ANALYTIC PERFORMANCE EVALUATION OF CUMULANT-BASED FIR SYSTEM IDENTIFICATION METHODS

José A. R. Fonollosa[1] and Jerry M. Mendel[2]

[1]E.T.S E. Telecomunicació, Universitat Politècnica de Catalunya, Apdo 30.002, Barcelona 08080, Spain.
[2]Signal and Image Processing Institute (EEB 400), University of Southern California, Los Angeles, CA 90089, USA

ABSTRACT

The covariances of the third- and fourth-order sample cumulants of stationary processes are derived. The resulting expressions are then used to obtain the analytical performance of FIR System Identification methods as a function of the coefficients and the statistics of the input sequence. The lower bound in the variance is also compared for different sets of sample statistics to provide insight about the information carried by each sample statistic. Finally the effect that the presence of noise has in the accuracy of the estimates is also studied analytically. The results are illustrated graphically with plots of the variance of the estimates as a function of the parameters or the signal to noise ratio. Monte Carlo simulations are also included to compare their results with the predicted analytical performance.

1. INTRODUCTION

Several parametric cumulant-based system identification methods are now available in the literature [1,2]. Nevertheless, in almost all cases, the performance of these methods has been evaluated only through Monte Carlo simulations and for a limited number of examples. Clearly, these simulations are insufficient to predict the general behavior of cumulant-based algorithms. The purpose of our work is to develop the analytic tools required to perform the asymptotic performance evaluation of algorithms based on second-, third-, and fourth-order sample cumulants. This paper does a fast presentation of the results obtained so far. In the journal version, we will include more details about the derivation of the formulas and more examples of their applications.

The first and most difficult step encountered in the analytic study of cumulant-based methods is the calculation of the covariances of higher-order moment estimates. This problem was addressed in part in [3] where the asymptotic covariances of second and third-order sample moments were derived. Here, those results are completed with the covariances of sample fourth-order moments and both the covariances of third- and fourth-order sample cumulants are obtained from the derived sample moments covariances. The effect of noise is also considered.

The derivation of the covariances requires the computation of up to eighth-order moments which involves several hundred terms. Although it is possible to work to these terms in a compact form [7], hand calculation and simplification of explicit expressions is a very tedious task. Symbolic algorithms for *Mathematica* [8] have been used to avoid hand derivations and obtain simplified expressions. With this symbolic math package it is not difficult to obtain the general expression of moments of any order as a function of the cumulants of the process. Then, we can program the steps followed in [3] and additional simplification rules to obtain expressions of the covariances of sample moments. Using the relation between sample cumulants and sample moments, the analytic simplified expression of the covariances of sample cumulants can also be obtained.

2. MOMENTS AND CUMULANTS

Let $\{x_1, x_2, ..., x_k\}$ be a collection of random variables. The moments of these random variables will be denoted as

$$mom[\{x_1, x_2, ..., x_m\}] = E[x_1 x_2 ... x_m]$$

With this notation, cumulants can be defined as a function of moments with the expression

$$cum[\{x_1,...,x_m\}] = \sum_{U_{p=1}^q I_p = I} (-1)^{q-1}(q-1)! \prod_{p=1}^q mom[I_p] \quad (2.1)$$

where the summation extends over all partitions of set I. This moment-to-cumulant equation can be inverted to obtain the cumulant-to-moment formula

$$mom[\{x_1, x_2, ..., x_m\}] = \sum_{U_{p=1}^q I_p = I} \prod_{p=1}^q cum[I_p] \quad (2.2)$$

A complete example which illustrates the use of (2.1) and (2.1) for the case $m=4$ can be found in the Appendix A of [2]. For that case, the number of partitions is 15, but it increases very fast with the order and for $m=8$, the number of partitions is 4140. For zero mean variables the

partitions that have a subset with one element can be discarded. Even in this case, the number of terms is still 715 for $m=8$.

Stationary Linear Processes

The mth-order moment of a strictly stationary random process $x(t)$, denoted by $M_{m,x}(i_1, i_2, ..., i_{m-1})$, is defined as the joint mth-order moment of the random variables $\{x(t), x(t+i_1), x(t+i_2), ..., x(t+i_{m-1})\}$

$$M_{m,x}(i_1, i_2, ..., i_m) = mom[\{x(t), x(t+i_1), ..., x(t+i_m)\}]$$
$$(2.3)$$

The mth-order cumulant is similarly defined as

$$C_{m,x}(i_1, i_2, ..., i_m) = cum[\{x(t), x(t+i_1), ..., x(t+i_m)\}]$$
$$(2.4)$$

Due to the stationarity of the process, the right side of (2.3) and (2.4) are independent of t, i.e., the mth-order moment and cumulant are only a function of the $m-1$ lags $i_1, i_2, ..., i_{m-1}$.

Linear processes are defined as the output $y(t)$ of a linear system whose input is a sequence of independent, identically distributed (i.i.d.) random variables.

$$y(t) = \sum_{n=-\infty}^{\infty} h(n)v(t-n)$$
$$(2.5)$$

We will assume that the impulse response $h(t)$ is exponentially stable, and the cumulants γ_m of the i.i.d. input $v(t)$ are finite. Then, the cumulants of the linear stationary process $y(t)$ are related to the impulse response $h(t)$ by the Bartlett-Brillinger-Rosenblatt formula

$$C_{m,y}(i_1, i_2, ..., i_{m-1}) = \gamma_m H_m(i_1, i_2, ..., i_{m-1}) \quad (2.6)$$

where

$$H_m(i_1, i_2, ..., i_{m-1}) = \sum_{n=-\infty}^{\infty} \prod_{k=0}^{m-1} h(n+i_k) \quad ; \ (i_0=0)$$
$$(2.7)$$

The output may be observed in presence of additive noise $w(t)$ which will be assumed to be independent of the input and summable. Under the above assumptions, it is well-known that the cumulants of the resulting process

$$x(t) = y(t) + w(t)$$
$$(2.8)$$

are the sum of the cumulants of each term

$$C_{m,x}(i_1, i_2, ..., i_{m-1}) = \gamma_m H_m(i_1, i_2, ..., i_{m-1}) + C_{m,w}(i_1, i_2, ..., i_{m-1})$$
$$(2.9)$$

It is also straightforward to check that $C_{m,x}(...)$ is finite and infinite summation over any or all of the indices is also finite.

3. SAMPLE MOMENTS AND CUMULANTS

The sample moments of a process $x(t)$ can be defined in different ways depending on the kind of data-windowing considered. Asymptotically, all those definitions are equivalent, i. e., they all converge with probability one to the true moments with the same asymptotic covariance.

Sample cumulants are then computed from (2.1, 2.3, 2.4) with the sample moments $m_i(...)$ instead of the true moments $M_{i,x}(...)$. In practice, the sample mean m_1 is subtracted from $x(t)$ and the simpler but equivalent formulas that follow are used to compute the sample cumulants:

$$x_o(t) = x(t) - m_1$$
$$(3.1)$$

$$c_2(i) = \frac{1}{N} \sum_{t=1}^{N} x_o(t)x_o(t+i)$$
$$(3.2)$$

$$c_3(i,j) = \frac{1}{N} \sum_{t=1}^{N} x_o(t)x_o(t+i)x_o(t+j)$$
$$(3.3)$$

$$c_4(i,j,k) = \frac{1}{N} \sum_{t=1}^{N} x_o(t)x_o(t+i)x_o(t+j)x_o(t+k) -$$
$$c_2(i)c_2(k-j) - c_2(j)c_2(k-i) - c_2(k)c_2(j-i)$$
$$(3.4)$$

Observe that, even if the process is known to be zero-mean, we will consider that the sample mean is subtracted from the given samples of the process. This procedure is usually recommended in the literature [1]. We have also observed that this subtraction simplifies the expressions of the covariances of sample cumulants.

Let $m_N=(m_i(...), m_j(...), ...)^t$ be a vector of sample moments, $M=(M_{i,x}(...), M_{j,x}(...), ...)^t$ the vector with the same-indexed true moments, and $g(m_N)$ be a vector-valued continuous function with continuous and bounded partial derivatives of first and second orders in some open neighborhood of M, then $s_N = g(m_N)$ converges almost surely to $S = g(M)$, i.e.

$$\lim_{N \to \infty} s_N = S \quad a.s.$$

and its asymptotic covariance is given by the following theorem

Theorem 1:

Under the above assumptions,

$$\lim_{N \to \infty} N E[(s_N - S)(s_N - S)^t] = G(M) \Sigma(M) G^t(M)$$

where $G(M)$ is the Jacobian matrix of $g(m)$, evaluated at $m = M$.

This theorem allows one to obtain the asymptotic performance of any method based on higher-order statistics form the asymptotic covariance of the sample moments $\Sigma(M)$ and the Jacobian of the method respect to these statistics $G(M)$.

Clearly, *Theorem 1* can also be applied to a function of s_N. Since most of the methods based on higher-order statistics deal directly with cumulants instead of moments, it will be easier to work with the covariances of the sample cumulants and the Jacobian of the method with respect to these statistics. The computation of the asymptotic covariances of the sample moments and cumulants is discussed in the following section. Then, in *Section 5*, several MA System Identification methods are analyzed. Due to the limitations of space, only the results in the third-order case are included.

4. ASYMPTOTIC COVARIANCES OF THE SAMPLE MOMENTS AND CUMULANTS OF STATIONARY PROCESSES.

4.1. Sample moments

The asymptotic cross-covariance of two sample moments $m_i(a_1,...,\ a_{i-1})$ and $m_j(b_1,...,\ b_{j-1})$ is given by the summation [3]

$$\lim_{N \to \infty} N\ cov\{m_i(a_1,...,\ a_{i-1}),\ m_j(b_1,...,\ b_{j-1})\} =$$

$$\sum_{t=-\infty}^{\infty} cov\{x(0)x(a_1)..x(a_{i-1}),\ x(t)x(t+b_1)...x(t+b_{j-1})\} =$$

$$\sum_{t=-\infty}^{\infty} [M_{i+j,x}(a_1,...,\ a_{i-1},\ t,\ t+b_1,...,\ t+b_{j-1}) -$$
$$M_{i,x}(a_1,...,\ a_{i-1})M_{j,x}(b_1,...,\ b_{j-1})] \qquad (4.1)$$

and it is finite under the assumptions of section 2.

Observe that the evaluation of the above formula requires the computation of moments whose order is equal to the sum of the orders of each sample moment. Hence, if we are interested in the covariances of sample fourth-order moments we have to be able to express the eighth-order moments of a process as a function of its parameters. For linear systems, the computation of these moments as a function of the impulse response is performed through the Bartlett-Brillinger-Rosenblatt and cumulant-to-moment formulas (2.2) and (2.7). Although the catalogue of complementary set partitions provided by McCullagh in [7] may simplify this task, symbolic packages as Mathematica [8] seem to be the best tool to deal with the huge number of cumulants required to compute sixth-, seventh- and eighth-order moments of linear processes.

Explicit expressions for sample moments of order three or less were derived in [3]. We have completed that work with the expressions of the covariances of sample fourth-

order moments and the cross-covariances of these moments with others of lower order. Even if the process is symmetric, i.e., its odd cumulants are zero, the resulting formulas in the fourth-order case have hundreds of terms. Nevertheless, since most of the methods based on higher-order statistics deal directly with cumulants instead of moments, the covariances of sample moments are interesting only as an intermediate step in the calculation of the covariances of the sample cumulants.

4.2. Sample cumulants

Theorem 1, with sample cumulants as the function s_N of the sample moments, gives the asymptotic expression of the covariances of sample cumulants.

For zero-mean processes the covariances of sample second-order moments are equal to those of sample cumulants, i.e., the subtraction of the sample mean does not affect the asymptotic covariance of the autocorrelation. This is not the case when third-order statistics are considered. The following equations show a reduced number of terms compared to those presented in [3] for third-order moments of zero mean processes. Moreover, the expressions presented here are valid for any stationary process and they do not depend on the mean of the process.

Let us denote

$$G_5(i, j, k) = \sum_{t=-\infty}^{\infty} C_{5,x}(i,\ t,\ t+j,\ t+k) \qquad (4.1)$$

$$G_{23}(i, j) = \sum_{t=-\infty}^{\infty} C_{2,x}(t)C_{3,x}(t+i,\ t+j) \qquad (4.2)$$

$$G_6(i, j, l, m) = \sum_{t=-\infty}^{\infty} C_{6,x}(i,\ j,\ t,\ t+l,\ t+m) \qquad (4.3)$$

$$G_{222}(i, j) = \sum_{t=-\infty}^{\infty} C_{2,x}(t)C_{2,x}(t+i)C_{2,x}(t+j) \qquad (4.4)$$

$$G_{24}(c, i, j) = \sum_{t=-\infty}^{\infty} C_{2,x}(t)C_{4,x}(c,\ t+i,\ t+j) \qquad (4.5)$$

$$G_{33}(c, i, j) = \sum_{t=-\infty}^{\infty} C_{3,x}(c,\ t)C_{3,x}(t+i,\ t+j) \qquad (4.6)$$

The asymptotic expression for the cross-covariances of the sample third-order cumulants and the sample second-order cumulants of stationary processes is

$$\lim_{N \to \infty} N\ cov\{c_2(k),\ c_3(l,m)\} = G_5(k,l,m) +$$
$$G_{23}(l-k,m-k) + G_{23}(-l-k,m-l-k) +$$
$$G_{23}(-m-k,l-m-k) + G_{23}(l+k,\ m+k) +$$
$$G_{23}(-l+k,m-l+k) + G_{23}(-m+k,l-m+k) \qquad (4.7)$$

and the asymptotic expression for the covariances of the sample third-order cumulants of stationary processes is

$$\lim_{N \to \infty} N\, cov\{c_3(k,l),\ c_3(l,m)\} = G_6(k,l,m,n) +$$

$$G_{222}(m\text{-}k,n\text{-}l) + G_{222}(n\text{-}k,m\text{-}l) +$$
$$G_{222}(m\text{+}k,\ n\text{+}k\text{-}l) + G_{222}(m\text{+}l,n\text{+}l\text{-}k) +$$
$$G_{222}(n\text{+}k,m\text{+}k\text{-}l) + G_{222}(n\text{+}l,m\text{+}l\text{-}k) +$$
$$G_{33}(k,\ m\text{-}l,\ n\text{-}l)\ + G_{33}(k,\text{-}l\text{-}m,n\text{-}l\text{-}m) +$$
$$G_{33}(k,\text{-}l\text{-}n,m\text{-}l\text{-}n) + G_{33}(l,m\text{-}k,n\text{-}k) +$$
$$G_{33}(l,\text{-}k\text{-}m,n\text{-}k\text{-}m) + G_{33}(l,\text{-}k\text{-}n,m\text{-}k\text{-}n) +$$
$$G_{33}(l\text{-}k,k\text{-}n,m\text{+}k\text{-}n) + G_{33}(l\text{-}k,k\text{-}m,n\text{+}k\text{-}m) +$$
$$G_{33}(l\text{-}k,k\text{+}m,k\text{+}n) + G_{24}(l\text{-}k,m\text{-}k,n\text{-}k) +$$
$$G_{24}(l\text{-}k,\text{-}m\text{-}k,n\text{-}m\text{-}k) + G_{24}(l\text{-}k,\text{-}n\text{-}k,m\text{-}n\text{-}k) +$$
$$G_{24}(l,m\text{+}k,n\text{+}k) + G_{24}(l,\text{-}m\text{+}k,n\text{-}m\text{+}k) +$$
$$G_{24}(l,\text{-}n\text{+}k,m\text{-}n\text{+}k) + G_{24}(k,m\text{+}l,n\text{+}l) +$$
$$G_{24}(k,\text{-}m\text{+}l,n\text{-}m\text{+}l) + G_{24}(k,\text{-}n\text{+}l,m\text{-}n\text{+}l) \qquad (4.8)$$

For linear processes, (2.9) is used to obtain the final expressions of the covariances as a function of the impulse response and the cumulants of the noise. If the noise is Gaussian, G_5 and G_6 can be expressed simply as

$$G_5(k,l,m) = \gamma_5 H_2(k)H_3(l,m)$$

$$G_6(k,l,m,n) = \gamma_6 H_3(k,l)H_3(m,n) \qquad (4.9)$$

The covariances of sample fourth-order cumulants will not be included here. Nevertheless we can point out that the resulting expressions show much fewer terms that those corresponding to sample moments.

5. PERFORMANCE ANALYSIS

Theorem 1 and the discussion following it allows one to compute the covariance of the estimates given by any cumulant-based method from the covariances of the sample cumulants and the Jacobian matrix of the estimates with respect to the sample cumulants. For an MA process and using a symbolic algorithm, it is not difficult to obtain the performance as a function of MA coefficients and the statistics of the noise.

Figure 1 compares the information carried by different sets of sample cumulants as a function of the coefficient b_1 of an exponential MA(1) process without noise. The three curves show the asymptotic variance in the estimation of b_1. L2 corresponds to the lower bound using the sample autocorrelation ($c_2(0)$ and $c_1(0)$), L3 to the lower bound using sample third-order cumulants ($c_3(0,0)$, $c_3(0,1)$ and $c_3(1,1)$) and L23 to the lower bound using both sets. From this figure, it is clear that second order statistics, apart from being phase-blind, do not provide good estimates when the zeros are close to the unit circle.
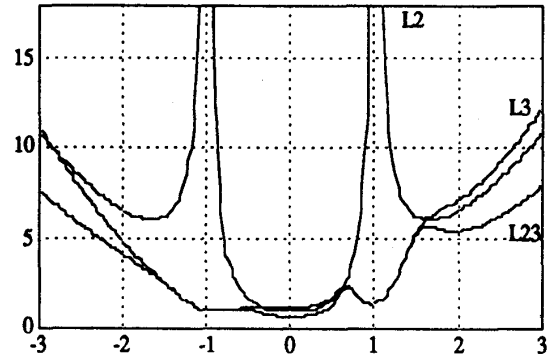

Figure 1. Asymptotic variance of the estimate as a function of the coefficient (b_1) of an MA(1) process. Lower bound using autocorrelation (L2), using third-order cumulants (L3), and using both (L23).

Figure 2 shows the performance, (asymptotic variance in the estimation of b_1), of three linear methods. GM is the method proposed by Giannakis and Mendel in [4] and GMT is the *Modification to Reformulated GM Algorithm* described in [5]. These two algorithms use second- and third-order cumulants whereas the WS method developed in [6] uses only third-order cumulants.
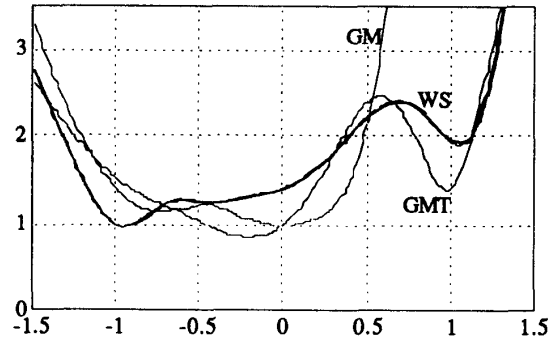

Figure 2. Asymptotic variance of the estimate as a function of the coefficient (b_1). MA(1) process. GM [4], GMT [5], and WS [6] methods.

The following table shows the agreement between the predicted performance by the analytic formulas and the simulation results for the WS method. For the Monte Carlo simulations the number of samples of each record was N=1000, and the value of the asymptotic standard deviation was computed from the estimates as $(N\, var(\hat{b}_1))^{1/2}$.

True Parameter	WS Analytical	WS Simulation
$b_1 = -0.80$	1.1604	1.1611
$b_1 = -1.25$	2.1834	2.1762
$b_1 = 1.25$	2.8146	2.8395

Figure 3 shows the performance of WS method as a function of the signal to noise ratio, for an MA(1) process with $b_1 = -1.25$. The lower bound of the asymptotic standard deviation is also included. The noise is white and Gaussian. SNR is computed here as the ratio between the power of the noise and the power of the output signal (MA process) and it ranges between 0 and 1.
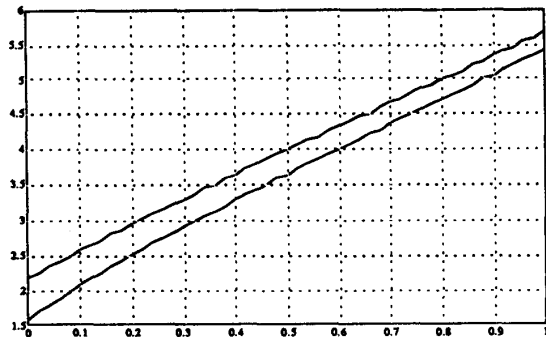

Figure 3. Asymptotic variance in the estimation of b_1 with the WS method [6] as a function of signal to noise ratio. MA(1) process with $b_1 = -1.25$. White Gaussian noise. The lower line corresponds to the lower bound (minimum variance of any method using the sample third-order cumulants $c_3(0,0)$, $c_3(0,1)$ and $c_3(1,1)$).
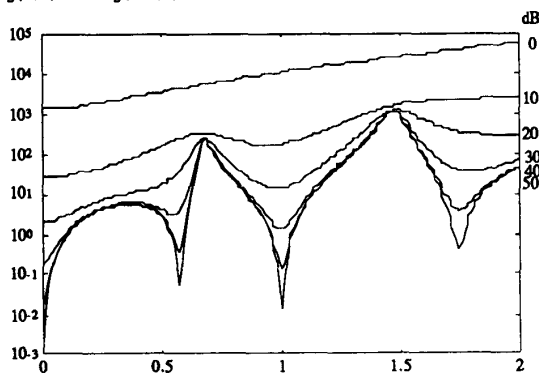

Figure 4. Number of sample required to estimate the coefficient b_1 with with an standard deviation equal to 0.01 as a function of the signal to noise ratio and b_1. BPSK signal transmitted through an MA(1) channel with additive white Gaussian noise at the output. Lower bound using sample fourth-order cumulants $c_4(0,0,0)$, $c_4(1,1,1)$ and $c_4(-1,-1,-1)$.

In the last example, a BPSK signal is used instead of the exponential processes considered in previous examples. A BPSK signal is transmitted through an MA(1) channel and we want to estimate b_1 from the sample fourth order cumulants $c_4(0,0,0)$, $c_4(1,1,1)$, and $c_4(-1,-1,-1)$ of the received signal. *Figure 4* illustrates the number of samples required to obtain an asymptotic standard deviation of 0.01 in the estimation of b_1 (lower bound). Observe that the number of samples ranges between hundreds and millions depending on the coefficient b_1 of the channel and the signal to noise ratio.

CONCLUSIONS

The expressions of the covariances of the sample cumulants of stationary linear processes in noise have been used to analyze the performance of MA system identification methods.

The lower bounds using different sets of sample cumulants have also been compared. From the results it is clear that higher-order statistics, apart from providing phase information, carry useful information to reduce the variance of the estimates.

The effect of noise has also been studied. We have found that for signal to noise ratios below 0-10 dB, the performance degrades very fast.

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

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

[6] José A. R. Fonollosa, J. Vidal, and A. Moreno, "FIR System Identification using a Linear Combination of Cumulants", in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. V, pp. 473-476, March 1992.

[7] P. McCullagh, "Tensor Methods in Statistics", Monographs on Statistics and Applied Probability, New York: Chapman and Hall, 1987.

[8] S. Wolfram, "Mathematica: A System for Doing Mathematics by Computer", Second Edition, Addison-Wesley, 1991.