Identification of MA processes using cumulants: several sets of linear equations

D.P.Ruiz
M.C.Carrión
A.Gallego
J.A.Morente

Indexing terms: Blind identification, High order spectral analysis, Nonminimum phase, System modelling, Moving average system

Abstract: Several new sets of linear equations relating the coefficients of a moving average (MA) system with its \((k-1)\text{th}\) and \(k\text{th}\)-order polyspectra, with \(k>2\), are presented. These equations are used to identify the nonminimum phase MA model parameters from the statistics of the noisy output. The system is driven by an (unobservable) independent and identically distributed nongaussian process and the noise is additive and coloured gaussian with unknown power spectrum. Four estimators based on the method of least-squares using linear relations between third- and fourth-order statistics are proposed and discussed. These methods make use of more higher order statistics information than other linear methods do, and the uniqueness of the solutions can normally be guaranteed by the linearity of the equations. Results from computer simulations confirm the expected theoretical advantages of the proposed methods in coloured noise environments and allow one to draw a comparison among proposed and other published linear methods.

1 Introduction

In the last decade increasing attention has been paid to identification of nonminimum phase systems using higher order statistics (HOS). There are many potential advantages to using HOS, two of them being [1–3]:

(i) The phase information of the signals is preserved, and therefore cumulants can be used to determine both magnitude and phase characteristics of nonminimum phase systems (NMP systems), unlike the autocorrelation sequence which is unable to detect NMP systems.

(ii) If the signal is corrupted by additive gaussian noise, uncorrelated with the input, the higher-order cumulants of the noisy signal are equal to the noiseless output, which allows one to suppress any gaussian noise regardless of its power spectrum.

Methods dealing with NMP systems can be applied to geophysics [4], data communications [5] and many other fields, such as sonar, radar, blind equalisation, time-delay estimation, image processing etc. (for an overview see [1, 3]). Recently several methods using cumulant statistics have been proposed for the identification of NMP systems. These methods may be classified into three categories [1]:

(a) Closed-form solutions, which are interesting from a theoretical point of view but are impractical because they do not smooth out the effects of measurement noise.

(b) Optimisation solutions, which normally provide the best estimations, but are computationally involved and the convergence to a global minimum is never guaranteed.

(c) Linear algebra solutions, based on the solution of a system of equations which are viewed to be linear in a set of parameters. These solutions have generated a great deal of interest owing to their computational simplicity and as good initial guesses to optimisation-based solutions.

This paper presents several sets of linear equations involving cumulants of arbitrary order and model parameters to identify a (possibly) NMP MA linear system driven by a nonmeasurable IID nongaussian sequence only from output measurements. These equations provide a general framework to combine \((k-1)\text{th}\) and \(k\text{th}\)-order cumulants \((k>2)\), which generalises existing linear methods such as the reformulated GM algorithm [8, 9] or those proposed in [10, 11] and allows one to obtain new sets of equations useful in recovering MA model coefficients. Since the use of the signal autocorrelation sequence is optimal in these equations, the proposed relationships will prove useful in estimating model coefficients from signal observations corrupted by additive coloured gaussian noise of unknown autocorrelation function.

2 Problem statement and preliminaries

Consider a \(2q\)th stationary random process \(x(i)\) which is the output of a stable linear time-invariant (LTI) (possibly NMP) MA system with parameters \(b(i), i = 0...q\). Consequently, the time series is described by

\[
x(n) = \sum_{i=0}^{q} b(i)w(n-i)
\]

where \(q\) denotes the MA system order and \(w(n)\) stands
for the input random process. The system output is corrupted with additive noise v(n). Therefore the observed signal y(n) is specified by

\[ y(n) = x(n) + v(n) \]  

(2)

The problem here is to determine the parameters of the system \( b(i), i = 0 \ldots q \) using only the observed noisy signal \( y(n) \) with a higher-order statistics-based approach.

The following conditions are assumed to hold:

**A.S1:** The driving noise sequence \( w(n) \) is unobservable, and is a zero-mean, independent and identically distributed (IID) non-Gaussian process with at least \( 0 < \gamma_k < \infty \) and \( 0 < \gamma_{k+1} < \infty \), \( k > 3 \), where \( \gamma_k \) denotes the \( k \)th order cumulant of the random variable \( w(n) \).

**A.S2:** The additive noise \( v(n) \) is an IID zero-mean gaussian process (white or coloured), independent of the input \( w(n) \) and hence of the output \( x(n) \).

**A.S3:** The LTI MA system \( \{ b(i), i = 1 \ldots q \} \) is causal, possibly NMP, with known order \( q \) and \( b(0) = 1 \). The last condition fixes the inherent scale ambiguity.

These assumptions, and among them, that the time series is non-gaussian, are often found in practical situations [4-6]. Since for a gaussian process \( \gamma_k = 0 \) for \( k > 2 \), output cumulants would vanish for \( k > 2 \), and therefore, one must assume a non-gaussian input signal.

It should also be noted that the IID assumption for \( w(n) \) is stronger than necessary, and that it suffices to have inputs with flat spectra over the non-zero frequency range of the system function.

For the output series defined in this section, the \( k \)th order output cumulant may be denoted as \( c_{xy}(\tau_1, \tau_2, \ldots, \tau_{k-1}) \) owing to the stationarity of the random process. Also, due to the fact that the \( k \)th-order cumulants of gaussian processes vanish for \( k > 2 \), we infer that (assumption 2):

\[ c_{xy}(\tau_1, \tau_2, \ldots, \tau_{k-1}) = c_{xx}(\tau_1, \tau_2, \ldots, \tau_{k-1}) \quad k \geq 3 \]  

(3)

Based on this observation, we use \( c_{xy} \) and \( c_{xx} \) interchangeably in the sequel. This insensitivity of cumulants to the output gaussian noise of unknown covariance function is the main point for using higher-order cumulants, even for the minimum-phase case.

On the other hand, because \( w(n) \) is IID, its \( k \)th order cumulants \( c_{ww}(\tau_1, \tau_2, \ldots, \tau_{k-1}) \) are \( \gamma_k \delta(\tau_1, \tau_2, \ldots, \tau_{k-1}) \). Taking this fact into account, and under assumptions A.S1–A.S3, the \( k \)th-order output cumulant sequence may be expressed in terms of the MA system coefficients by the Brillinger–Rosenblatt formula [15] as

\[ c_{xy}(\tau_1, \tau_2, \ldots, \tau_{k-1}) = \gamma_k \sum_{i=0}^{q} b(i) c_{xy}(i + \tau_1, \ldots, i + \tau_{k-1}) \]  

(4)

Taking the \( k \)-1 dimensional Z-transform, this relation turns out to be

\[ c_{xy}(z_1, z_2, \ldots, z_{k-1}) = \gamma_k \sum_{i=0}^{q} b(i) c_{xy}(z_1, z_2, \ldots, z_{k-1}) \]  

(5)

\[ \cdots \cdot B(\tau_{k-1}) \cdot B(\tau_{k-1})^{-1} \]

where \( B(z) \) is the model's transfer function and \( C_{xy}(z_1, z_2, \ldots, z_{k-1}) \) is the \( k \)th-order polynomials.

## 3 General relationship between the kth- and (k–1)th-order polynomials

### 3.1 First set of equations

Consider the output signal \( y(n) \) defined in the previous Section, with \( \gamma_{k-1} \) and \( \gamma_{k-2} \) not equal to zero \( (k > 3) \), and the following cumulant sequences:

\[ s_1(\tau_1, \tau_2, \ldots, \tau_{k-1}) = c_{xy}(\tau_1, \tau_2, \ldots, \tau_{k-1}) \]  

(6)

\[ s_2(\tau_1, \tau_2, \ldots, \tau_{k-1}) = c_{xy}(\tau_1, \tau_2, \ldots, \tau_{k-1}) \cdot B(\tau_{k-1})^{-1} \]  

(7)

\[ s_3(\tau_1, \tau_2, \ldots, \tau_{k-1}) = c_{xy}(\tau_1, \tau_2, \ldots, \tau_{k-1}) \cdot B(\tau_{k-2})^{-1} \]  

(8)

These equations may be considered as the generalisation of the reformulated GM equation [7] and the linear equation proposed by Tugnait [8], eqn. 10 to the \( k \)th and \((k-1)\)th-order cumulants domain. Setting \( k = 3 \) in eqns. 11 and 12, and assuming that no noise corrupts the system output, one can rewrite the above equations as

\[ \sum_{i=0}^{q} b(i) c_{xy}(i + \tau, \ldots, i + \tau_{k-1}) = \gamma_k \sum_{i=0}^{q} b(i) c_{xy}(i + \tau, \ldots, i + \tau_{k-1}) \]  

(9)

where \( \gamma_k \) is the standard notation \( \gamma_k(\tau) = c_k(\tau) \) has been used for the autocorrelation sequence. These sets of equations were combined in [9] to identify MA system parameters, and this method is referred to as the T-method in the literature. The drawback to these equations is that they are sensitive to additive gaussian noise of unknown autocorrelation sequence, since they employ the autocorrelation sequence of the system output. When coloured gaussian noise contaminates the output, a useful method may be obtained by setting \( k = 4 \) in eqns. 11 and 12, as stated in the following proposition:

---

Proposition 1: Consider the output noisy signal \( y(n) \) defined in eqn. 2. Under assumptions AS1–AS3 with \( k \) sets up to four the following relations between the third- and fourth-order output cumulants sequences hold:

\[
\sum_{i=0}^{q} b(i)c_{4y}(\tau_1 - i, \tau_2 - i, i, l) = \varepsilon \sum_{i=0}^{q} b^2(i)c_{3y}(\tau_1 - i, l) \quad (15)
\]

\[
\sum_{i=0}^{q} b(i)c_{4y}(\tau_1 - i, \tau_2 - 0, q) = \varepsilon b(q)c_{3y}(\tau_1, -\tau_2) \quad (16)
\]

where the symmetry properties of the cumulants have been used, \( \varepsilon = \gamma_k \gamma_l \), and \( l \) ranges from \(-q\) to \( q \), standing for the specific slice used in eqn. 15.

Remark 1: The set of equations in (eqns. 15 and 16) are insensitive to additive Gaussian noise of unknown autocoherence sequence contaminating the system output, unlike eqns. 13 and 14, which employ second-order information. This point suggests that better parameter estimates may be obtained with eqns. 15 and 16 and when the additive noise is coloured Gaussian. On the other hand, when estimated sampled cumulants are used instead of the true statistics, as occurs in any practical case, higher-order sampled cumulants have larger variances, so the use of eqns. 13 and 14 is recommended for high-fidelity signals. When the Gaussian noise is white, the zero autocorrelation lag can be removed, as stated in the T-method, but when the Gaussian noise is coloured, the estimates obtained by this method are strongly affected by noise, and eqns. 15 and 16 prove to be useful, as shown in Section 5.

3.2 Second set of equations

In addition, a different set of equations using \((k-1)\)th- and \(k\)th-order cumulants sequences may be obtained. Consider the sequences

\[
s_1(\tau_1, \tau_2, \ldots, \tau_k) = c_{3y}(\tau_1, \tau_2, \ldots, \tau_k) \delta(\tau_k) + \sum_{i=0}^{q} b^2(i)c_{3y}(\tau_1 - i, l, 0)
\]

\[
s_1(\tau_1, \tau_2, \ldots, \tau_k) = c_{k-1y}(\tau_1, \tau_2, \ldots, \tau_k - 1) \delta(\tau_k - 1) + \sum_{i=0}^{q} b^2(i)c_{k-1y}(\tau_1 - i, \tau_2 + i, \ldots, \tau_k - 1 + i, 0)
\]

In the \(k\)-dimensional Z-domain, the \(k\)th-order poly-spectrum of these sequences may be written as

\[
S_k(z_1, z_2, \ldots, z_k) = C_{k-1y}(z_1, z_2, \ldots, z_k) + \sum_{i=0}^{q} b^2(i)C_{3y}(z_1 - i, z_2, \ldots, z_k)
\]

\[
S_k(z_1, z_2, \ldots, z_k) = C_{k-1y}(z_1, z_2, \ldots, z_k - 2) + \sum_{i=0}^{q} b^2(i)C_{k-1y}(z_1 - i, z_2 + i, \ldots, z_k - 1 + i)
\]

Combining the pair of equations obtains

\[
\gamma_{k-1}B(\tau_2, z_2 - 2)B(z_1, \tau_1 - 2, z_k - 1) = \gamma_{k-1}B(z_1, \tau_1 - 2, z_k - 1)B(z_1, \tau_1 - 2, z_k - 1)
\]

Taking the inverse \(k\)-dimensional Z-transform on both sides of eqn. 19, the following time-domain equation relating the \((k-1)\)th- and \(k\)th-order cumulants sequences appears:

\[
\gamma_{k-1}B(\tau_k - 1) = \gamma_{k-1}b(\tau_k - 1) + \sum_{i=0}^{q} b(i)b(\tau_k - 2 - i, \tau_1 + i, \ldots, \tau_k - 2 + i, \tau_1 + i)
\]

\[
\gamma_{k-1}b(\tau_k - 1) = \gamma_{k-1}b(\tau_k - 1) + \sum_{i=0}^{q} b(i)b(\tau_k - 2 - i, \tau_1 + i, \ldots, \tau_k - 2 + i, \tau_1 + i)
\]

Result 2: The coefficients of the MA system can be obtained from the family of equations resulting from eqn. 20 by imposing \( \tau_k = \tau_{k-2} = \tau_{k-3} = 0 \) or \( \tau_k = \tau_{k-2} = q \), \( \tau_{k-1} = 0 \), respectively.

\[
\gamma_{k-1} = \gamma_{k-2} = \gamma_{k-3} = 0
\]

\[
\gamma_{k-1}b(q)c_{k-1y}(\tau_1 + i, \tau_2 + i, \ldots, \tau_k - 3 + i, 0)
\]

\[
\gamma_{k-1}b(q)c_{k-1y}(\tau_1 + i, \tau_2 + i, \ldots, \tau_k - 3 + i, 0)
\]

In the case of third- and fourth-order cumulants, this result may be summarised in the following proposition:

Proposition 2: Output cumulants of MA processes under assumptions AS1–AS3, with \( k \) sets up to four, obey the recursive equations

\[
\sum_{i=0}^{q} b(i)c_{3y}(\tau + i, 0) = \sum_{i=0}^{q} b^2(i)c_{3y}(\tau + i, 0)
\]

\[
\gamma_{k-1}b(q)c_{k-1y}(\tau_1 + i, \tau_2 + i, \ldots, \tau_k - 3 + i, 0)
\]

\[
\gamma_{k-1}b(q)c_{k-1y}(\tau_1 + i, \tau_2 + i, \ldots, \tau_k - 3 + i, 0)
\]

Remark 2: The pair of equations given by proposition 2 constitutes an alternative to the equations given by proposition 1. But unlike eqns. 15 and 16, they only make use of the horizontal slice of fourth-order cumulants \( c_{4y}(\tau, 0, 0) \) and the horizontal slices \( c_{3y}(\tau, 0) \) and \( c_{3y}(\tau, q) \) of third-order cumulants, whereas eqns. 15 and 16 allow a greater range of slices for both third- and fourth-order cumulants. Another fact that deserves consideration is that when the system of eqns. 23 and 24 is treated as linear in both \( b(i) \) and \( b^2(i) \), the parameter estimates are obtained by taking the cubic root on the system solution \( b(i) \), which could make this procedure more sensitive to errors in the estimated sampled cumulant sequences.

3.3 Third set of equations based on last horizontal slice (q-slice)

Several families of equations may be obtained from the last horizontal slice of \((k-1)\)th- and \(k\)th-order cumulants. These relationships arise from the Brillinger–Rosenblatt formula, which relates the output cumulants of a MA filter driven by non-Gaussian IID noise with its coefficients eqn. 4. Using the fact that \( b(i) = 0 \) for \( i > q \), the last horizontal (multidimensional) slice for \((k-1)\)th- and \(k\)th-order cumulants is given by

\[
C_{k-1y}(\tau_1, \ldots, \tau_{k-2}, q) = \gamma_{k-1}b(0)b(\tau_1) \cdots b(\tau_{k-2})b(q)
\]

\[
C_{k-1y}(\tau_1, \ldots, \tau_{k-3}, q) = \gamma_{k-1}b(0)b(\tau_1) \cdots b(\tau_{k-3})b(q)
\]

obtained by making \( \tau_{k-2} = q \) and \( \tau_{k-1} = q \), respectively, in the Brillinger–Rosenblatt formula.

We are interested in looking for a relation between the last horizontal slices of \((k-1)\)th- and \(k\)th-order cumulants (or equivalently \((k+1)\)th and \(k\)th-order...
cumulants). To accomplish this, take powers of \( c_{k}^{p}(\tau_{1}, \tau_{2}, \ldots, q) \) and look for an equation which relates it with the last slice of \((k-1)\)-th order cumulants \( c_{k-1}^{q}(\tau_{1}, \tau_{2}, \ldots, q) \). Many options are possible, but the following result summarises the most useful cases for the foregoing applications:

**Result 3:** \( a^{k}_{x} \left( \tau_{1}, \tau_{2}, \ldots, q \right) \) and \( c_{k}^{q}(\tau_{1}, \tau_{2}, \ldots, q) \). The following equations for the last horizontal slice of the \((k-1)\)-th order cumulant sequence hold:

\[
\begin{align*}
\sum_{i=0}^{q} b(i) c_{k-1}^{q-1}(\tau_{i} + i, \tau_{2}, \ldots, q) &= \gamma_{k} \sum_{i=0}^{q} b(i) c_{k-1}^{q-2}(\tau_{i} + i, \tau_{2}, \ldots, q) \\
&= \gamma_{k} \frac{\beta^{k-2}(\tau_{1}) \cdots b^{k-2}(\tau_{q}) c_{k}^{q}(\tau_{1}, \tau_{1}, \ldots, \tau_{1})}{\gamma_{k}} \\
&= \gamma_{k} \frac{\beta^{k-2}(\tau_{1}) \cdots b^{k-2}(\tau_{q}) c_{k}^{q}(\tau_{1}, \tau_{1}, \ldots, \tau_{1})}{\gamma_{k}} \quad (27)
\end{align*}
\]

**Proof:** Taking the \((k-1)\)th power in eqn. 26 obtains

\[
\frac{\beta^{k-1}(\tau_{1} + i, \tau_{2}, \ldots, q)}{\gamma_{k}^{k-1}} = \gamma_{k}^{k-1} \frac{\beta^{k-1}(\tau_{1} + i, \tau_{2}, \ldots, q)}{\gamma_{k}^{k-1}} \quad (28)
\]

Multiplying both sides of eqn. 28 by \( b(i) \) and summing over the nonzero range of \( b(i) (i = 0 \ldots q) \), gives

\[
\begin{align*}
\sum_{i=0}^{q} b(i) c_{k-1}^{q-1}(\tau_{i} + i, \tau_{2}, \ldots, q) &= \gamma_{k} \sum_{i=0}^{q} b(i) c_{k-1}^{q-2}(\tau_{i} + i, \tau_{2}, \ldots, q) \\
&= \gamma_{k} \frac{\beta^{k-2}(\tau_{1}) \cdots b^{k-2}(\tau_{q}) c_{k}^{q}(\tau_{1}, \tau_{1}, \ldots, \tau_{1})}{\gamma_{k}} \quad (29)
\end{align*}
\]

Keeping in mind eqn. 4, the summation in the second member in eqn. 29 can be recognized as the sequence \( c_{p,1}^{q}(\tau_{1}, \tau_{1}, \ldots, \tau_{1}) \gamma_{k}^{q} \), which allows one to obtain the first equation in eqn. 27.

Similarly, one can prove the second equation in eqns. 27 by taking the \((k-2)\)th power in eqn. 26 and substituting \( \tau_{1} \) with \( \tau_{1} + i \):

\[
\begin{align*}
\frac{\beta^{k-2}(\tau_{1} + i, \tau_{2}, \ldots, q)}{\gamma_{k}^{k-2}} &= \gamma_{k} \frac{\beta^{k-2}(\tau_{1} + i, \tau_{2}, \ldots, q)}{\gamma_{k}^{k-2}} \\
&= \gamma_{k} \frac{\beta^{k-2}(\tau_{1} + i, \tau_{2}, \ldots, q)}{\gamma_{k}^{k-2}} \quad (30)
\end{align*}
\]

Multiplying both sides by \( b(i) \) and taking the summation from \( i = 0 \) to \( q \) gives

\[
\begin{align*}
\sum_{i=0}^{q} b(i) c_{k-1}^{q-2}(\tau_{i} + i, \tau_{2}, \ldots, q) &= \gamma_{k} \sum_{i=0}^{q} b(i) c_{k-1}^{q-1}(\tau_{i} + i, \tau_{2}, \ldots, q) \\
&= \gamma_{k} \frac{\beta^{k-2}(\tau_{1}) \cdots b^{k-2}(\tau_{q}) c_{k}^{q}(\tau_{1}, \tau_{1}, \ldots, \tau_{1})}{\gamma_{k}} \quad (31)
\end{align*}
\]

Taking into account eqn. 4, the expression can be written as

\[
\begin{align*}
\sum_{i=0}^{q} b(i) c_{k-1}^{q-2}(\tau_{i} + i, \tau_{2}, \ldots, q) &= \gamma_{k} \frac{\beta^{k-2}(\tau_{1}) \cdots b^{k-2}(\tau_{q}) c_{k}^{q}(\tau_{1}, \tau_{1}, \ldots, \tau_{1})}{\gamma_{k}} \\
&= \gamma_{k} \frac{\beta^{k-2}(\tau_{1}) \cdots b^{k-2}(\tau_{q}) c_{k}^{q}(\tau_{1}, \tau_{1}, \ldots, \tau_{1})}{\gamma_{k}} \quad (32)
\end{align*}
\]

which completes the proof.

The relationship between third- and fourth-order cumulants may be obtained by setting \( k = 4 \) in result 3.

The equations that arise in this case are collected in proposition 3.

**Proposition 3:** Under assumptions AS1–AS3, the following relationships involving the \((k-1)\)-th order cumulants \( c_{k-1}^{q}(\tau_{1}, \tau_{2}, \ldots, q) \) of an \( MA(q) \) system hold:

\[
\begin{align*}
\sum_{i=0}^{q} b(i) c_{k}^{q}(\tau + i, q) &= \gamma_{k} \frac{\beta^{q}(\tau, \tau, \tau)}{\gamma_{k}} \\
&= \gamma_{k} \frac{\beta^{q}(\tau, \tau, \tau)}{\gamma_{k}} \\
&= \gamma_{k} \frac{\beta^{q}(\tau, \tau, 0)}{\gamma_{k}} \quad (34)
\end{align*}
\]

**Proof:** These equations are the third- and fourth-order counterparts of those appearing in result 3.

**Remark 3:** The foregoing result establishes a family of equations relating the \( q \)-slice with other slices of the \((k-1)\)-th or \( k \)-th order cumulant sequence for an \( MA(q) \) system which, in addition to those obtained in the previous Sections, may be used to obtain \( MA \) model parameters. The set of equations relating those slices of the \( c_{k} \) and \( c_{k-1} \) cumulants is wider and may be obtained along the same lines. The case collected here is the most useful for the third- and fourth-order case, which will be used in the simulations.

### 4 Estimation of MA coefficients: proposed algorithms

In the previous Sections, several families of equations relating the \((k-1)\)-th and \( k \)-th order cumulant sequences have been developed. These equations may be used alone or in combination to estimate \( MA \) model parameters, either in a least-squares approach or by a recursive closed-form solution. Some recursive closed-form solutions are given in the appendix of this paper. However, these solutions are more important theoretically than in practice, since they do not smooth out the effect of additive noise on the estimated sampled cumulants. In this Section we propose some combinations that are useful for obtaining robust estimates of the \( MA \) system coefficients. These coefficients are obtained by solving, in the least-squares sense, an overdetermined system of equations which is viewed to be linear in a set of parameters. The proposed algorithms, considering the usual third- and fourth-order case for notational simplicity (extension to \((k-1)\)-th and \( k \)-th order cumulants is straightforward from the preceding results) are as follows:

**Algorithm 1: Least-squares 1:** Consider the equations given by proposition 1. Taking \( l = 0 \) in eqn. 15 and \( \tau_{1} = \tau_{2} \) (diagonal slice), obtain the pair of equations:

\[
\begin{align*}
\sum_{i=0}^{q} b(i) c_{n}^{q}(\tau - i, \tau - i, 0) &= \gamma \sum_{i=0}^{q} b(i) c_{n}^{q}(\tau - i, 0) \\
\sum_{i=0}^{q} c_{n}^{q}(i - \tau, i - \tau, q) &= e b(q) c_{n}^{q}(\tau, \tau - \tau, \tau) \quad (36)
\end{align*}
\]

Forming the system of equations that result from eqn. 35 for \( \tau = \tau_{1} \ldots \tau_{q} \) and from eqn. 36 for \( \tau = \tau_{1} \ldots \tau_{q} \) obtains an overdetermined system of equations with \( 2q + 2 \) unknowns \( x = (e, eb(1), e, eb(2), \ldots, eb(q), e(b(q), b(1), \ldots, b(q))) \):
where matrix $A$ denotes the matrix shown in Fig. 1 and vector $b$ stands for

$$
\begin{bmatrix}
c_{3}(q, -q, 0) \\
c_{3}(q + 1, -q, 0) \\
c_{3}(q + 2, -q, 0) \\
\vdots \\
c_{3}(q, q, 0) \\
c_{4}(q, q, q) \\
c_{4}(q - 1, q - 1, q) \\
c_{4}(q - 2, q - 2, q) \\
\vdots \\
0
\end{bmatrix}
$$

The least-squares solution is then found as

$$
x = (A^T A)^{-1} A b
$$

and the last $q$ elements of the row vector $x$ are taken as the MA model parameter estimates. In addition, the uniqueness of the solution is guaranteed since matrix $A$ is full rank, as proved in the following proposition.

Proposition: Given the true output cumulants of the MA process, matrix $A$ defined in Fig. 1 is full rank.

Proof: Matrix $A$ can be rewritten as comprising the block matrices $A_1$, $A_2$ and $A_3$ as follows:

$$
A = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}
$$

where $A_1$ is the matrix constructed by taking the first $3q + 1$ rows and $q + 1$ columns of matrix $A$, $A_2$ contains the remaining $q + 1$ columns, and $A_3$ consists of the last $2q + 1$ rows and the last $q + 1$ columns.

The first $q + 1$ rows of $A_1$ constitute an upper triangular matrix whose determinant is $c_{3q+1}(-q, 0)$, which is nonzero since it lies on the domain of support for the cumulant sequence $c_{3}(t_1, t_2)$ for an MA($q$) system. This implies that $A_1$ is full rank. Similarly, the first $q + 1$ rows of $A_3$ are an upper triangular matrix with the determinant $c_{3q+1}(q, q, q) c_{3q+1}(-q, -q)$, which is also nonzero for an MA($q$) system and thus $A_3$ has full rank. Since $A_1$ and $A_3$ have full rank, $A$ also has full rank.

Algorithm 2: Least-squares 2: Consider eqn. 35 in algorithm 1 and eqns. 33 and 34 given in proposition 3, i.e.

$$
\sum_{i=0}^{q} b(i) c_{3q}(\tau - i, \tau - i, 0) = \sum_{i=0}^{q} b^{2}(i) c_{3q}(\tau - i, 0) \tag{41a}
$$

$$
\sum_{i=0}^{q} b(i) c_{3q}(\tau - i, \tau - i, 0) = \frac{1}{2} \sum_{i=0}^{q} b^{2}(i) c_{3q}(\tau - i, 0) \tag{41b}
$$

$$
\sum_{i=0}^{q} b^{2}(i) c_{3q}(\tau - i, \tau - i, 0) = \frac{1}{2} \sum_{i=0}^{q} b^{2}(i) c_{3q}(\tau - i, 0) \tag{41c}
$$

where $c_{3q}(q, q) = \gamma_{3q} b^{2}(q)$ has been used. Forming the system of equations than result from concatenating eqns. 41a for $\tau = -q \ldots 2q$, eqn. 41b for $\tau = -q \ldots q$, and eqn. 41c for $\tau = -q \ldots q$ obtains a system of $7q + 3$ equations with $2q + 2$ unknowns $x = (e, b^{2}(1), b^{2}(2), \ldots, b^{2}(q), \gamma_{3q} b^{2}(q)/\epsilon, b(1), \ldots, b(q))$. Again, the least-squares solution is found and the MA parameters are obtained from the last $q$ elements of vector $x$. Furthermore, the above equation yields a unique solution since matrix $A$, as can be easily proven, has full rank.

Algorithm 3: least-squares 3: Another alternative for the estimation of MA model parameters may be to solve eqn. 15 from proposition 1 by considering a specific set of slices. Thus, if this equation is concatenated for $-q \leq \tau \leq 2q$ and $1$ varying from $0$ to $l_1$, $l_1$ being the last horizontal slice in the third- or fourth-order cumulants space used to estimate the parameters $(l_1 \leq q)$. In this case we have $2q + 1$ unknowns $x = (e, b^{2}(1), b^{2}(2), \ldots, b^{2}(q), b(1), \ldots, b(q))$ and $(3q + 1)(l_1 + 1)$ equations, which can be solved in the least-squares sense. As in the preceding algorithm, the consistency (uniqueness) of this approach can be proven noting that the coefficient matrix is full rank. Then, the MA parameter estimates are obtained from the last $q$ elements in the least-squares solution for the system.

Algorithm 4: least-squares 4: To estimate the model parameters, consider the equations given by proposition 2. Concatenating eqn. 23 for $-q \leq \tau \leq 2q$ and eqn. 24 for $-q \leq \tau \leq q$, obtains a system of $5q + 2$ equations with $2q + 2$ unknowns $x = (e', e'^{2}(1), \ldots, e'^{2}(q), e'^{2}(q), e'(q), b^{2}(1), \ldots, b^{2}(q))$ and $(3q + 1)(l_1 + 1)$ equations. As before, parameter vector $x$ is obtained by solving the system in the least-squares sense. Estimated model parameters are then taken from the last $q$-elements of $x$. The system of equations gives a unique solution since the coefficient matrix is full rank, as can be proven by noting that this matrix has nonzero ‘diagonal’ elements for an MA($q$) model.

Remark 4: Other combinations not shown could be used to estimate model parameters. For example, some equations used before may be used alone, combined or added in different ways to obtain a new overdetermined system of equations. This procedure would lead to a different set of equations for estimating signal
parameters and when true statistics are available would yield correct MA parameter estimates. However, for the finite data case and noisy observations, sampled cumulants have to be used and not all situations are identical. In this case, several facts have a practical influence in the good behaviour of the estimates and should be taken into consideration, as is shown in the simulations performed in Section 5. The parameter estimates obtained as the least-squares solutions of an overdetermined system of equations depend on the specific set of third- or fourth-order cumulant slices and the relative amount of third- and fourth-order statistical information used in the estimation. Moreover, adding more equations to a given system of equations (i.e. using more slices or new equations and so increasing the size of the coefficient matrix) may not result in better estimates and in some cases will increase the estimated parameter variances. The reason is that sampled cumulants have larger variances at higher lags and fourth-order sampled cumulants show a greater variance than in the third-order case. Taking into account these points, the combinations used in algorithms 1-4 make use of a wide range of slices and, thus, of third- and fourth-order statistical information without redundancy and constitute a set of practical methods to recover MA model coefficients from noisy data.

5 Simulation results

In this Section we apply the algorithms described in Section 4 in the third- and fourth-order case to the NMP MA system identification from noisy output data. The results reported refer to the coloured gaussian noise case, since these algorithms use only higher-than-second-order cumulants and thus show their potentiality when applied to the coloured noise case. In the following simulations we also compare the performance of the proposed algorithms LS-1, LS-2, LS-3 (with \( l_i = 2 \)) and LS-4 with two existing well-studied linear approaches, the modified version of the Giannakos-Mendel method proposed by Tugnait [9] and the so-called least-squares 1 algorithm proposed in [10] and also used in [11] which involves third-order statistics over the \( q \)-slice. In the simulations below, the effects of additive noise level and data length are studied. For each test case, the system input has been chosen to be a zero-mean IID exponentially distributed series with a theoretical skewness \( \gamma_3 = 2 \) and theoretical kurtosis equal to 6. The observation \( y(n) \) is given by eqn. 2 where the additive zero-mean coloured gaussian noise has been obtained by passing zero-mean white gaussian noise through the ARMA filter given by the AR parameters \( [1, -2.2, 1.77, -0.52] \) and the MA parameters \( [1, -1.25] \) used in [11]. This noise is added to the output to produce the prescribed signal-to-noise ratio (SNR) defined

\[
SNR = 10 \log \frac{\sum_{n=1}^{N} x^2(n)}{\sum_{n=1}^{N} v^2(n)}
\]

where \( N \) stands for the data record length. The \( N \) out-
put samples were used to estimate third- and fourth-order cumulants by replacing expectations with sampled averages using the biased estimator. These estimates substituted the corresponding true third- and fourth-order cumulants in the equations in Section 4.

Test case 1 (MA (2) case): The following second-order MA model is employed:

\[ y(n) = w(n) - 2.333w(n-1) + 0.667w(n-2) + v(n) \]

used in [4, 7, 10] \[ (43) \]

with zeros located at 2 and 1/3. To reduce the realisation dependency of the simulations, 500 Monte Carlo runs of \( y(n) \) were generated and the mean of the parameter estimates and the variance were calculated for the proposed algorithms (LS-1 to LS-4) and those proposed in [9–10]. Table 1 shows the results of our simulations for three different record lengths (\( N = 1024 \), \( N = 2048 \) and \( N = 4096 \)) and three SNR values (\( SNR = 10 \text{dB}, SNR = 5 \text{dB} \) and \( SNR = 0 \text{dB} \)). As can be observed from these tables, each estimator provides better results when data record length increases and the variance of the estimates increases as the SNR decreases. Among the proposed algorithms, LS-2 and LS-3 are more robust to coloured Gaussian noise, whereas LS-4 is quite biased for this example. These algorithms, and especially LS-2, outperform the other linear methods used in [9, 10] in terms of bias and variance. To show this performance clearly, for each independent run we have also computed the average mean squared error (MSE) between the true \( b(i) \) and estimated \( \hat{b}(i) \) MA coefficients, i.e.

\[
MSE = \text{mean over all realisations} \left\{ \sum_{i=1}^{q} (b(i) - \hat{b}(i))^2 \over \sum_{i=1}^{q} \hat{b}(i)^2 \right\} \tag{44}
\]

Figs. 2 and 3 show the MSE plotted against the SNR when the proposed LS-1, LS-2 and LS-3 algorithms are employed (LS-4 is quite biased in this test and is not shown in these figures) and those of [9, 10]. Based on the results obtained we found that

- The proposed LS-1, LS-2 and LS-3 algorithms have the lowest MSE.
- LS-2 appears to outperform the other methods in terms of bias and variance.
- Estimations given by [9] are clearly biased since this method makes use of the signal autocorrelation sequence affected by the coloured Gaussian noise.
- The proposed algorithms LS-2 and LS-3 show a greater degree of robustness than the least-squares 1 in [10], which employs third-order cumulants over the \( q \)-slice. The reason for this performance is that the proposed algorithms make use of more cumulant statistics than [10], thus obtaining better accuracy on the estimates.

Test case 2 (MA (5) case): The observed signal is given by the NMP MA(5) model taken from [11]:

\[ y(n) = w(n) - 2.333w(n-1) + 0.75w(n-2) + 0.5w(n-3) \]
\[ + 0.3w(n-4) - 1.4w(n-5) + v(n) \tag{45} \]

For brevity’s sake Table 2 shows the result of a Monte Carlo experiment for three different record lengths and SNR sets up to 5 and 0dB. As this table indicates, worse results appear to be obtained for all the linear algorithms for a larger model order. As can be observed, the proposed LS-2 and LS-1 algorithms give a lower MSE and better estimates in terms of bias and variance than the other studied algorithms. Among them, the LS-2 algorithm once again provides the best results.

![Fig. 2: MSE of estimated parameters for MA(2) process \{1, -2.333, 0.667\} in test case 1 for \( N = 1024 \) (500 Monte Carlo runs) ]

![Fig. 3: MSE of estimated parameters for MA(2) process \{1, -2.333, 0.667\} in test case 1 for \( N = 2048 \) (500 Monte Carlo runs) ]

6 Concluding remarks

In this paper the relationship between the \((k-1)\text{th}\) and \(k\text{-th}\) order polyspectra for MA systems has been analysed. From a general framework, several sets of equations involving cumulants have been developed for the identification of minimum or nonminimum phase MA processes. Based on these equations we have obtained through propositions 1–3 the case of third- and fourth-order cumulants, for which both recursive closed-form solutions and batch least-squares algorithms have been presented. The least-squares versions of the proposed methods are proven to be useful for MA system identification using output measurements corrupted by coloured Gaussian noise of unknown autocorrelation sequence. As a result of the linearity of the methods, the uniqueness of solutions can usually be guaranteed.

Extensive simulations have been performed comparing the four presented algorithms with some published linear methods. From the simulations shown and from others performed by the authors, several conclusions can be drawn:

- All the studied linear methods show less bias and variance when the data record length or the SNR increases.
Table 2: MSE and estimated parameters (mean ± standard deviation) for MA(5) process in test case 2

<table>
<thead>
<tr>
<th>SNR=5dB</th>
<th>N=1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b(1)$=2.330</td>
<td>1.520±0.562</td>
</tr>
<tr>
<td>$b(2)$=0.750</td>
<td>0.239±0.441</td>
</tr>
<tr>
<td>$b(3)$=0.500</td>
<td>0.182±0.340</td>
</tr>
<tr>
<td>$b(4)$=0.300</td>
<td>0.226±0.302</td>
</tr>
<tr>
<td>$b(5)$=1.400</td>
<td>-0.864±0.451</td>
</tr>
<tr>
<td>MSE</td>
<td>0.237±0.168</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SNR=5dB</th>
<th>N=2048</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b(1)$=2.330</td>
<td>-1.780±0.289</td>
</tr>
<tr>
<td>$b(2)$=0.750</td>
<td>0.384±0.367</td>
</tr>
<tr>
<td>$b(3)$=0.500</td>
<td>0.336±0.306</td>
</tr>
<tr>
<td>$b(4)$=0.300</td>
<td>0.302±0.169</td>
</tr>
<tr>
<td>$b(5)$=1.400</td>
<td>-1.054±0.299</td>
</tr>
<tr>
<td>MSE</td>
<td>0.100±0.006</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SNR=0dB</th>
<th>N=4096</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b(1)$=2.330</td>
<td>-1.256±0.430</td>
</tr>
<tr>
<td>$b(2)$=0.750</td>
<td>0.222±0.386</td>
</tr>
<tr>
<td>$b(3)$=0.500</td>
<td>0.239±0.322</td>
</tr>
<tr>
<td>$b(4)$=0.300</td>
<td>0.274±0.246</td>
</tr>
<tr>
<td>$b(5)$=1.400</td>
<td>-0.901±0.346</td>
</tr>
<tr>
<td>MSE</td>
<td>0.247±0.168</td>
</tr>
</tbody>
</table>

Results are for 500 Monte Carlo runs.

- The proposed methods perform better overall and among them, the LS-2 algorithm appears to be the best in terms of bias, variance and MSE. This algorithm shows a good degree of robustness to noise and cumulant estimation errors, whereas LS-4 appears to be quite sensitive to cumulant estimation errors, especially for those examples whose estimated cumulant sequences provide relatively smaller absolute values compared with the true statistics.

- For the coloured gaussian noise case, the method in [9] is quite biased since it makes use of the autocorrelation sequence of the output signal, strongly perturbed by noise.

- The proposed combinations of third- and fourth-order cumulants given by the LS-1, LS-2 and LS-3 algorithms, and especially LS-2, outperform the algorithm in [10, 11], since the proposed methods exploit much higher-order statistics information than it does.

The algorithms implemented in Section 4 use the equations given in Section 3 evaluated over a specific set of slices. In addition, these equations could be implemented using all the slices allowed for an MA($q$) model. This procedure was suggested in [13] for the GM equation and it has the advantage of using much larger data sets of output statistics. In like manner, this method could be extended to the proposed equations as well. Furthermore, the proposed methods can be applied to the identification of the MA part of an ARMA model using the residual time series method [1], estimating the AR parameters using any one of the methods proposed in [16-19], or to compute both the AR and MA parameters of a noncausal ARMA model using the ‘double MA method’ [20].

Future work will be directed towards new applications of the sets of equations given in Section 3 and the use of the performance analysis in [21] for comparison among different approaches.

7 Acknowledgments

This work was partially supported by the Comisión Interministerial de Ciencia y Tecnología (CICYT) of Spain under project TIC95-0465. The authors would also like to thank the anonymous referees for carefully reading the paper and for their valuable comments that greatly improved the presentation of this paper.

8 References


9 Appendix: Closed-form solutions

Under assumptions AS1-AS3, MA model coefficients can be estimated in closed-form (using $q$ equations) from the family of equations given in Section 3. For example, set $\tau_1 = -q$ in eqn. 12 to obtain

$$c_k(q, -\tau_2, \ldots, -\tau_k - q) = c_{k+1}(q, -\tau_2, \ldots, -\tau_k - q)$$

Next, set $\tau_1 = -q + 1$ in eqn. 12 and use eqn. 46 to obtain

$$b(1)c_{k+1}(q, 1, -\tau_2, \ldots, -\tau_k - q) + c_{k+2}(q, 1, -\tau_2, \ldots, -\tau_k - q) = cb(q)c_{k+1}(q, 1, -\tau_2, \ldots, -\tau_k - q)$$

$\Rightarrow b(1) = c_{k+1}(q, -\tau_2, \ldots, -\tau_k - q)c_{k+1}(q, 1, -\tau_2, \ldots, -\tau_k - q)
 \quad - \quad \frac{c_{k+2}(q, 1, -\tau_2, \ldots, -\tau_k - q)}{c_{k+1}(q, 1, -\tau_2, \ldots, -\tau_k - q)}$  

Continuing this way, one can successively obtain $b(m)$ $(1 \leq m \leq q)$ from the previously obtained $b(k)$ $1 \leq k \leq m-1$ as follows:

$$b(m) = c_{k+1}(q, -\tau_2, \ldots, -\tau_k - q)c_{k+1}(q, m - \tau_2, \ldots, -\tau_k - q)
 \quad - \quad \sum_{i=0}^{m-1} b(i)c_{k+1}(q, m + i, -\tau_2, \ldots, -\tau_k - q)$$

This solution is well-conditioned since the only things done are with $c_k(q, m - \tau_2, \ldots, -\tau_k - q)$, $m = 0, \ldots, q$ which are nonzero provided that the specific slice $(\tau_2, \tau_3, \ldots, \tau_k - 1)$ is chosen to lie in the domain of support for an MA$(q)$ model $(m - q \leq \tau_2, \tau_3, \ldots, \tau_k - 1 \leq m)$.

Similarly, one can use eqn. 22 to establish another closed-form solution:

$$\tau_1 = q$$

$$\gamma_{k-1w} b(q) = c_{k-1w}(q, \tau_2, \ldots, -\tau_k - q)$$

$$\tau_1 = q - m$$

$$b^2(m) = \gamma_{k-1w} b(q)$$

and the MA parameter are obtained from eqn. 49 by taking $(b^2(m))^{1/2}$. Similar closed-form solutions can be obtained from eqn. 27a repeating the above derivations. In like manner, equations involving $b^2(i)$ $i = 1, \ldots, q$ like eqn. 27b may be used in conjunction with eqn. 11, successively obtaining $b^2(i)$ $i = 1, \ldots, q$. Once $b^2(i)$ is obtained by the above procedure, one can recover $b$ and $b(i)$ $i = 1, \ldots, q$ from eqn. 11 recursively. These procedures are always well-conditioned since we only take divisions by nonzero valued cumulants for an MA$(q)$ system.