Cumulant Based Identification Approaches for Nonminimum Phase FIR Systems

Saleh A. Alshebeili, *Member, IEEE*, A. N. Venetsanopoulos, *Fellow, IEEE*, and A. Enis Çetin, *Member, IEEE*

Abstract—In this paper, recursive and least squares methods for identification of nonminimum phase linear time-invariant (NMP-LTI) FIR systems are developed. The methods utilize the second- and third-order cumulants of the output of the FIR system whose input is an independent, identically distributed (i.i.d.) non-Gaussian process. Since knowledge of the system order is of utmost importance to many system identification algorithms, new procedures for determining the order of an FIR system using only the output cumulants are also presented. To illustrate the effectiveness of our methods, various simulation examples are presented.

I. INTRODUCTION

IN this paper, we introduce new parametric approaches to the identification of nonminimum phase linear time-invariant (NMP-LTI) systems. The parametric model that we consider is an MA model described by

$$y(n) = \sum_{i=0}^{q} h(i)x(n - i) \tag{1}$$

where $y(n)$ is the output of an FIR system whose input is $x(n)$. The system input $x(n)$ is assumed to be non-Gaussian, i.i.d., random process with $E\{x(n)\} = 0$, $E\{x(n)x(n + \tau)\} = \beta_2\delta(\tau)$, and $E\{x(n)x(n + \tau_1)x(n + \tau_2)\} = \beta_3\delta(\tau_1, \tau_2)$. This model is a special case of the general ARMA model; however, it has been widely studied not only for its own interest but also due to the existence of some methods which use MA estimation techniques to identify ARMA models [7].

Most of the standard system identification algorithms available in the literature estimate only a spectrally equivalent minimum phase system because these techniques exploit only the second-order statistics which suppress all phase information of the underlying process; thus they are incapable of identifying the nonminimum phase structure of the system. A recent approach to the identification problem of NMP-LTI systems is the use of polyspectra or higher order spectra. This approach exploits the fact that

the $(k + 1)$st-order spectrum $(k > 1)$ which is the spectrum of the $(k + 1)$st-order cumulant sequence contains information regarding both the phase and magnitude of the Fourier transform of the system; and thus it has an advantage over procedures relying solely on the power spectrum in that perfect reconstruction is possible even when the system is nonminimum phase [10].

Several methods utilizing cumulant statistics for the identification of NMP-LTI systems have been proposed in the literature (see, for example, [9]). Recently, Giannakis and Mendel [5], and Tugnait [12] developed recursive and least squares identification methods by using the second-order statistics and a 1-D slice of the output cumulants. In this paper, we present new cumulant based identification methods by exploiting a much larger data set of output statistics. First, we propose a method that uses all samples of the second- and third-order cumulants to reconstruct the unknown system impulse response $h(n)$. Second, we develop cumulant based identification methods using the power spectrum and a 1-D slice of the bispectrum.

The methods presented in this paper require solution of a system of linear equations. Identification methods that exploit all the relevant statistics and use nonlinear optimization techniques are developed in the literature; see [3], [9], and [14]. Nonlinear solutions are computationally expensive and may converge to a local minimum. However, the estimates obtained via nonlinear solutions are generally better than the estimates obtained via linear solutions provided that the first are properly initialized. A good linear method can provide such an initialization that not only may lead to global convergence but may reduce the computational complexity as well.

An important problem in system identification is the determination of the system order. In this paper, this problem is also considered. New procedures for determining the order of an FIR system using only the output cumulants are proposed. In order to demonstrate the effectiveness of our methods, various simulation examples are presented.

II. A FUNDAMENTAL RELATIONSHIP

In this section, we derive a relationship between the second- and third-order spectra. This relationship is of fundamental importance because it is the basis of our system identification algorithms.

Let us assume that we are given the second- and third-order cumulants of the output sequence (in practice they can be estimated), then we can reconstruct the following 2-D sequence

$$s(\tau_1, \tau_2) = \begin{cases} c_2(\tau_1) & \text{if } \tau_1 = \tau_2 \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Let us define the sequence $g(\tau_1, \tau_2)$ such that

$$g(\tau_1, \tau_2) * s(\tau_1, \tau_2) = c_3(\tau_1, \tau_2) \quad (3)$$

where $*$ denotes convolution operation. Equation (3) is the key equation for reconstructing $h(n)$ from the second- and third-order statistics. In the frequency domain, (3) becomes

$$G(\omega_1, \omega_2) = \frac{C_3(\omega_1, \omega_2)}{C_2(\omega_1 + \omega_2)} \quad (4)$$

where $C_2(\omega)$ and $C_3(\omega_1, \omega_2)$ are the power spectrum and bispectrum, respectively. It is well known that the bispectrum is related to the system transfer function $H(\omega)$ by

$$C_3(\omega_1, \omega_2) = \beta_3 H(\omega_1) H(\omega_2) H(-\omega_1 - \omega_2) \quad (5)$$

and the power spectrum is related to the system transfer function by

$$C_2(\omega) = \beta_2 H(\omega) H(-\omega). \quad (6)$$

By substituting (5) and (6) into (4), $G(\omega_1, \omega_2)$ can be written in terms of the system transfer function as

$$G(\omega_1, \omega_2) = \frac{\epsilon H(\omega_1) H(\omega_2)}{H(\omega_1 + \omega_2)} \quad (7)$$

where $\epsilon = \beta_3/\beta_2$. Bearing in mind (4) and taking the inverse Fourier transform of both sides of (7), we obtain the relationship

$$\sum_{i=0}^{q} h(i) c_3(\tau_1 - i, \tau_2 - i)$$

$$= \sum_{i=0}^{q} \epsilon h(i) h(\tau_2 - \tau_1 + i) c_2(\tau_1 - i) \quad (8)$$

which relates the system response coefficients to the second- and third-order cumulants. Equation (8) is not new. It has been first derived in [3] and [12]. In the next section, we develop a new approach that exploits all the relevant statistics in the solution of (8); and in Section IV, we use (7) in the development of new approaches for the identification of NMP-LTI systems using slices of higher order spectra.

The basic equations used in [5] and [12] can be obtained from (8). Giannakis and Mendel's method utilizes the second-order statistics and the diagonal slice $c_3(\tau, \tau)$ of third-order cumulants. By setting $\tau_1 = \tau_2 = \tau$ in (8), we get

$$\sum_{i=0}^{q} h(i) c_3(\tau - i, \tau - i) = \sum_{i=0}^{q} \epsilon h^2(i) c_2(\tau - i). \quad (9)$$

If we multiply both sides of (9) by $1/\epsilon$, the resulting equation is called Giannakis and Mendel's equation or simply GM-equation.

Tugnait's method, on the other hand, utilizes the second-order statistics and the 1-D slice $c_3(q, \tau)$ of third-order cumulants. By setting $\tau_1 = q$ and $\tau_2 = \tau_1 + q$ in (8), we obtain Tugnait's equation (also called T-equation), that is,

$$\sum_{i=0}^{q} h(i) c_3(q, i - \tau) = \epsilon h(0) h(q) c_2(\tau) \quad (10)$$

where we have used the fact that $c_3(i, i + q) = c_3(-i, q) = c_3(q, -i)$. It is of interest to note that while both methods utilize a 1-D slice of third-order cumulants, (9) makes use of more correlation information than (10) does. Methods based on simultaneous use of both (9) and (10) are considered in [13].

III. Reconstruction from the Second- and Third-Order Cumulants

A. Least Squares (LS) Method

In this section, we develop least squares method for reconstructing the system impulse response coefficients $\{h(i)\}$ from the second- and third-order statistics of the output sequence $y(n)$. The method is based on using (8). Concatenating (8) for $\tau_1, \tau_2 \in S$, where S is a region shown in Fig. 1, we obtain the following system of linear equations:

$$d = Mr \quad (11)$$

where

$r = (h(1) \cdots h(q) \; \epsilon \; \epsilon h(1) \cdots \epsilon h(q) \; \epsilon h^2(1) \cdots \epsilon h(1) h(q) \cdots \epsilon h^2(q))^T - (q^2 + 5q + 2)/2$ column vector,

$d = (d_1 d_2)^T - 5q^2 + 4q + 1$ column vector,

$d_1 = (c_3(-q, -q) \cdots c_3(-q, 0) \; c_3(-q + 1, -q) \cdots c_3(-q + 1, 1) \cdots c_3(0, -q) \cdots c_3(0, q))$,

$d_2 = (c_3(1, -q + 1) \cdots c_3(1, q) \cdots c_3(q, 0) \cdots c_3(q, q) \; 0 \cdots 0)$

and M is a matrix of size $(5q^2 + 4q + 1) \times (q^2 + 5q + 2)/2$ whose entries are determined according to (8). The least squares solution of this overdetermined system of equations is

$$r = (M^T M)^{-1} M^T d. \quad (12)$$

In solving (11), we assume without loss of generality that the impulse response has been scaled so that $h(0) = 1$. The unknown coefficients $h(1), h(2), \cdots, h(q)$ can then be determined as the first q elements of the vector r. This would be the end of the matter when there are no measurement noise and estimation errors. If this is not the case, we propose an alternative approach which exploits all the available information provided by the vector r.
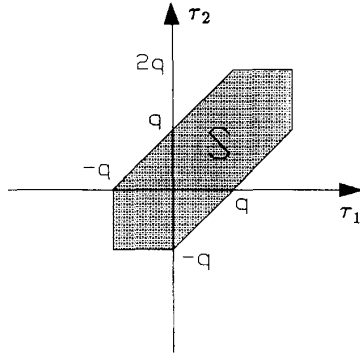
Fig. 1. The region of set S.

First, we form the matrix R from the vector r as follows:

$$
R = \begin{pmatrix}
1 & h(1) & h(2) & \cdots & h(q) \\
\epsilon & \epsilon h(1) & \epsilon h(2) & \cdots & \epsilon h(q) \\
\epsilon h(1) & \epsilon h^2(1) & \epsilon h(1)h(2) & \cdots & \epsilon h(1)h(q) \\
\epsilon h(2) & \epsilon h(2)h(1) & \epsilon h^2(2) & \cdots & \epsilon h(2)h(q) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\epsilon h(q) & \cdots & \cdots & \cdots & \epsilon h^2(q)
\end{pmatrix}.
$$

(13)

It is clear from the structure of R that its rank is one. R can be written in the following form:

$$
R = h_\epsilon h^T
$$

$$
= \begin{pmatrix} 1 \\ \epsilon \\ \epsilon h(1) \\ \vdots \\ \epsilon h(q) \end{pmatrix} (1 \quad h(1) \quad h(2) \quad \cdots \quad h(q)).
$$

(14)

The unknown impulse response sequence $h(n)$ can now be identified from the matrix R using one of the elegant techniques in numerical algebra; the singular value decomposition (SVD). That is,

$$
R = ZVU^T
$$

(15)

where V is a diagonal matrix, the diagonal elements of which are the singular values of R. The columns of the unitary matrix Z, that is, $z_1, z_2, \cdots, z_{2+q}$, are the left singular vectors of R, and the columns of the second unitary matrix U, that is, $u_1, u_2, \cdots, u_{1+q}$, are the right singular vectors of R. Since R is of rank one, it follows that there is only one nonzero singular value, whose corresponding singular vectors determine the impulse response. From the properties of the SVD, it can be shown that [8]

$$
h(n) = k_1 u_1(n) \qquad 0 \leq n \leq q
$$

(16)

and

$$
h_\epsilon(n) = k_2 z_1(n) \qquad 0 \leq n \leq q + 1
$$

(17)

where k_1 and k_2 are constants. By scaling the singular vectors z_1 and u_1 so that $z_1(0) = u_1(0) = 1$, the unknown parameters ϵ, $h(1), \cdots, h(q)$ can be determined from (16) and (17) in a straightforward manner. Alternatively, we could use the following expressions (which follow from (14) and (15)) to avoid division by a small value:

$$
h(n) = z_1(0) V(0, 0) u_1(n) \qquad 0 \leq n \leq q
$$

(18)

$$
h_\epsilon(n) = u_1(0) V(0, 0) z_1(n) \qquad 0 \leq n \leq q + 1.
$$

(19)

It is relevant at this point to mention that theoretically only one singular value of R is nonzero. In practice, due to noise and estimation errors, there may be many nonzero singular values, but only a single dominant one. In fact, when the number of realizations approaches infinity, this dominance will be more pronounced.

Extension of the above described method to the fourth-order cumulant is provided in the Appendix.

B. Uniqueness of the LS Solution

The least squares method described in the previous section yields a unique (least squares) solution if the matrix M has full rank. In this section, we show that the matrix M is of full rank. Towards this objective, we first show that the unknown parameters $\{h(i)\}$ and $\{\epsilon h(i)h(\tau_2 - \tau_1 + i)\}$, elements of the vector r, can be uniquely determined from (8) using a recursive algorithm. By setting $\tau_1 = \tau_2 = -q$ in (8), we obtain

$$
\epsilon = \frac{c_3(-q, -q)}{c_2(-q)}
$$

(20)

where we have used $h(0) = 1$. Similarly, by setting $\tau_1 = -q$, we obtain

$$
\epsilon h(\tau_2 + q) = \frac{c_3(-q, \tau_2)}{c_2(-q)} \qquad \tau_2 = -q + 1, \cdots, 0
$$

(21)

and

$$
h(\tau_2 + q) = \frac{c_3(-q, \tau_2)}{\epsilon c_2(-q)} = \frac{c_3(-q, \tau_2)}{c_3(-q, -q)}
$$

$$
\tau_2 = -q + 1, \cdots, 0.
$$

(22)

The unknown parameters $\{\epsilon h(i)h(\tau_2 - \tau_2 + i)\}$ can be obtained from (8) as follows. First, we compute $\epsilon h^2(q)$ by setting $\tau_1 = \tau_2 = 2q$, that is,

$$
\epsilon h^2(q) = \frac{h(q)c_3(q, q)}{c_2(q)}.
$$

(23)

Then, we use the following recursive formulas:

For $n = 1$ to $\lfloor q/2 \rfloor$

$$\epsilon h(n) h(\tau_2 + q)$$

$$= \frac{1}{c_2(-q)} \left(\sum_{i=0}^{n} h(i) c_3(\tau_1 - i, \tau_2 - i) \right.$$

$$\left. - \sum_{i=0}^{n-1} \epsilon h(i) h(\tau_2 - \tau_1 + i) c_2(\tau_1 - i) \right)$$

$$\tau_1 = -q + n$$

$$\text{and} \quad \tau_2 = -q + n, \cdots, 0 \quad (24)$$

$$\epsilon h(q - n) h(\tau_2 - q)$$

$$= \frac{1}{c_2(q)} \left(\sum_{i=q-n}^{q} h(i) c_3(\tau_1 - i, \tau_2 - i) \right.$$

$$\left. - \sum_{i=q-n+1}^{q} \epsilon h(i) h(\tau_2 - \tau_1 + i) c_2(\tau_1 - i) \right)$$

$$\tau_1 = 2q - n$$

$$\text{and} \quad \tau_2 = 2q - n, \cdots, 2q \quad (25)$$

end

where $\lfloor q/2 \rfloor = q/2$ if q is even, and $\lfloor q/2 \rfloor = (q - 1)/2$ is q is odd. The above recursive algorithm computes $\{h(i)\}$ and $\{\epsilon h(i) h(\tau_2 - \tau_1 + i)\}$ independently. It only uses (8) for certain values of τ_1 and τ_2. It follows then that there are $(q^2 + 5q + 2)/2$ linearly independent rows of the coefficient matrix M. Since the number of linearly independent rows equals the number of linearly independent columns, therefore the rank of the matrix M is $(q^2 + 5q + 2)/2$. Alternatively, if the rank of the matrix M is less than $(q^2 + 5q + 2)/2$, then we have more than one solution for r. But, from (20)–(25), we have only one unique solution for r. Thus, we have a contradiction since all solutions of (11) must satisfy (20)–(25); therefore, the rank of the matrix M is $(q^2 + 5q + 2)/2$.

C. Robustness to Additive Noise

In practical applications, the received signal is usually a noise corrupted version of the original one. In this section, we consider the signal model

$$d(n) = y(n) + w(n) \quad (26)$$

where $d(n)$ is the received signal, and $w(n)$ is an additive Gaussian noise.

For Gaussian processes only, cumulants of order greater than two are identically zero. This property can be exploited in estimating the third-order cumulants of noisy observations. Under the assumption that $w(n)$ is independent of $y(n)$, the third-order cumulants of $d(n)$ are equal to the third-order cumulants of $y(n)$. Symbolically,

$$c_{3_d}(\tau_1, \tau_2) = c_{3_y}(\tau_1, \tau_2). \quad (27)$$

This indicates that $c_{3_y}(\tau_1, \tau_2)$ are not affected by additive Gaussian noise. The second-order cumulants, on the other hand, appear to be affected by presence of noise, because

$$c_{2_d}(\tau) = c_{2_y}(\tau) + c_{2_w}(\tau). \quad (28)$$

If the second-order cumulants of the additive noise are nonzero only for lags in the range $|\tau| < \bar{q}$ where

$$\bar{q} = \begin{cases} (q/2) - 1 & \text{if } q \text{ is even} \\ (q - 1)/2 & \text{if } q \text{ is odd} \end{cases} \quad (29)$$

the recursive method developed in the previous section for reconstructing the unknown parameters $\{h(i)\}$ and $\{\epsilon h(i) h(\tau_2 - \tau_1 + i)\}$ from the second- and third-order cumulants will not be affected by presence of noise since it uses samples of $c_{2_d}(\tau)$ for which $\bar{q} < |\tau| \le q$. Consequently, uniqueness and consistency of the LS solution will remain unaffected if the rows of the matrix M which contain the samples of $c_{2_w}(\tau)$ are removed.

IV. RECONSTRUCTION FROM THE POWER SPECTRUM AND A 1-D SLICE OF THE BISPECTRUM

Another possible approach to reconstructing $h(n)$ from the output data $\{y(n)\}$ is to use slices of higher order spectra [1], [2]. In this section, we present recursive and least squares methods for reconstructing $h(n)$ from the power spectrum and a 1-D slice of the bispectrum. The methods utilize the 1-D slice corresponding to $\omega_1 = \omega_2$. By substituting $\omega_1 = \omega_2 = \omega$ in (7), the following relationship results:

$$G(\omega, \omega) = \frac{C_3(\omega, \omega)}{C_2(2\omega)} = \frac{\epsilon H^2(\omega)}{H(2\omega)}. \quad (30)$$

Multiplying both sides of (30) by $C_2(2\omega) H(2\omega)$, we obtain

$$H(2\omega) C_3(\omega, \omega) = \epsilon H^2(\omega) C_2(2\omega). \quad (31)$$

In the time domain, (31) takes the form

$$\sum_{i=0}^{2q} h_1(i) c'(\tau - i) = \sum_{i=0}^{2q} \epsilon h_2(i) s'(\tau - i) \quad (32)$$

where

$$h_1(n) = \mathfrak{F}^{-1}[H(2\omega)]$$

$$h_2(n) = \mathfrak{F}^{-1}[H^2(\omega)]$$

$$c'(\tau) = \mathfrak{F}^{-1}[C_3(\omega, \omega)]$$

$$s'(\tau) = \mathfrak{F}^{-1}[C_2(2\omega)].$$

Equation (32) is the basis for our recursive and least squares methods.

A. Recursive (RC) Method

Since $h(n)$ is nonzero only over $0 < n < q$, the length of the sequence $c'(\tau)$ is $4q + 1$ where $-2q \leq \tau < 2q$. Now, the parameter ϵ can be computed from (32) by setting $\tau = -2q$. That is

$$\epsilon = \frac{c'(-2q)}{c_2(-q)} \tag{33}$$

where we have used $h(0) = 1$. Once we obtain ϵ, we can proceed to compute the first parameter $h(1)$ by setting $\tau = -2q + 1$. By doing so, we obtain

$$h_2(1) = \frac{c'(-2q + 1)}{\epsilon c_2(-q)} = \frac{c'(-2q + 1)}{c'(-2q)}. \tag{34}$$

Since we know that

$$h_2(n) = h(n) * h(n) = \sum_{i=0}^{q} h(i)h(n - i) \tag{35}$$

therefore

$$h(1) = \tfrac{1}{2} h_2(1). \tag{36}$$

$c_2(\tau)$ and the phase sensitive statistic $c'(\tau)$ as follows:

For $n = 1$ to q
    term 1 $= \sum_{i=0}^{n} h_1(i)c'(-2q + n - i)$
    term 2 $= \epsilon \sum_{i=0}^{n-1} h_2(i)s'(-2q + n - i)$
    $h_2(n) = ($term 1 $-$ term 2$)/c'(-2q)$
    $h(n) = (h_2(n) - \sum_{i=1}^{n-1} h(i)h(n - i))/2$
end

The above algorithm is well conditioned in the sense that all divisions are done through $c_2(-q)$ and $c'(-2q)$ which are nonzero for an $MA(q)$ model.

B. Least Squares (LS) Method

The recursive method presented in the previous section introduces propagating errors when the initial estimates of $h(n)$ are inaccurate; in addition, it does not smooth out the effects of measurement noise. To avoid such problems, we develop in this section a least squares solution for (32). Generally, (32) can be written in a matrix form as follows:

$$d = Db \tag{37}$$

where

$d = (c'(-2q) \quad c'(-2q + 1) \quad \cdots \quad c'(2q) \quad 0 \quad \cdots \quad 0)^T - 6q + 1 \quad$ column vector

$b = (h(1) \quad \cdots \quad h(q) \quad \epsilon \quad \epsilon h_2(1) \quad \epsilon h_2(2) \quad \cdots \quad \epsilon h_2(2q))^T - 3q + 1 \quad$ column vector

$D - (6q + 1) \times (3q + 1) \quad$ matrix

$$D = \begin{bmatrix}
0 & \cdots & 0 & -s'(-2q) & 0 & \cdots & 0 \\
0 & \cdots & 0 & -s'(-2q + 1) & -s'(-2q) & \cdots & 0 \\
c'(-2q) & \cdots & 0 & -s'(-2q + 2) & -s'(-2q + 1) & \cdots & \\
c'(-2q + 1) & \cdots & 0 & -s'(-2q + 3) & -s'(-2q + 2) & \cdots & \vdots \\
\vdots & & \vdots & \vdots & \vdots & & \\
c'(2q - 2) & \cdots & c'(0) & -s'(2q) & -s'(2q - 1) & \cdots & -s'(0) \\
c'(2q - 1) & \cdots & c'(1) & 0 & -s'(2q) & \cdots & -s'(1) \\
c'(2q) & \cdots & c'(2) & 0 & 0 & \cdots & -s'(2) \\
0 & \cdots & c'(3) & 0 & 0 & \cdots & \vdots \\
\vdots & & \vdots & \vdots & \vdots & & -s'(2q - 1) \\
0 & \cdots & c'(2q) & 0 & 0 & \cdots & -s'(2q)
\end{bmatrix} \tag{38}$$

Similarly, by setting $\tau = -2q + 2$, we can compute $h_2(2)$ from which we can get $h(2)$. In general, the unknown parameters of the MA model described by (1) can be recovered recursively from the magnitude sensitive statistic

The least squares solution of this overdetermined system of equations is

$$b = (D^T D)^{-1} D^T d. \tag{39}$$

Once we obtain b, the sequence $h(n)$ can be recovered from the first q elements of b. That is, $h(n + 1) = b(n)$ for $0 \leq n \leq q - 1$; or it can be recovered from the laser $2q + 1$ elements of b using recursive and nonrecursive methods. If we define z to be as follows:

$$z = (1 \quad b(q + 1)/b(q) \quad \cdots \quad b(3q)/b(q))^T \quad (40)$$

then

$$h(n) = \mathfrak{F}^{-1}[\mathfrak{F}[z(n)]]; \quad z(0) = 1 \quad (41)$$

or recursively

$$h(n) = \left(z(n) - \sum_{i=1}^{n-1} h(i)h(n - i) \right) \bigg/ 2 \quad 1 \leq n \leq q \quad (42)$$

where $h(0) = 1$. Let us, however, note that using (40) may lead to numerically unstable solution (very high magnitude of the estimated coefficients) if $|b(q)| (= |\epsilon|)$ is to be severely underestimated.

C. Modified LS Solution

The uniqueness of the LS solution of (39) is guaranteed provided that the matrix D has full rank. However, it is not straightforward to show that D is of full rank. First, the matrix D is not in one of the standard forms that can be easily analyzed. Second, the recursive method developed in Section IV-A cannot be used in analyzing the rank of D, since it used (32) in addition to (35).

In the following, we present a modified version of the above LS method. This modification, which is very much like the approach given in [13], will enable us to analyze the rank of the coefficient matrix provided that q is the true system order. Concatenating (10) for $|\tau| \leq q$, we obtain

$$d' = D'b' \quad (43)$$

where

and $\eta = \epsilon h(0)h(q)$. Combining (37) and (43) yields

$$\begin{pmatrix} d \\ d' \end{pmatrix} = \begin{pmatrix} \mathbf{0}_{(6q+1) \times 1} & D \\ D' & \mathbf{0}_{(2q+1) \times (2q+1)} \end{pmatrix} \begin{pmatrix} \eta \\ b \end{pmatrix} \quad (45)$$

where

$\mathbf{0}_{M \times N} - M \times N$ matrix whose entries are identically zeros.

The coefficient matrix of the above system of linear equations is a full rank matrix, since the elements: η, $\{h(n)\}$, and $\{\epsilon h_2(n)\}$ of the unknown vector of (45) can be recursively determined from (10) and (32) as shown below.

First, we compute η and $h(q)$ by setting $\tau = -q$ and $\tau = q$ in (10), respectively:

$$\eta = \frac{c_3(q, q)}{c_2(q)} \quad (46)$$

$$h(q) = \frac{\eta c_2(q)}{c_3(0, q)}. \quad (47)$$

Once η is determined, the unknown MA coefficients $\{h(n)\}$ can be recovered from (10) as follows [12]:

For $n = 1$ to $\lfloor q/2 \rfloor$

$$h(n) = (\eta c_2(q - n) - \sum_{i=0}^{n-1} h(i)c_3(i + q - n, q))/c_3(q, q)$$

$$h(q - n) = (\eta c_2(q - n) - \sum_{i=q-n+1}^{q} h(i)c_3(i - q + n, q))/c_3(q, q)$$

end

Similarly, the unknown coefficients $\{\epsilon h_2(n)\}$ can be recovered from (32) as follows. First, we compute ϵ as in (33), and $\epsilon h_2(2q)$ by setting $\tau = 4q$ in (32). That is

$$\epsilon h_2(2q) = \frac{h(q)c'(2q)}{c_2(q)}. \quad (48)$$

$$b' = (\eta \quad h(1) \quad h(2) \quad \cdots \quad h(q))^T - (q + 1) \quad \text{column vector}$$

$$d' = (-c_3(q, q) - c_3(q, q - 1) \quad \cdots \quad -c_3(q, 1) \quad 0 \quad \cdots \quad 0)^T - (2q + 1) \quad \text{column vector}$$

$$D' - (2q + 1) \times (q + 1) \quad \text{matrix}$$

$$D' = \begin{bmatrix} -c_2(-q) & 0 & 0 & \cdots & 0 \\ -c_2(-q + 1) & c_3(q, q) & 0 & \cdots & 0 \\ -c_2(-q + 2) & c_3(q, q - 1) & c_3(q, q) & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -c_2(-1) & c_3(q, 2) & c_3(q, 3) & \cdots & 0 \\ -c_2(0) & c_3(q, 1) & c_3(q, 2) & \cdots & c_3(q, q) \\ \vdots & \vdots & \vdots & & \vdots \\ -c_2(q) & 0 & 0 & \cdots & c_3(q, 0) \end{bmatrix} \quad (44)$$

Then, we use the following recursion:

For $n = 1$ to q

term1 $= \Sigma_{i=0}^{n} h_1(i) c'(-2q + n - i)$

term2 $= \Sigma_{i=0}^{n-1} \epsilon h_2(i) s'(-2q + n - i)$

$\epsilon h_2(n) = (\text{term1} - \text{term2})/c_2(-q)$

term1 $= \Sigma_{i=2q-n}^{2q} h_1(i) c'(4q - n - i)$

term2 $= \Sigma_{i=2q-n-1}^{2q} \epsilon h_2(i) s'(4q - n - i)$

$\epsilon h_2(2q - n) = (\text{term1} - \text{term2})/c_2(q)$

end

From the above recursive algorithms, it follows that there are $(3q + 2)$ linearly independent rows of the matrix

$$B = \begin{pmatrix} 0_{(6q+1) \times 1} & D \\ D' & 0_{(2q+1) \times (2q+1)} \end{pmatrix} \quad (49)$$

which implies that there are $(3q + 2)$ linearly independent columns. Therefore, the matrix B is of full rank. Alternatively B can be rewritten as

$$B = \begin{pmatrix} D_{c'} & D_s \\ D' & 0_{(2q+1) \times (2q+1)} \end{pmatrix} \quad (50)$$

where $D_{c'}$ and D_s are matrices of size $(6q + 1) \times (q + 1)$ and $(6q + 1) \times (2q + 1)$, respectively. The matrices D_s and D' are full rank matrices. Therefore, the matrix B is of full rank.

D. Robustness to Additive Noise

In the presence of additive Gaussian noise, the relationships (32) and (10) take the forms

$$\sum_{i=0}^{2q} h_1(i) c'_y(\tau - i) = \sum_{i=0}^{2q} \epsilon h_2(i) s'_d(\tau - i)$$

$$= \sum_{i=0}^{2q} \epsilon h_2(i) \{ s'_y(\tau - i)$$

$$+ s'_w(\tau - i) \} \quad (51)$$

and

$$\sum_{i=0}^{q} h(i) c_{3y}(q, i - \tau) = \eta c_{2d}(\tau)$$

$$= \eta \{ c_{2y}(\tau) + c_{2w}(\tau) \} \quad (52)$$

where we have used (26)–(28). To obtain consistent parameter estimation, (51) should not be used for $0 \leq \tau \leq 2q$ if $w(n)$ is a *white* Gaussian noise, and for $-2\bar{q} \leq \tau \leq 2q + 2\bar{q}$ if $w(n)$ is a *colored* Gaussian noise with

$$E\{w(n)w(n + \tau)\} = \begin{cases} c_{2w}(\tau) \neq 0 & |\tau| \leq \bar{q} \\ 0 & \text{otherwise.} \end{cases} \quad (53)$$

Similarly, (52) should not be used for $\tau = 0$ if $w(n)$ is a white Gaussian noise, and for $|\tau| \leq \bar{q}$ if $w(n)$ is a colored Gaussian noise whose second-order cumulant is given by (53).

The recursive method developed in Section IV-A utilizes (32) for lags in the range $-2q \leq \tau \leq -q$. And the recursive method developed in Section IV-C utilizes (10) for lags in the range $|\tau| > \bar{q}$, and (32) for lags in the range $-2q \leq \tau \leq -q$ and $3q \leq \tau \leq 4q$. Therefore, they are not affected by presence of a Gaussian noise with a second-order cumulant as in (53). As a consequence, the LS solution of (45) will also remain unaffected if the rows of B that contain the samples of $c_{2w}(\tau)$ are removed. In contrary, the LS solution of (37) is affected by presence of a colored Gaussian noise. It can only handle a white Gaussian noise, and that is by removing $2q + 1$ rows of the matrix D which contain the sample $c_{2w}(0)$.

V. PROPERTIES OF THE NEW METHODS

In this section, we present the properties of the NMP system reconstruction methods described in Sections III and IV.

1) There is a close relationship between the LS solution of Sections III and IV. Specifically the matrix R can be decomposed into two matrices as shown below:

$$R = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \quad (54)$$

where

$$R_1 = (1 \quad h(1) \quad h(2) \quad \cdots \quad h(q)) - q + 1 \quad \text{row vector}$$

$$R_2 - (q + 1) \times (q + 1) \quad \text{matrix}$$

$$= \begin{pmatrix} \epsilon & \epsilon h(1) & \epsilon h(2) & \cdots & \epsilon h(q) \\ \epsilon h(1) & \epsilon h^2(1) & \epsilon h(1)h(2) & \cdots & \epsilon h(1)h(q) \\ \epsilon h(2) & \epsilon h(2)h(1) & \epsilon h^2(2) & \cdots & \epsilon h(2)h(q) \\ \vdots & \vdots & \vdots & & \vdots \\ \epsilon h(q) & \cdots & \cdots & \cdots & \epsilon h^2(q) \end{pmatrix} . \quad (55)$$

By projecting the elements of the matrix R_2 onto its diagonal, we can form the following vector:

$$v = (\epsilon \quad 2\epsilon h(1) \quad \epsilon(h^2(1) + 2h(2))$$
$$\cdots \quad 2\epsilon h(q-1)h(q) \quad \epsilon h^2(q)). \quad (56)$$

Examining (56), it is clear that

$$v(n) = \sum_{i=0}^{n} R_2(i, n-i)$$

$$= \epsilon(h(n) * h(n)) = \epsilon h_2(n) \quad (57)$$

where $v(0) = R_2(0, 0) = \epsilon$, and $v(2q) = R_2(q, q) = \epsilon h^2(q)$.

2) Instead of applying the SVD to the matrix R given by (13), we could have applied it to the matrix R_2 given by (55). R_2 has rank one and can always be written as

$$R_2 = h_\epsilon h_o^T$$

$$= \begin{pmatrix} \epsilon \\ \epsilon h(1) \\ \vdots \\ \epsilon h(q) \end{pmatrix} (1 \quad h(1) \quad h(2) \quad \cdots \quad h(q)). \quad (58)$$

In theory, $\{R(0, n)\}$, $\{h_\epsilon(n)/h_\epsilon(0)\}$, and $\{h_o(n)\}$ are identical. In practice, due to noise and estimation errors, they are different; and as such, we have three candidates to the unknown coefficients $\{h(n)\}$. They are

$$h^{(1)}(n) = R(0, n)$$

$$h^{(2)}(n) = h_o(n)$$

$$h^{(3)}(n) = h_\epsilon(n)/h_\epsilon(0) \quad 0 \le n \le q. \quad (59)$$

To choose one candidate out of the others, we select $h^{(i)}$ ($i = 1, 2, 3$) which minimizes the squared differences between the observed cumulants and the cumulants of the proposed model.

3) It is possible to remove the redundancy present in the unknown parameters of (11) and (37). Consider, for instance, the LS method of Section III. Multiplying both sides of (7) by $H(w_1 + w_2)$ and taking the inverse Fourier transform, we obtain

$$\sum_{i=0}^{q} h(i)g(\tau_1 - i, \tau_2 - i) = \epsilon h(\tau_1)h(\tau_2) \quad (60)$$

where $g(\tau_1, \tau_2) = \mathfrak{F}^{-1}[G(\omega_1, \omega_2)]$. The right-hand side of (60) is nonzero only in the region $0 \le \tau_1, \tau_2 \le q$. If $\tau_1, \tau_2 < 0$ or $\tau_1, \tau_2 > q$, (60) reduces to

$$\sum_{i=1}^{q} h(i)g(\tau_1 - i, \tau_2 - i) = -g(\tau_1, \tau_2) \quad (61)$$

where we have used $h(0) = 1$. The above equation can be written in a matrix form and then be solved for the unknown parameters $h(1)$, $h(2)$, $\cdots$, $h(q)$. Let us, however, note that the solution so obtained is affected by presence of additive Gaussian noise since it uses all sam-

ples of the second-order cumulants. In addition, $H(z)$ should not have zero(s) on the unit circle.

4) The RC and LS methods in Section IV can be extended in a straightforward manner to the fourth-order spectra. If $C_4(\omega, \omega, \omega)$ is the 1-D slice of the fourth-order spectrum, $G(\omega)$ will be related in this case to the system transfer function $H(\omega)$ by

$$G(\omega) = \frac{\epsilon H^3(\omega)}{H(3\omega)} \quad (62)$$

where $G(\omega) = C_4(\omega, \omega, \omega)/C_2(3\omega)$, and $C_2(3\omega)$ is the Fourier transform of the sequence

$$s'(\tau) = \begin{cases} c_2(\tau/3) & \text{if } \frac{\tau}{3} = \text{integer} \\ 0 & \text{otherwise.} \end{cases} \quad (63)$$

In the time domain, (62) takes the form

$$\sum_{i=0}^{3q} h_1(i)c'(\tau - i) = \sum_{i=0}^{3q} \epsilon h_2(i)s'(\tau - i) \quad (64)$$

where

$$h_1(n) = \mathfrak{F}^{-1}[H(3\omega)]$$

$$h_2(n) = \mathfrak{F}^{-1}[H^3(\omega)]$$

$$c'(\tau) = \mathfrak{F}^{-1}[C_4(\omega, \omega, \omega)].$$

Equation (64) resembles (32). Therefore, it can be solved in a procedure similar to that in Section IV using recursive and least squares methods. Extension of (10) to the fourth-order cumulants is straightforward and is given by [12]

$$\sum_{i=0}^{q} h(i)c_4(q, q, i - \tau) = \eta c_2(\tau). \quad (65)$$

VI. System Order Determination

In this section, we address the problem of system order determination using cumulant statistics. Knowledge of the system order is of utmost importance to many system identification algorithms. In [6], two methods were suggested for determining the order of an FIR system using the third-order cumulants. The methods are based on visual inspection and statistical tests. In the former, one searches for the order q for which $c_3(q, 0) \ne 0$ and $c_3(q + 1, 0) = 0$; and in the second, one tests the null hypothesis that the FIR order is q when $|c_3(q + 1, 0)| < t_c$ by computing

$$\Pr\{|c_3(q + 1, 0)| \le t_c\}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-t_c}^{t_c} \exp(-c^2/2\sigma^2) \, dc = 1 - \epsilon_0 \quad (66)$$

for a given probability of error ϵ_0, where σ^2 is the variance of the random variable $c_3(q + 1, 0)$ and t_c is a threshold. Obviously, the first method is impractical, while the second is dependent on statistics of a single random variable $c_3(q + 1, 0)$.

In the following, we propose a new procedure for determining the order of an FIR system using the second- and third-order cumulants of the output sequence. The method presented depends neither on visual inspection nor on statistical tests. In addition, it utilizes a much larger data set of output statistics than does (66).

It is well known that the third-order cumulants of an FIR system are identically zero for lags outside the region defined by $-q \leq \tau_1, \tau_2 \leq q$ (see Fig. 2 for an exact definition). From (10), the unknown impulse response $h(n)$ of an FIR system is related to the second-order cumulants $c_2(\tau)$ and the 1-D slice $c_3(q, \tau_2)$ of the third-order cumulants $c_3(\tau_1, \tau_2)$ by

$$\sum_{i=0}^{q} h(i) c_3(i - \tau, q) = \eta c_2(\tau). \tag{67}$$

Equation (67) can be rewritten in a matrix form as

$$d' = D' b' \tag{68}$$

where d', D', and b' are defined as in (43).

If $\beta_3 = 1$ and q is the true system order, the third-order cumulants can be computed as follows [9]:

$$c_3'(\tau_1, \tau_2, q) = \sum_{i=0}^{q} h(i) h(i + \tau_1) h(i + \tau_2) \tag{69}$$

where $c_3'(\tau_1, \tau_2; q)$ denotes the third-order cumulants computed by using the estimated impulse response coefficients $h(1) \cdots h(q)$ obtained from solving (68).

Let us now define the function $e(\tau_1, \tau_2; q')$ as follows:

$$e(\tau_1, \tau_2; q') = c_3(\tau_1, \tau_2) - c_3'(\tau_1, \tau_2; q'). \tag{70}$$

Theoretically, $e(\tau_1, \tau_2; q')$ should be zero when $q' = q$. The system order determination algorithm that we propose is based on the following steps:

1) From the given data, compute the second-order cumulants $c_2(\tau; p)$ in the region defined by $0 \leq \tau \leq p$ where $p > q$.

2) From the given data, compute the third-order cumulants $c_3(\tau_1, \tau_2; p)$ in the region bounded by the lines $\tau_1 = 0$, and $\tau_1 = \tau_1; 0 \leq \tau_1, \tau_2 \leq p$.

3) From (68), compute the unknown coefficients $h(1) \cdots h(q')$ assuming the order of the system is $q'; 1 \leq q' \leq p$.

4) Find the value of q' which minimizes the following performance measure:

$$E(q', p) = \sum_{\tau_1=0}^{p} \sum_{\tau_2=0}^{\tau_1} e^2(\tau_1, \tau_2; q')$$

$$= \sum_{\tau_1=0}^{q'} \sum_{\tau_2=0}^{\tau_1} (c_3(\tau_1, \tau_2; p) - c_3'(\tau_1, \tau_2; q'))^2$$

$$+ \sum_{\tau_1=q'+1}^{p} \sum_{\tau_2=0}^{\tau_1} c_3^2(\tau_1, \tau_2; p)$$

$$1 \leq q' \leq p. \tag{71}$$

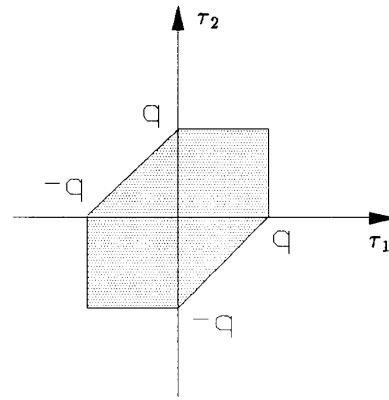Note that $c_3'(\tau_1, \tau_2; q') = 0$ for $\tau_1 > q'$. Theoretically,

Fig. 2. The region of support of third-order cumulant of an MA(q) system.

$E(q', p) = 0$ when $q' = q$. In practice, due to noise and estimation errors, $E(q, p) \neq 0$ but of minimum value. In deriving the above algorithm, we have used $\beta_3 = h(0) = 1$. If $h(0) \neq 1$ (or $\beta_2 \neq 1$), the sequences $c_3(\tau_1, \tau_2)$ and $c_3'(\tau_1, \tau_2)$ should be properly scaled to ensure that $e(\tau_1, \tau_2; q) = 0$ when $c_3(\tau_1, \tau_2)$ is the true cumulant sequence. One possible approach to doing this is to normalize $c_3(\tau_1, \tau_2)$ and $c_3'(\tau_1, \tau_2)$ by their zero samples at the origin. The samples $c_3(0, 0)$ and $c_3'(0, 0)$ should not in this case be equal to zero.

The above described algorithm is affected by presence of additive Gaussian noise since it utilizes the second-order cumulants. However, if the second-order cumulants of the additive noise are nonzero only for lags in the range $|\tau| \leq \bar{q}$, the effect of the noise can be eliminated by deleting the rows of the matrix D corresponding to $|\tau| \leq \bar{q}$. The matrix D so modified remains of full rank [12].

Instead of computing $c_3'(\tau_1, \tau_2; q)$ using the second- and third cumulants, we could have computed it using the third-order cumulants only. In [4], a relationship between the system impulse response $h(n)$ and the 1-D slice $c_3(q, \tau_2)$ has been derived. This relationship is given by

$$h(n) = \frac{c_3(q, n)}{c_3(-q, -q)} \qquad n = 0, 1, \cdots, q. \tag{72}$$

By substituting (72) into (69), we obtain

$$c_3'(\tau_1, \tau_2; q) = \frac{\sum_{i=0}^{q} c_3(q, i) c_3(q, i + \tau_1) c_3(q, i + \tau_2)}{c_3^3(-q, -q)}. \tag{73}$$

Equation (73) is useful when the output data is contaminated by additive, colored Gaussian noise with unknown statistics.

It is relevant at this point to mention that the system order determination algorithm described above can be applied to AR(q)-type processes, as well. If $\{ y(n) \}$ corresponds to an AR(q) model of order q, then its inverse statistics correspond to an MA(q) process. To see this, let

$B(\omega_1, \omega_2)$ denote the bispectrum of an AR (q) process described by

$$y(n) = x(n) - \sum_{i=1}^{q} h(i) y(n - i). \qquad (74)$$

Then

$$B(\omega_1, \omega_2) = \frac{\beta_3}{H(\omega_1, \omega_2)} \qquad (75)$$

where $H(\omega_1, \omega_2)$ is the bispectrum of $h(n)$. The inverse bispectrum of $y(n)$ is the bispectrum of $h(n)$ scaled by $\gamma = 1/\beta_3$. That is,

$$\frac{1}{B(\omega_1, \omega_2)} = \gamma H(\omega_1, \omega_2)$$

$$= \gamma H(\omega_1) H(\omega_2) H(-\omega_1 - \omega_2). \qquad (76)$$

By computing the finite extent third-order cumulants of $h(n)$ which is the inverse Fourier transform of $H(\omega_1, \omega_2)$, the algorithm of this section can be applied to determining the unknown order q. It is relevant at this point to notice that the third-order cumulants of $h(n)$ can also be computed from the third-order cumulants of $y(n)$ using a set of linear equations [11]:

$$\sum_{i,j=-q}^{q} c_{3(h)}(i, j) c_{3(y)}(\tau_1 - i, \tau_2 - j) = \beta_3 \delta(\tau_1, \tau_2) \qquad (77)$$

where (77) is derived by computing the inverse Fourier transform of both sides of the equation $B(\omega_1, \omega_2) H(\omega_1, \omega_2) = \beta_3$. Equation (77) can be written in a matrix form and then solved for the unknown cumulant samples of $h(n)$ for lags in the region $0 \leq \tau_1, \tau_2 \leq q$ and $\tau_1 \leq \tau_2$.

VII. SIMULATION EXAMPLES

A. System Identification

In this simulation example, the available data $\{d(n)\}$ is generated by the signal model [12]:

$$d(n) = y(n) + w(n)$$

$$= x(n) + 0.1x(n - 1) - 1.87x(n - 2)$$

$$+ 3.02x(n - 3)$$

$$- 1.435x(n - 4) + 0.49x(n - 5) + w(n) \qquad (78)$$

where the input signal $x(n)$ is zero-mean exponentially distributed i.i.d. noise generated by the GGEXN subroutine of the IMSL library with $\beta_3 = 1$. The zeros of the unknown system transfer function $H(z)$ are located at -2, $0.7 \pm j0.7$, and $0.25 \pm j0.433$. The signal $w(n)$ is an additive noise which is zero-mean white Gaussian process generated by the GGNML subroutine of the IMSL library.

Three different lengths of output data was used in this study: $N = 1024$, 2048, and 4096. The impulse response coefficients of the unknown system were computed for 100 output realizations at signal-to-noise ratio (SNR) $= \infty$ (noise-free case) where SNR $= E\{y^2(n)\}/E\{w^2(n)\}$. Besides the noise-free case, the impulse response coefficients were also computed for $N = 4096$ at SNR $= 100$ and 10. The second- and third-order cumulant sequences were computed using the indirect estimation method described in [10]. For each run, we also calculated the mean squared error (MSE) defined as

$$\text{MSE} = \frac{\sum_{n=0}^{q} (h(n) - h_r(n))^2}{\sum_{n=0}^{q} h^2(n)} \qquad (79)$$

where $h(n)$ and $h_r(n)$ are the true and reconstructed impulse responses, respectively. Tables I–III show the mean values and standard deviations of the various parameters when the LS solutions of Section III-A (LS-I), Section IV-C (LS-II), and Tugnait's method [13] (LS-T) are employed. For the noisy cases, rows of the coefficient matrix that contain the sample $c_{2d}(0)$ are removed. Based on the results obtained, we observe that: i) LS-II has lower MSE than LS-T, and ii) the estimates obtained via LS-I are generally better (in terms of variance and MSE) than the estimates obtained via LS-II and LS-T. This is, however, expected since the data set of output statistics exploited by LS-I is much larger than the data set of output statistics exploited by LS-II or LS-T; for example, LS-I utilizes all slices of third-order cumulant whereas LS-T utilizes only two slices of it.

B. Order Determination

In this section, we present an example demonstrating the use of the order determination algorithm described in Section VI. Seven independent realizations were generated using the signal model [6]:

$$d(n) = x(n) + 0.9x(n - 1) + 0.385x(n - 2)$$

$$- 0.771x(n - 3) + w(n) \qquad (80)$$

where $x(n)$ and $w(n)$ are defined as in (78). Table IV shows the value of $E(q', p)$ computed at $q' = 1, 2, \cdots, 5$; with $p = 5$, SNR $= 10$, and $N = 1024$. Based on the results obtained, it is clear that $E(q', p)$ attains its minimum when $q' = q = 3$.

In order to illustrate the effectiveness of our algorithm, 100 Monte Carlo runs for determining the order of the system described in (80) were performed at SNR $= 10$ and $N = 512$, 1024, 2048, and 4096. Table V shows the number of successful selections of the MA order when the algorithm of Section VI was implemented using (69) and (73). For comparison, the Giannakis–Mendel's algorithm [6] was also tested using the same data. The parameter ϵ_0 was fixed to 0.05 and the variance σ^2 was computed as

TABLE I
THE RECONSTRUCTED IMPULSE RESPONSE OF THE SIMULATION EXAMPLE USING LS-I AND 100 MONTE CARLO RUNS (μ = MEAN, σ = STANDARD DEVIATION, AND MSE = MEAN SQUARED ERROR)

| True Value | SNR = ∞ | | | | | | N = 4096 | | | |
| | N = 1024 | | N = 2048 | | N = 4096 | | SNR = 100 | | SNR = 10 | |
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
0.100	−0.0196	0.2407	−0.1294	0.2284	−0.1251	0.2259	−0.1025	0.1688	−0.0873	0.1973
−1.870	−0.7005	0.4349	−0.9627	0.4099	−1.2402	0.3934	−0.8926	0.4207	−0.8150	0.4255
3.020	1.3730	0.6648	1.8499	0.5575	2.2705	0.4670	1.2265	0.5742	1.0992	0.5530
−1.435	−0.7510	0.3224	−0.9748	0.2430	−1.1461	0.1641	0.4930	0.3096	−0.4281	0.2741
0.490	0.2182	0.1137	0.2950	0.1043	0.3455	0.0820	0.1137	0.0994	0.0943	0.0930
MSE	0.3431	0.2246	0.1940	0.1677	0.0984	0.1004	0.3690	0.2159	0.4151	0.2168

TABLE II
THE RECONSTRUCTED IMPULSE RESPONSE OF THE SIMULATION EXAMPLE USING LS-II AND 100 MONTE CARLO RUNS (μ = MEAN, σ = STANDARD DEVIATION, AND MSE = MEAN SQUARED ERROR)

| True Value | SNR = ∞ | | | | | | N = 4096 | | | |
| | N = 1024 | | N = 2048 | | N = 4096 | | SNR = 100 | | SNR = 10 | |
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
0.100	0.2839	0.7190	0.2453	0.6807	0.2578	0.6211	1.2010	0.7356	1.1615	0.7285
−1.870	−1.3817	0.7001	−1.5877	0.5620	−1.7870	0.4328	−0.5068	0.9958	−0.5222	0.9731
3.020	1.7987	0.6245	2.1656	0.5477	2.5424	0.3890	1.3117	0.0395	1.2876	0.9961
−1.435	−0.4836	0.5523	−0.7231	0.5478	−0.9542	0.4992	−0.5989	0.5070	−0.5564	0.5083
0.490	0.0031	0.4167	0.0999	0.3813	0.2215	0.3487	0.1908	0.2047	0.1780	0.2238
MSE	0.2991	0.2143	0.1883	0.1682	0.1034	0.1163	0.6066	0.4086	0.6005	0.3785

TABLE III
THE RECONSTRUCTED IMPULSE RESPONSE OF THE SIMULATION EXAMPLE USING LS-T AND 100 MONTE CARLO RUNS (μ = MEAN, σ = STANDARD DEVIATION, AND MSE = MEAN SQUARED ERROR)

| True Value | SNR = ∞ | | | | | | N = 4096 | | | |
| | N = 1024 | | N = 2048 | | N = 4096 | | SNR = 100 | | SNR = 10 | |
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
0.100	−0.2909	0.5064	−0.1873	0.5773	0.0027	0.5502	0.6468	0.6362	0.5328	0.5381
−1.870	−0.5784	0.6101	−0.8704	0.6552	−1.2176	0.6145	−0.3376	0.9576	−0.3988	1.0141
3.020	0.8711	0.6721	1.2669	0.7381	1.7313	0.6618	0.8240	1.0982	0.7267	1.3145
−1.435	−0.1518	0.5276	−0.2877	0.6739	−0.4393	0.5687	−0.4911	0.5839	−0.4311	0.5603
0.490	0.0342	0.2639	0.1012	0.3018	0.1437	0.2724	0.2223	0.2169	0.2127	0.1845
MSE	0.6095	0.3519	0.4681	0.3351	0.2956	0.2823	0.7107	0.4646	0.7569	0.4912

TABLE IV
SYSTEM ORDER DETERMINATION USING SEVEN INDEPENDENT REALIZATIONS WITH p = 5, SNR = 10, AND N = 1024

| q' | $E(q', p)$ | | | | | | |
	Run-1	Run-2	Run-3	Run-4	Run-5	Run-6	Run-7
1	1.4200	1.0365	1.5786	1.6723	1.6873	2.7133	1.9147
2	2.7602	1.6311	2.7367	3.9544	3.0917	4.3189	3.2453
3	0.3767	0.3671	0.1238	0.3711	0.1375	0.5422	0.3347
4	2.2084	1.5434	2.8321	1.0652	2.8356	2.3786	3.1369
5	2.3805	2.8165	1.5221	1.9823	0.3146	3.8384	5.2588

TABLE V
PERFORMANCE EVALUATION OF SYSTEM ORDER DETERMINATION ALGORITHMS WITH p = 5, SNR = 10, AND 100 MONTE CARLO RUNS

| Approach | Number of Successful Selections | | | |
	N = 512	N = 1024	N = 2048	N = 4096
Eq. (69)	83	85	96	100
Eq. (73)	84	91	93	99
[6]	11	30	52	74

follows:

$$\sigma^2[c_3(\tau_1, \tau_2)] = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=-q+\tau_2}^{q+\tau_1} \left(1 - \frac{|j|}{N}\right)$$

$$\cdot [d(i)d(i + \tau_1)d(i + \tau_2) - c_3(\tau_1, \tau_2)]$$

$$\cdot [d(i + j)d(i + j + \tau_1)d(i + j + \tau_2)$$

$$- c_3(\tau_1, \tau_2)]$$

$$\tau_1 = q + 1 \quad \text{and} \quad \tau_2 = 0 \quad (81)$$

where $c_3(\tau_1, \tau_2)$ is the observed third-order cumulant. The results of this simulation example are shown in Table V.

From Table V, we observe that the Giannakis–Mendel's algorithm deteriorates when N is small. For $N = 512$, the algorithm picked the correct order only 11 times, as opposed to the algorithm of Section VI (69), which was successful 83 times. For $N = 1024$, 2048, and 4096, we also observe that the new methods outperform the Giannakis–Mendel's approach. Performance of the two new methods described in Section VI are comparable to each other.

VIII. Conclusion

In this paper, new methods for the identification of non-Gaussian white-noise-driven NMP-LTI FIR systems are proposed. Recent developments on NMP-LTI FIR system identification include works by Giannakis and Mendel [5], and Tugnait [12], [13] who considered FIR parameter estimation using the second- and one-dimensional (1-D) slices of output cumulants. In this paper, we show how to solve this problem using the second-order and all samples of third-order cumulants in an appropriate domain of support. Identification methods using the second-order cumulants and the diagonal slice of bispectrum are also developed. It is shown that the methods presented yield consistent parameter estimation in a class of colored

Gaussian noise. Both recursive closed form and batch least squares versions of the parameter estimators are presented. Extension of these methods to the fourth-order cumulants is also addressed.

A relevant problem in parametric modeling of higher order statistics is the determination of system order. In this paper, two new methods for determining the order of an FIR system using only output cumulants are also presented. The first method utilizes the second- and third-order cumulants, whereas the second method utilizes the third-order cumulants only. It is shown by computer simulation that our system order determination algorithms outperform the existing cumulant based algorithms [6].

Appendix

The least squares method presented in Section III-A can be extended to the fourth-order cumulant. The function $G(\omega_1, \omega_2, \omega_3)$ can be expressed in terms of the system transfer function as follows:

$$G(\omega_1, \omega_2, \omega_3) = \frac{C_4(\omega_1, \omega_2, \omega_3)}{C_2(\omega_1 + \omega_2 + \omega_3)}$$

$$= \frac{\epsilon H(\omega_1)H(\omega_2)H(\omega_3)}{H(\omega_1 + \omega_2 + \omega_3)} \quad (82)$$

where $\epsilon = \beta_4/\beta_2$ and $C_4(\omega_1, \omega_2, \omega_3)$ is the trispectrum of the system output. Multiplying (82) by $C_2(\omega_1 + \omega_2 + \omega_3)H(\omega_1 + \omega_2 + \omega_3)$ and taking the inverse Fourier transform, we find

$$\sum_{i=0}^{q} h(i)c_4(\tau_1 - i, \tau_2 - i, \tau_3 - i)$$

$$= \sum_{i=0}^{q} \epsilon h(i)h(\tau_2 - \tau_1 + i)h(\tau_3 - \tau_1)c_2(\tau_1 - i).$$

$$(83)$$

From (83), we can form an overdetermined system of linear equations whose unknown vector r is given by

$$r = (h_0 \quad \epsilon r_0 \quad h(1)r_1 \quad \cdots \quad h(q)r_q)^T \quad (84)$$

where

$$h_0 = (h(1) \quad h(2) \quad \cdots \quad h(q)),$$

$$r_i = (\epsilon h^2(i) \quad \epsilon h(i)h(i + 1) \quad \cdots \quad \epsilon h(i)h(q)$$

$$\epsilon h^2(i + 1) \quad \cdots \quad \epsilon h(i + 1)h(q) \quad \cdots \quad \epsilon h^2(q)),$$

$$h(0) = 1 \quad \text{and} \quad i = 0, 1, 2, \cdots, q.$$

Once the vector r is determined, the matrix R is formed as follows:

$$R = \begin{pmatrix} \epsilon & \epsilon h(1) & \cdots & \epsilon h(q) & \epsilon h^2(1) & \cdots & \epsilon h(1)h(q) & \cdots & \epsilon h^2(q) \\ \epsilon h(1) & \epsilon h(1)h(1) & \cdots & \epsilon h(1)h(q) & \epsilon h(1)h^2(1) & \cdots & \epsilon h^2(1)h(q) & \cdots & \epsilon h(1)h^2(q) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \epsilon h(q) & \epsilon h(q)h(1) & \cdots & \epsilon h(q)h(q) & \epsilon h(q)h^2(1) & \cdots & \epsilon h(1)h^2(q) & \cdots & \epsilon h(q)h^2(q) \end{pmatrix}. \quad (85)$$

The matrix R has rank one, and can be written as

$$R = h_1 \hat{r} = \begin{pmatrix} \epsilon \\ \epsilon h(1) \\ \vdots \\ \epsilon h(q) \end{pmatrix} (1 \quad h(1) \quad \cdots \quad h(q) \quad h^2(1)$$

$$\cdots \quad h(1)h(q) \quad \cdots \quad \cdots \quad h^2(q)). \quad (86)$$

The vectors h_1 and $\hat{r}$ can be determined from R using the SVD.

A matrix $\hat{R}$ can also be formed from the elements of the

vector $\hat{r}$ as follows:

$$\hat{R} = \begin{pmatrix} 1 & h(1) & h(2) & \cdots & h(q) \\ h(1) & h^2(1) & h(1)h(2) & \cdots & h(1)h(q) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h(q) & h(q)h(1) & h(q)h(2) & \cdots & h^2(q) \end{pmatrix}.$$

(87)

By applying the SVD to the matrix $\hat{R}$, we find

$$\hat{R} = h_2 h_3 = \begin{pmatrix} 1 \\ h(1) \\ \vdots \\ h(q) \end{pmatrix} (1 \quad h(1) \quad \cdots \quad h(q)). \quad (88)$$

Therefore, by using the fourth-order cumulant, four candidates to the unknown system impulse response $\{h(n)\}$ are obtained. They are: $\{h_0(n)\}$, $\{h_1(n)/h_1(0)\}$, $\{h_2(n)\}$, and $\{h_3(n)\}$. With the exact knowledge of second- and fourth-order cumulant samples, all the candidates are identical. The uniqueness of the (least squares) solution is guaranteed, since the coefficient matrix of (83) is of full rank. This follows from the fact that the unknown parameters $\{h(n)\}$ and $\{\epsilon h(i)h(\tau_2 - \tau_1 + i)h(\tau_2 - \tau_1 + i)\}$ can be uniquely determined from (83) using procedures analogous to those of the third-order cumulant described in Section III-C.

REFERENCES

[1] S. A. Alshebeili and A. E. Cetin, "A phase reconstruction algorithm from bispectrum," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, no. 2, pp. 166–170, 1990.

[2] A. E. Cetin, "Algorithm for signal reconstruction from bispectrum," in *Proc. 14th Biennial Symp. Commun. Signal Processing*, Kingston, Ontario, Canada, May 1988, pp. D.2.5–8.

[3] B. Friedlander and B. Porat, "Asymptotically optimal estimation of MA and ARMA parameters of non-Gaussian processes from higher order moments," *IEEE Trans. Automat. Contr.*, vol. 35, no. 1, pp. 27–35, 1990.

[4] G. B. Giannakis, "Cumulants: A powerful tool in signal processing," *Proc. IEEE*, vol. 75, no. 9, pp. 1333–1334, 1987.

[5] G. B. Giannakis and J. M. Mendel, "Identification of nonminimum phase systems using higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 3, pp. 360–377, 1989.

[6] G. B. Giannakis and J. M. Mendel, "Cumulant-based order determination of non-Gaussian ARMA models," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 8, pp. 1411–1423, 1990.

[7] G. B. Giannakis and A. Swami, "On estimating noncausal nonminimum phase ARMA models of non-Gaussian processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 3, pp. 478–495, 1990.

[8] G. H. Golub and C. F. Van Loan, *Matrix Computation*. Baltimore, MD: Johns Hopkins University Press, 1983.

[9] 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, 1991.

[10] C. L. Nikias and M. R. Raghuveer, "Bispectrum estimation: A digital signal processing framework," *Proc. IEEE*, vol. 75, no. 7, pp. 869–891, 1987.

[11] M. Rangoussi and G. B. Giannakis, "On the use of second- and higher order inverse statistics," in *Proc. Workshop Higher Order Spectral Analysis*, Vail, CO, pp. 7–12, June 1989.

[12] 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, 1990.

[13] J. K. Tugnait, "New results on FIR system identification using higher order statistics," in *Proc. Fifth ASSP Workshop Spectrum Estimation Modeling*, Rochester, NY, Oct. 1990, pp. 212–216; also in *IEEE Trans. Signal Processing*, vol. 39, no. 10, pp. 2216–2221, 1991.

[14] J. K. Tugnait, "Identification of linear stochastic systems via second- and fourth-order cumulant matching," *IEEE Trans. Inform. Theory*, vol. 33, no. 3, pp. 393–407, 1987.

Saleh A. Alshebeili (S'89–M'91) received the Ph.D. degree from the University of Toronto in 1991. He is currently with King Saud University, Saudi Arabia.

Anastasios N. Venetsanopoulos (S'66–M'69–SM'79–F'88) received the B.S. degree from the National Technical University of Athens (NTU), Greece, in 1965, and the M.S., M.Phil., and Ph.D. degrees in electrical engineering, all from Yale University, in 1966, 1968, and 1969, respectively.

He joined the University of Toronto, Canada, in September 1968, where he is now Professor in the Department of Electrical Engineering. He also served as Chairman of the Communications Group (1974 to 1978 and 1981 to 1986), and as Associate Chairman of the Department of Electrical Engineering (1978 to 1979). He was on research leave at the Swiss Federal Institute of Technology, the University of Florence, the Federal University of Rio de Janeiro, the National Technical University of Athens, and the Imperial College of Science and Technology, and was Adjunct Professor at Concordia University. He served as Lecturer of numerous short courses to industry and continuing education programs; he is a contributor to eleven books and has published over 300 papers in digital signal and image processing, and digital communications; he also served as consultant to several organizations, and as Editor of the *Canadian Electrical Engineering Journal* (1981 to 1983).

Dr. Venetsanopoulos is a member of the New York Academy of Sciences, Sigma Xi, the International Society for Optical Engineering, and the Technical Chamber of Greece; he is a Registered Professional Engineer in Ontario and Greece, a Fellow of the Engineering Institute of Canada, and a Fellow of the Institute of Electrical and Electronics Engineers for contributions to digital signal and image processing.
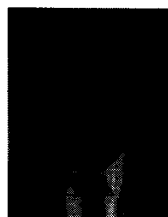
A. Enis Çetin (S'85–'87) received the B.Sc. degree in electrical engineering from the Middle East Technical University, Ankara, Turkey, in 1984, and the M.S.E. and Ph.D. degrees in systems engineering from the Moore School of Electrical Engineering, University of Pennsylvania, in 1986 and 1987, respectively.

From September 1987 to July 1989 he was Assistant Professor of Electrical Engineering at the University of Toronto. During the summers of 1988 and 1991 he was a consultant and a resident visitor, respectively, at Bellcore, Morristown, NJ. He is presently Associate Professor of Electrical and Electronics Engineering at Bilkent University, Ankara, Turkey.