ANALYTIC PERFORMANCE EVALUATION OF CUMULANT-BASED ARMA SYSTEM IDENTIFICATION METHODS

José A. R. Fonollosa and Josep Vidal

E.T.S E. Telecomunicació, Universitat Politècnica de Catalunya, Apdo. 30.002, Barcelona 08080, Spain.

ABSTRACT

In this paper, we perform an analytic study of some of the recently developed cumulant-based methods for estimating the AR parameters of ARMA processes. Our analysis includes new AR identifiability results for pure AR process and the analytic performance evaluation of system identification methods based on cumulants. We present examples of pure AR process that are not identifiable via the normal equations based on the diagonal third-order cumulant slice. The results of the performance evaluation are illustrated graphically with plots of the variance of the estimates as a function of the parameters of the process.

1. INTRODUCTION

Recently, several system identification methods based on higher-order statistics have been proposed in the literature [1,2]. In almost all the cases, the performance of these methods has been evaluated only through Monte Carlo simulations and for a limited number of examples. These simulations are clearly insufficient to predict the general behavior of the algorithms. The development of analytic analysis tools seems the best way to gain an in-depth understanding of cumulant-based methods.

In [4,8], the expressions of the covariances of sample cumulants were derived and several MA System Identification methods were compared for complete ranges of parameter values and signal to noise ratios. The information carried by different sets of statistics was also studied. In this paper, we consider the estimation of the AR coefficients of ARMA process instead of the MA coefficients.

Apart form the asymptotic performance evaluation of the algorithms we present new AR identifiability results. It is well-known that, in general, consistent AR estimates cannot be obtained via the normal equations from a single cumulant slice of an ARMA process [6]. Nevertheless, no examples involving pure AR processes had been presented until now, and the issue of whether or not more than one slice is needed to estimate the coefficients of a *truly* AR(p) system was an open question [2]. *Section 2* introduces the notation of the paper and the cumulant-based normal equations. In *Section 3* we present examples of AR processes that are not identifiable using only the diagonal slice in the cumulant-based normal equations.

2. ARMA SYSTEM IDENTIFICATION

Let us consider a causal stable ARMA process x(n) described by

$$\sum_{i=0}^{p} a(i)x(n-i) = \sum_{i=0}^{q} b(i)u(n), \quad a(0) = b(0) = 1 \tag{1}$$

where the input u(n) is an i.i.d. non-Gaussian sequence with m-th order cumulants γ_m.

Assuming there are no pole-zero cancellations, the irreducible transfer function of the ARMA(p,q) model is given by

$$H(z) = \frac{\sum_{i=0}^{q} b(i)z^{-i}}{\sum_{i=0}^{p} a(i)z^{-i}} = \frac{B(z)}{A(z)} \tag{2}$$

The Barlett-Brillinger-Rosenblatt summation formula

$$C_{m,x}(i_1, i_2, \ldots, i_{m-1}) = \gamma_m \sum_{n=-\infty}^{\infty} \prod_{k=0}^{m-1} h(n+i_k), \; i_0=0$$

relates the m-th order cumulant of $x(n)$ to the impulse response $h(n)$. From this formula, it is easy to obtain the following equations

$$\sum_{j=0}^{p} a(j)C_{m,x}(t-j, i_2, \ldots, i_{m-1})) = 0 \quad t > q \tag{3}$$

which collected for $t = q+1, \ldots, q+p$, form the cumulant-based normal equations

$$R(i)\, \boldsymbol{a} = -\boldsymbol{b}(i)$$
$$\boldsymbol{a} = (a(1), \ldots, a(p))^t \tag{4}$$

This work was supported by the Spanish government, under grant TIC92-0800-C05-04

where $R(i)$ is a p by p matrix with $C_{m,x}(q+j-i,\ i_2,\ ...,\ i_{m-1})$ as its (j,i) element, and the vector $b(i)$ has $C_{m,x}(q+j,\ i_2,\ ...,\ i_{m-1})$ as its j-th element.

For $m = 2$ we have the well-known and widely used autocorrelation-based normal equations. Nevertheless, these equations are not useful if the system have inherent all-pass factors, or we have additive Gaussian noise of unknown power spectral density. In these cases, we can make use of higher-order cumulants if the process is nonGaussian.

Apart from the normal equations, the AR parameters can be determined by other approaches [2]. The cumulant-matching method is a general procedure based on the minimization of squared differences between the observed cumulants and the cumulants of the proposed method. In *Section 4* we will compare the performance of linear methods based on the normal equations and the cumulant-matching method.

3. AR IDENTIFIABILITY

When using the cumulant-based normal equations to estimate the AR parameters we can consider a single 1-D cumulant slice, i.e., cumulants with fixed lags i_k $(k=2,\ ...,\ m-1)$ or a set of cumulant slices. A 1-D slice is called a *full-rank* slice when the associated matrix $R(i)$ has rank p. From (4) it is clear that consistent estimates cannot be obtained from a single slice if it is not a full-rank slice.

Until recently, no result was available about the minimum number of slices required to assure a consistent estimation. Swami and Mendel [6], as well as Giannakis [7] showed that the AR coefficients of an ARMA(p,q) process can always be determined using p+1 cumulant slices of the *m*th-order cumulant. In [6], ARMA examples were presented showing that 1) every 1-D cumulant slice need not be a full rank slice, and 2) a full rank cumulant slice may not exist.

For pure AR models, the issue of whether or not more than one slice is needed to estimate the coefficients was an open question. In this paper, we complete those previous results with examples of AR processes that are not identifiable via the normal equations based on a single slice.

Our approach to study the rank of the 1-D cumulant slices is basically the same followed by Swami and Mendel in [6].

Let us define the z-transform of a 1-D cumulant slice

$$C_{m,x}(t;\ k) = C_{m,x}(-t,\ k,\ 0,\ ...,\ 0) \tag{5}$$

as

$$C_{m,x}(z;\ k) = \sum_{t=-\infty}^{\infty} C_{m,x}(t;\ k) z^{-t} = \gamma_m H_{m-1}(z;\ k) H(z^{-1}) \tag{6}$$

where

$$H_1(z;\ 0) = H(z) \tag{7}$$

and

$$H_{m-1}(z;\ k) = \sum_{t=-\infty}^{\infty} h^{m-2}(t) h(t+k) z^{-t} = z^k H(z) * H_{m-2}(z;\ 0) \tag{8}$$

where $*$ denotes complex convolution.

From equation (6) we can express the 1-D slice normal equations in the z domain as

$$C_{m,x}(z;\ k) A(z^{-1}) = \gamma_m H_{m-1}(z;\ k) B(z^{-1}) \tag{9}$$

We note that if $A(z^{-1})$ and the numerator of $H_{m-1}(z;\ k)$ have common factors, this slice will not be sufficient for the identification of the AR parameters, i.e., the recursion will hold with a minimal order less than p. In the following example, we study the identifiability of AR processes.

Example. Let us consider an AR(3) model

$$H(z) = \frac{1}{(1 - p_1 z^{-1})(1 - p_2 z^{-1})(1 - p_3 z^{-1})} \tag{10}$$

For this model, evaluating $H_2(z;\ 0)$ with the help of a symbolic mathematical package, we obtain

$$H_2(z;\ 0) = \frac{z^6 + (p_1p_2+p_1p_3+p_2p_3)z^5 - p_1p_2p_3(p_1+p_2+p_3)z^4 + p_1^2p_2^2p_3^2z^3}{(z - p_1^2)(z - p_1p_2)(z - p_2^2)(z - p_1p_3)(z - p_2p_3)(z - p_3^2)} \tag{11}$$

If the numerator of $H_2(z;0)$ has the factor $(1 - p_1^{-1}z^{-1})$ the diagonal slice $C_{3,x}(t;\ 0)$ will not be a *full-rank* slice. We can search for this family of models if we set $z = p_1^{-1}$ in the numerator, and we solve the resulting equation

$$1 + p_1^2(p_2+p_3) + p_1p_2p_3 - p_1^4p_2p_3 + p_1^3p_2p_3(p_2+p_3) + p_1^5p_2^2p_3^2 = 0 \tag{12}$$

We have found solutions to equation (12) with a real pole $p_1 = p$, and a pair of complex conjugates poles

$$p_2 = r + i\,s; \quad p_3 = p_2^* = r - i\,s; \quad (i = \sqrt{-1})$$

where r and s are the real and the imaginary part of p_2, respectively. This substitution leads to the following biquadratic equation in s

$$a\,s^4 - b\,s^2 + c = 0 \tag{13}$$

where

$$a = -p^5 \tag{14}$$

$$b = p - 2rp^3 - p^4 - 2r^2p^5 \tag{15}$$

$$c = 1 + r^2p + 2rp^2 - 2r^3p^3 - r^2p^4 - r^4p^5 \tag{16}$$

Equation (13) have real solutions for p, r and s that satisfy $0 < p^2 < 1$ and $0 < (r^2 + s^2) < 1$. For example, if we set $p = -0.8$ and $r = -0.15$, one of the solutions for s is 0.8327. The resulting model

$$p_1 = -0.8, \quad p_{2,3} = -0.1500 \pm 0.8327 \tag{17}$$

$$(\,a(1) = 1.1,\ a(2) = 0.9559,\ a(3) = 0.5727\,) \tag{18}$$

is not identifiable via the third-order diagonal normal equations.

Figure 1 shows, for values of the real pole p_1 ranging between -0.5 and -0.9, a plot of all the locations of p_2 on the z-plane for which $H_2(z;\,0)$ has a zero in $z = p_1^{-1}$. For all those families of AR(3) models, the third-order diagonal slice is not full-rank slice.

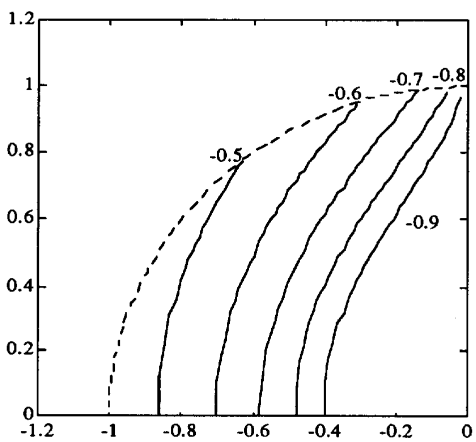

Figure 1. Families of AR(3) models $H(z) = 1/((1-p_1z^{-1})(1-p_2z^{-1})(1-p_2^*z^{-1}))$, for which the third-order diagonal slice $C_3(n,n)$ is not a full-rank slice. Each plot represents the positions of p_2 in the z-plane for a fixed value of the real pole p_1= -0.5, -0.6, -0.7, -0.8, -0.9.

4. ANALYTICAL PERFORMANCE EVALUATION

The first and most difficult step encountered in the analytic study of cumulant-based methods is the computation of the covariances of higher-order sample cumulants. This problem was first addressed in [3] for third-other sample moments. In [4,8], that work is completed with the expressions for fourth-order moments and cumulants, and including an analysis of the effect of noise.

The expressions of the covariances derived in [4,8] were used to perform a comparative analysis of several MA system identification methods based on cumulants. For MA processes the terms appearing in the expressions of the covariances can be computed directly since only finite summations are involved. For ARMA processes the exact or symbolic computation of the covariances terms is quite more complicated. The covariances of sample moments of ARMA processes were obtained in [3] in the third-order case using a state-space notation for $h(n)$ and the Kronecker product. Approximate results may be also obtained considering only a finite number of terms in the summations. For ARMA processes, this truncation of the cumulant sequences seem the easiest way to obtain the covariances of fourth-order cumulants and to include the effect of noise in both the third- and fourth-order case.

Let us consider the analytic performance evaluation of a system identification method defined as a vector-valued function $\boldsymbol{g}(.)$ of the sample cumulants. Let $\theta_N = \boldsymbol{g}(\boldsymbol{c})$ be the estimated parameters and $\theta = \boldsymbol{g}(\boldsymbol{C})$ the true parameters of a process, where $\boldsymbol{c}$ are the estimated cumulants and $\boldsymbol{C} = \boldsymbol{f}(\theta)$ the true cumulants. The asymptotic covariance of the estimated parameters is given by

$$\boldsymbol{P}(\theta) = \lim_{N \to \infty} N\,E[(\theta_N - \theta)(\theta_N - \theta)^t]$$
$$= \boldsymbol{G}(\theta)\,\Sigma(\theta)\,\boldsymbol{G}(\theta)^t \tag{19}$$

where $\boldsymbol{G}(\theta)$ is the Jacobian matrix of $\boldsymbol{g}(\cdot)$, evaluated at $\boldsymbol{C} = \boldsymbol{f}(\theta)$; and $\Sigma(\theta)$ is the asymptotic covariance of the sample cumulants, i.e.,

$$\Sigma(\theta) = \lim_{N \to \infty} N\,E[(\boldsymbol{c} - \boldsymbol{C})(\boldsymbol{c} - \boldsymbol{C})^t] \tag{20}$$

The expression of the Jacobian matrix is usually easy to obtain by standard differential analysis. Explicit expressions for the cumulant-matching method and the least squares and minimum-norm solutions of linear methods can be found in [8]

Example. We have compared the performance of five different system identification methods in the estimation of the parameters of a AR(3) process. The driving input is a zero-mean, exponentially distributed, i.i.d sequence.

Let us define the following two sets of cumulants:

$$D = \{C_{3,x}(i,i);\ i = -3, -2, ..., 3\}$$

$$P = \{C_{3,x}(i,j);\ i = 0,1, 2, 3;\ j = i, i+1, i+2, i+3\}$$

Figure 2 compares the performance of three methods based on the normal equations. YWD corresponds to the cumulant-based normal equations using only the diagonal slice, YWP to the cumulant-based normal equations using $p+1=4$ slices [2], and WSP to the approach described in [9]. As the YWP method, the WSP method combines $p+1$ slices, but with a smaller computational cost. The YWD methods uses the set D of sample cumulants, while both the YWP and WSP use the set P.

Figure 3 analyzes the performance of the cumulant-matching approach in the estimation of the AR coefficients. The plots labelled CMD and CMP corresponds to the cumulant-matching method based on the diagonal set D and the cumulant-matching method based on the set P, respectively.

The asymptotic standard deviation in the estimation of a_3 is compared in both figures for different values of the coefficients. For these plots AR(3) models of equal reflection or lattice coefficients were considered, i.e.,

$$-1 < k_1 = k_2 = k_3 = a_3 < 1$$

Observe that the performance of the normal equations using only the diagonal slice (YWD) has an important degradation when the coefficient a_3 is greater than 0.45. In fact, the peak in the variance corresponds to the example of *Section 3* of an AR(3) system no identifiable via the diagonal normal equations. On the other hand, when $p+1$ slices are considered (YWP and WSP), the performance is close to the nonlinear cumulant-matching approach for almost all the range of parameter values.

Figure 2 indicates that the AR(3) model is still identifiable from the diagonal cumulants if we use the cumulant-matching approach.

In general, the cumulant-matching methods do not seem to provide a significant increase in performance respect to the normal equations in the estimation of the AR parameters. Moreover, it suffers from the usual problems associated with nonlinear optimization methods: computational complexity and convergence to a local minimum.

5. CONCLUSIONS

The cumulant-based methods for the estimation of the AR parameters of ARMA processes have been studied analytically. The results confirm the need of considering $p+1$ slices in the normal equations for both ARMA and pure AR processes. Other system identification approaches as the cumulant-matching approach does seem to provide more accurate estimates for causal AR processes.

References

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

[2] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system sheory: theoretical results and some applications", *Proc. IEEE*, vol. 79, no. 3, pp. 278-305, March 1991.

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

[4] José A. R. Fonollosa and J. M. Mendel, "Analytic Performance Evaluation of Cumulant-based FIR System Identification Methods", *Proc. IEEE Sixth SP Workshop on Statistical Signal & Array Processing*, 1992.

[5] J. Tugnait, "Fitting noncausal AR signal plus noise models to noisy nonGaussian linear processes", *IEEE Trans. Automat. Contr.*, vol. 32, pp. 547-552, 1987.

[6] A. Swami and J. M. Mendel, "AR identifiability using cumulants", in *Proc. Workshop on Higher-Order Spectral Analysis,* Vail, CO, 1989, pp. 1-6.

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

[8] José A. R. Fonollosa, "Sample cumulants of Stationary Processes: Asymptotic Results", submitted for publication, 1993.

[9] Josep Vidal and José A. R. Fonollosa, "Causal AR modelling using a linear combination of cumulant slices", submitted for publication, 1992.

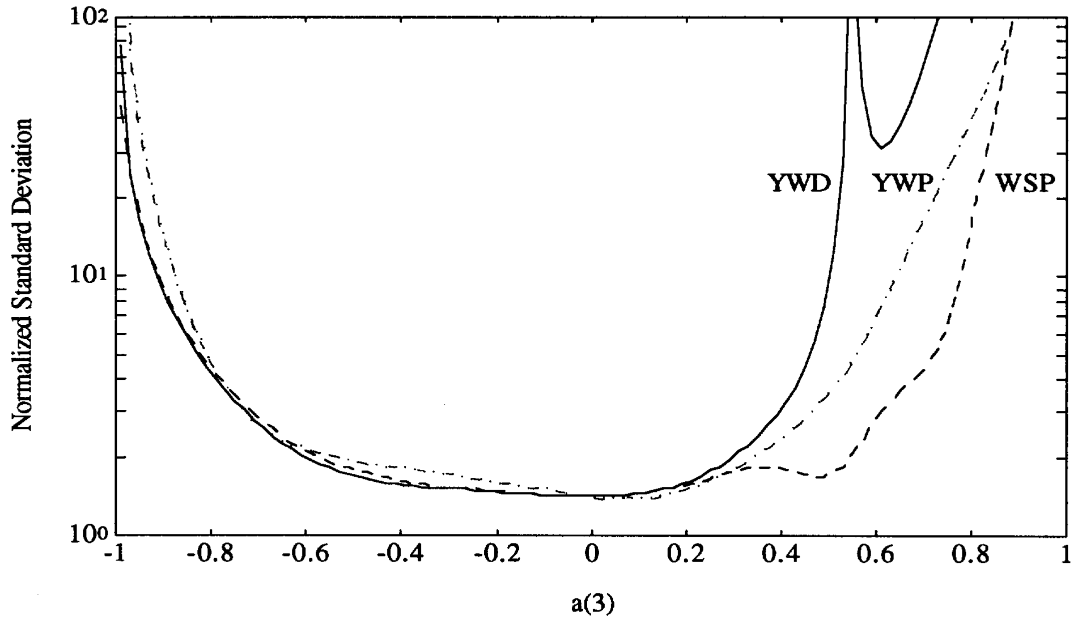

Figure 2. Asymptotic standard deviation in the estimation of a_3 as a function of its value. YWD and YWP: cumulant-based normal equations using the diagonal slice and four slices respectively. WSP: w-slice approach.

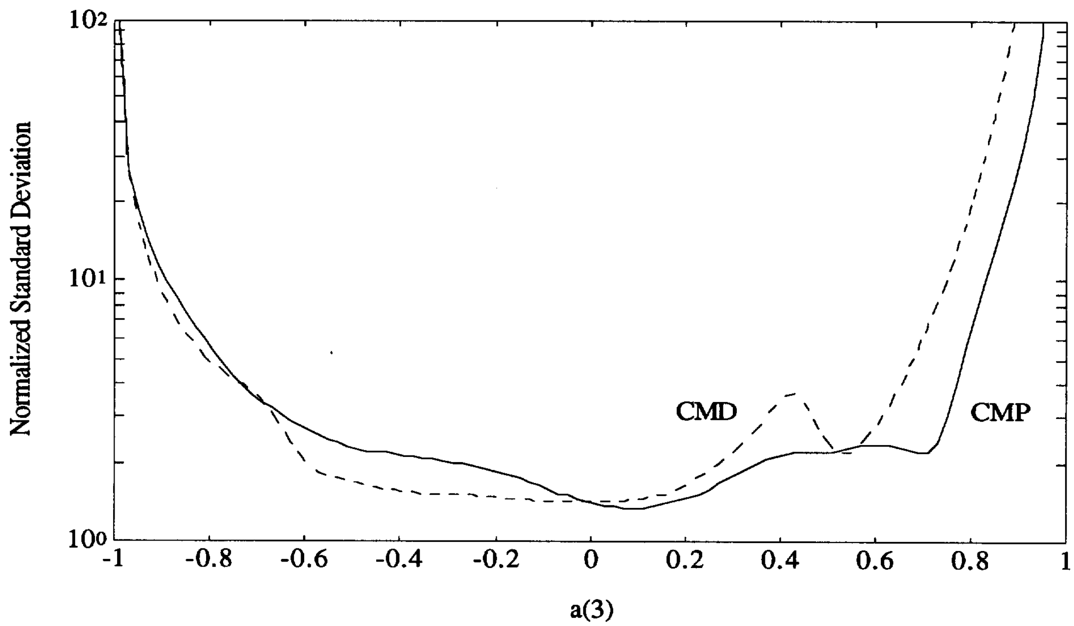

Figure 3. Asymptotic standard deviation in the estimation of a_3 as a function of its value. CMD and CMP: cumulant-matching approach based on the cumulants of the diagonal slice and *p+1* slices respectively.