$	GMT1	GMT2	WS	Statistics
q	$q+1$	$3q-1$	$3q$	$(q+1)(q+2)/2$	cumulants
	—	q	$q+1$	—	autocorrelations
3	4	8	9	10	cumulants
	—	3	4	—	autocorrelations
5	6	14	15	21	cumulants
	—	5	6	—	autocorrelations
7	8	20	21	36	cumulants
	—	7	8	—	autocorrelations

APPENDIX

1. The Modified GM Algorithm (GMT1)

The GM algorithm [6] was modified by Tugnait in [8], [10] to remove certain deficiencies. The modification consisted in augmenting the GM equations

$$\sum_{k=1}^{q} b^2(k)\, C_{2,x}(\tau - k) - \sum_{k=0}^{q} [\epsilon b(k)]\, C_{3,x}(\tau - k, \tau - k) = -C_{2,x}(\tau) \tag{A1}$$

with the equations

$$b(q)\, C_{3,x}(-\tau, 0) - \sum_{k=1}^{q} b^2(k)\, C_{3,x}(k - \tau, q) = C_{3,x}(-\tau, q). \tag{A2}$$

The two sets of equations are treated as a linear system with $2q + 2$ unknowns: ϵ, $b(q)$, $\epsilon b(k)$, $b^2(k)$ $(k = 1, 2, \cdots, q)$. The simulations were performed with the *Hi-Spec* implementation of this algorithm (*ma_est*) [12], that combines (A1) for $-q \le \tau \le -1$ and $q + 1 \le \tau \le 2q$, and (A2) for $1 \le |\tau| \le q$. The least squares solution of the resulting system of $4q$ equations is then used to compute the estimates as

$$b(k) = \sqrt{\frac{([\epsilon b(k)]/\epsilon)^2 + b^2(k)}{2}} \qquad (k = 1, 2, \cdots, q) \tag{A3}$$

if all the estimated $b^2(k)$ are positive. If not, the alternative solution $b(k) = [\epsilon b(k)]/\epsilon$ is used. The above set of equations are still valid in the presence of i.i.d. noise, but the performance is clearly degraded in respect to the case where all the equations are considered.

2. Modification to Reformulated GM Algorithm (GMT2)

With the aim of avoiding numerical ill conditioning the GM equations were recently reparametrized [11] as

$$\sum_{k=0}^{q} [\epsilon' b^2(k)]\, C_{2,x}(\tau - k) - \sum_{k=1}^{q} b(k)\, C_{3,x}(\tau - k, \tau - k) = C_{3,x}(\tau, \tau) \tag{A4}$$

and combined with the equations

$$[\epsilon' b(q)]\, C_{2,x}(\tau) - \sum_{k=1}^{q} b(k)\, C_{3,x}(k - \tau, q) = C_{3,x}(-\tau, q) \tag{A5}$$

to form a linear system with $2q + 2$ unknowns: ϵ', $\epsilon' b(k)$, $b(k)$, $\epsilon' b^2(k)$ $(k = 1, 2, \cdots, q)$. We have considered for the simulations the noise-free case in which (A4) is used for $-q \le \tau \le 2q$ and (A5) for $-q \le \tau \le q$. The least squares solution for $b(k)$ $(k = 1, 2, \cdots, q)$ is used directly as the estimated coefficients.

REFERENCES

[1] C. L. Nikias and M. R. Raghuveer, "Bispectrum estimation: A digital signal processing framework," *Proc. IEEE*, vol. 75, pp. 869–891, July 1987.

[2] 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.

[3] B. Porat and B. Friedlander, "Performance analysis of parameter estimation algorithms based on high-order moments," *Int. J. Adaptive Contr. Signal Processing*, vol. 3, pp. 191–229, 1989.

[4] G. B. Giannakis, "On the identifiability of non-Gaussian ARMA models using cumulants," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 18–26, 1990.

[5] G. B. Giannakis, "Cumulants: A powerful tool in signal processing," *Proc. IEEE*, vol. 75, pp. 1333–1334, 1987.

[6] G. B. Giannakis and J. M. Mendel, "Identification of nonminimum phase systems using higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 360–377, 1989.

[7] R. Pan and C. L. Nikias, "The complex cepstrum of higher order moments," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 186–205, 1988.

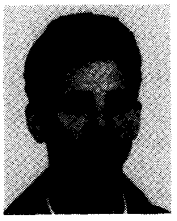
[8] J. K. Tugnait, "Approaches to FIR system identification with noisy data using higher-order statistics," in *Proc. Workshop Higher-Order Spectral Anal.*, Vail, CO, June 1989, pp. 13–18.

[9] J. A. R. Fonollosa, E. Vidal, and E. Masgrau, "Adaptive system identification based on higher-order statistics," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, May 1991, pp. 3437–3440.

[10] J. K. Tugnait, "Approaches to FIR system identification with noisy data using higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 7, pp. 1307–1317, July 1990.

[11] J. K. Tugnait, "New results on FIR system identification using higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 39, no. 10, pp. 2216–2221, Oct. 1991.

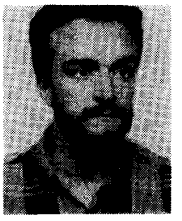
[12] *Hi-Spec*™, *Computer Software for Signal Processing with Higher-Order Spectra*, United Signals and Systems, Inc., P.O. Box 2374, Culver City, CA 90231.



José A. R. Fonollosa (S'86–M'89) was born in Barcelona, Spain, on April 19, 1963. He received the M.S. and Ph.D. degrees in electrical engineering from the Polytechnic University of Catalonia (UPC), Spain, in 1986 and 1989, respectively.

Since 1986 he has been with the Department of Signal Theory and Communications of the Polytechnic University of Catalonia, where he is now an Associate Professor. From August 1991 to July 1992 he was visiting with the Signal and Image Processing Institute, University of Southern California. His current research interests include higher order statistics and their applications in communication systems and speech processing algorithms.

Dr. Fonollosa is member of the European Speech Communication Association and the Acoustical Society of America. Currently, he is serving as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.



Josep Vidal (S'91) was born in Barcelona, Spain, on April 12, 1965. He received the Enginyer de Telecomunicació degree from the Polytechnic University of Catalonia, Barcelona, Spain, in 1989. He is currently pursuing the Ph.D. degree at the Polytechnic University of Catalonia.

From 1988 to 1989 he worked at Cognivision Research, S.A., Barcelona, as a Development Engineer. From 1989 to 1991 he joined the Signal Processing Laboratory of the Ecole Polytécnique Fédérale de Lausanne, Switzerland, as a Research Engineer. In early 1991 he came to the Department of Signal Theory and Communications, Polytechnic University of Catalonia, Barcelona. Since then he has been working towards his Ph.D. degree, granted by the Spanish Education Ministry. He is currently Assistant Lecturer in this department. His research interests are mainly in the time-series analysis, adaptive systems, nonlinear dynamics, higher order modeling and estimation theory, with applications to speech, image, communications, and biomedical signals.

Mr. Vidal is a student member of the Spanish Association of Telecommunication Engineers.