August 1986 (revised)
September 1985

LIDS-P-1521

Dual Estimation of the Poles and Zeros of an ARMA(p,q) Process[1]

M. Isabel Ribeiro[2], José M. F. Moura[3]

Abstract: Identification of the p poles and q zeros of an autoregressive moving average process ARMA(p,q) is considered. The method described departs from approaches frequently reported in the litterature in two main respects. First, it substitutes the sample covariance lags by the sequence of estimated reflection coefficients. Second, it provides a direct estimation procedure for the MA component, which is not contingent upon prior identification of the AR structure. In distinction to current practice, both tasks are directly addressed, avoiding that errors in one contaminate the other. The algorithm explores the linear dependencies between corresponding coefficients of successively higher order linear filters fitted to the time series: linear predictors are used for the estimation of the MA component and linear innovation filters for the identification of the AR part. The overdimensioned system of linear equations derived from these dependencies provides statistical stability to the procedure. The paper establishes these dependencies and derives from them a recursive algorithm for ARMA identification. The recursiveness is on the number of (sample) reflection coefficients used. As it turns out, the MA procedure is asymptotic in nature, the rate of convergence being established in terms of the second power of the zeros of the process. Simulation examples show the behavior of the algorithm, illustrating how peaks and valleys of the power spectrum are resolved. The quality of the estimates is established in terms of the bias and mean square error, whose leading terms are shown to be of order T^{-1}, where T is the data lenght.

1 The work of the first author was partially supported by INIC (Portugal). The work of the second author was supported by the Army Research Office under contract DAAG-29-84-k-005.

2 Teaching Assistant, CAPS, Instituto Superior Técnico, Lisbon, Portugal.

3 Professor, Department of Electrical and Computer Engineering, Carnegie-Mellon University. The work was carried out while visiting the Massachusetts Institute of Technology and Laboratory for Information and Decision Systems, on leave from Instituto Superior Técnico, Lisbon, Portugal.

1. INTRODUCTION

In many problems an important question is that of fitting models to a series of measurements points. Tha available a priori information and the ultimate purpose of the model may change with the specifics of each application. Accordingly, there are alternative standard classes of models to fit the time series. The Autoregressive Moving-Average processes with p poles and q zeros, ARMA(p,q), plays a relevant role in the study of stationary time series with rational spectrum. The ARMA process is the output of a linear discrete time invariant system driven by an uncorrelated sequence of Gaussian random variables. The system exhibits a feedback or Autoregressive part AR(p) and a feedforward or Moving-Average part MA(q). In the absence of the moving-average component (q=0), the work of several authors shows that the estimate of the process parameters has well established methods, see for example appropriate comments in [1], [2], [3]. These methods explore the linear dependence between (p+1)-successive autocorrelation lags of the process (Yule-Walker equations). Taking into account the Toeplitz structure of the associated autocorrelation matrix, when a new lag is available, the Levinson algorithm updates the AR parameters via a numerically efficient recursion. The Burg modification [4] of the Levinson algotithm substitutes the knowledge of the autocorrelation lags by the estimation of the reflection coefficients.

When the MA component is present, the moving-average parameters are nonlinearly related to the autocorrelation lags. The methods reported in the literature either require nonlinear optimization techniques, or first estimate the AR parameters, then remove the (estimated) AR component, and finally obtain the MA parameters by factorization of the spectrum of the residual process. Whatever procedure used, experience has shown that the MA parameter estimates are as a rule of lower quality as compared to the AR estimates, e.g. [1], [2], [3].

By the Wold decomposition [5], [6], [7], a finite order ARMA(p,q) is equivalent to an MA(∞) or an AR(∞). The presence of zeros may be traded for the inclusion of higher order poles. This is particularly penalizing in narrowband applications, where an acceptable accuracy of the spectral notches may require a larger number of poles. Albeit contradicting the parsimony principle of statistics, this is a common procedure in practice. Noting that the Burg recursion provides at intermediate stage i the "best" (in the mean square sense) i^{th} order linear predictor approximation to the process, higher order AR models are successively fitted to the time series, till a "reasonable" approximate model is found.

The aim of the present paper is to discuss a method where the AR and MA component of a multivariable ARMA(p,q) process are estimated by a dualized type algorithm. This dualization is in the sense that the mechanics of the MA estimation algorithm parallel those of the AR part. Actually, the proposed equations do not use the autocorrelation lags, but the elements of a conveniently defined square root of the Toeplitz matrix of autocorrelation lags associated with the process. The algorithm accomplishes the following:

i) The MA coefficients are determined from the linear dependence exhibited by corresponding coefficients of successively higher order linear predictor filters.

ii) The AR coefficients are determined from the linear dependence exhibited by corresponding coefficients of successively higher order innovations filters.

The above statements need clarification. In point i) the coefficients of the alluded linear dependence are not the MA parameters. The important fact is that asymptotically they converge to the MA parameters (Corollary 1). The quality of the procedure is connected to the rate of this convergence (Theorem 3). In point ii) the coefficients of the linear dependence are in fact the AR parameters. From a different perspective, this explains why the MA determination may have lower quality than the AR estimation.

In order to establish the linear dependence mentioned in i) and ii) above, the ARMA model is given in Section 2 an internal description (state variable framework). The model fitting becomes a linear prediction problem of the Kalman-Bucy type [8], where the measurement noise is completely correlated with the driving noise sequence. For this problem, Section 2 adapts the asymptotic results of the associated Riccati equation obatined in [9], where the limiting behavior is in the sense that the measurement noise becomes totally correlated with the driving noise. The limiting results of [9] parametrize the coefficients of the successively higher order linear predictor and linear innovation filters in terms of the state variable matrices describing the process in terms of the solution of a (degenerate) Riccati equation. Section 3 shows how the coefficients of those filters are linearly related by the MA and AR parameters. Section 4 presents the details of the dual algorithm for AR and MA estimation. In section 5, some simulation results are presented together with a discussion of the performance of the algorithm. Finally Section 6 concludes the paper.

2. MODELLING CLASS

An s-dimensional, zero mean stationary sequence, $\{\ldots, y(n-1), y(n), y(n+1),\ldots\}$ is given, being assumed to be a multivariable ARMA process, output of the linear discrete time invariant (LDTI) system

$$y(n) + \sum_{i=1}^{p} A_i y(n-i) = \sum_{i=0}^{q} B_i e(n-i) . \qquad (1)$$

Define the z transforms

$$A(z) = \sum_{i=0}^{p} A_i z^{-i}, \; A_o=I \qquad\qquad B(z) = \sum_{i=0}^{q} B_i z^{-i}. \qquad (2)$$

The following are in force:

(H1) $\{e_n\}$ is an s-dimensional discrete white Gaussian noise sequence, with zero mean, and identity covariance matrix,

(H2) $p\geq q$, det $A_p\neq 0$, det $B_o\neq 0$, $B_q\neq 0$,

(H3) detA(z) is an asymptotically stable polynomial,

(H4) the polynomial matrices A(z) and B(z) are left coprime.

By (H2)

$$\deg(\det A(z)) = d = ps . \qquad (3)$$

We will denote by a_i, $i=0,\ldots,d$ the coefficients of the characteristic polynomial,

$$a(z) = \det A(z) = \sum_{i=0}^{d} a_i z^{-i}, \quad a_o = 1 , \qquad (4)$$

and by

$$N(z) = \sum_{i=0}^{u} N_i z^{-i}, \quad N_o = B_o \tag{5}$$

the numerator polynomial of the system transfer matrix, $A^{-1}(z)B(z)$ written as $\phi(z)=N(z)/a(z)$, where, by (H2) and (H4), $u=d-(p-q)$.

With (1), we associate the state-variable model,

$$x_{n+1} = F x_n + G e_n \qquad n \geq 0 \tag{6}$$

$$y_n = H x_n + B_o e_n \tag{7}$$

which, under (H1)-(H4) can be chosen so that,

(M1) - it has minimal dimension d, i.e. $x \in R^d$,

(M2) - F is an asymptotically stable matrix,

(M3) - (F,G) and (F,H) are completely controllable and completely observable pairs,

(M4) - F is nonsingular.

The dimensions of F, G and H follow from (M1). The initial state x_o is a Gaussian, zero mean vector, with covariance matrix P_o, solution of the discrete Lyapounov equation

$$P_o = F P_o F^T + G G^T. \tag{8}$$

The Kalman-Bucy filter associated with (6)-(7) leads to

$$\hat{x}_{n+1} = F \hat{x}_n + K_n v_n, \qquad \hat{x}_o = 0 \tag{9}$$

$$y_n = H \hat{x}_n + v_n \tag{10}$$

where $\hat{x}_n$ is the one-step ahead linear prediction of the state vector x at time n given the past observations and v_n is the innovation process. The gain matrix K_n is given by

$$K_n = (F P_n H^T + G B_o^T) D_n^{-1} , \tag{11}$$

where

$$D_n = H P_n H^T + B_o B_o^T , \tag{12}$$

is the power of the innovation sequence. The one step ahead prediction error covariance matrix P_n is given by the Riccati equation

$$P_{n+1} = F P_n F^T - K_n D_n K_n^T + G G^T, \quad n=0,1,2,... \tag{13}$$
$$P_o = P_o , \tag{8}$$

that corresponds to the prediction problem where the observation and driving noises are totally correlated. The behavior of this equation can be established by a straightforward modification of the results of [9], which considers the prediction problem when there is no noise in the observation. In particular, we obtain the following three results that will be needed later on.

Lemma 1: $P = 0$ is a fixed point of the Riccati equation (13).

#

Lemma 2: $\lim P_n = 0$ as n goes to infinity.

Proof of Lemma 2: Lemma 1 is obtained by direct verification. By the controllability and observability properties of (6)-(7), (13) has a unique nonnegative definite solution. Then Lemma 2 follows from Lemma 1.

#

Lemma 3:

$$P_n (F^{-T})^j H^T = 0 \qquad n \geq j, \; 1 \leq j \leq p-q \tag{14}$$

$$K_n^T (F^{-T})^{j+1} H^T = 0 \qquad 1 \leq j \leq p-q+1, \; n \geq j \tag{15}$$

Proof of Lemma 3: the Markov parameters of the system (1)-(2) satisfy

$$H F^{-1} G = B_o - N_d / a_d \tag{16a}$$
$$H F^{-j} G = \underline{0} \qquad 2 \leq j \leq p-q \tag{16b}$$

where, $N_k=0$ for $k>u$. The proof of this Lemma then follows by induction using (13) and (11) together with (16a)-(16b); it is a straighforward modification of the proof for the scalar case contained in [10]

#

Define

$$Y_r^k = \text{vect } (y_i)_{\, r\leq i\leq k} = [\; y_r^T \ldots y_k^T \;]^T$$

and similarly V_r^k. Considering the normalized variance representation of the innovation process, $\tilde{v}_i = D_i^{-1/2} v_i$ (*), the input/output relations (9)-(10) may be writen in a matrix format as

$$Y_0^N = \tilde{W}_N \, \tilde{V}_0^N \,, \tag{17}$$

where $\tilde{W}_N$ is the (N+1)x(N+1) block lower triangular matrix

$$\tilde{W}_N = \begin{bmatrix} D_o^{1/2} & & & & \\ HK_oD_o^{1/2} & D_1^{1/2} & & \underline{0} & \\ HFK_1D_o^{1/2} & HK_1D_1^{1/2} & D_2^{1/2} & & \\ \vdots & \vdots & & \ddots & \\ HF^{N-1}K_oD_o^{1/2} & HF^{N-2}K_1D_1^{1/2} & & \cdots & D_N^{1/2} \end{bmatrix}, \tag{18}$$

each block being a square matrix of order s. The block line i of $\tilde{W}_N$ (line and column blocks are numbered starting from zero) is the i^{th} order normalized innovation representation of the process. Conversely, the block entries on line i of the inverse matrix $\tilde{W}_N^{-1}$ are the i^{th} order normalized prediction error filter coefficients associated with the process. An expression of $\tilde{W}_N^{-1}$ matrix entries

---

(*) The square root $A^{1/2}$ of a matrix $A>0$ is defined as the lower triangular matrix such that $A = A^{1/2} \, (A^T)^{1/2}$.

$$\tilde{W}_N^{-1} = \begin{bmatrix} D_o^{-1/2} & & & & \\ -D_1^{-1/2}HK_o & D_1^{-1/2} & & & \underline{0} \\ -D_2^{-1/2}HF_1^cK_o & \cdots & D_2^{-1/2} & & \\ -D_3^{-1/2}HF_2^cF_1^cK_o & & & \ddots & \\ \vdots & & & & \\ -D_N^{-1/2}HF_{N-1}^c\cdots F_1^cK_o & -D_N^{-1/2}HF_{N-1}^c\cdots F_2^cK_1 & & \cdots & D_N^{-1/2} \end{bmatrix} \quad (19)$$

where

$$F_i^c = F - K_i H \quad (20)$$

is the closed loop filter matrix, were presented in [9]. From (17)

$$R_N = \tilde{W}_N \ \tilde{W}_N^T = E\ Y_o^N\ (Y_o^N)^T$$

where R_N is the matrix of the autocovariance lags of the process $\{y_n\}$, truncated at lag N. In the next section, linear relations on the entries of consecutive lines of $\tilde{W}_N$ and $\tilde{W}_N^{-1}$ will be established.

3. LINEAR DEPENDENCE OF THE COEFFICIENTS OF SUCCESSIVE ORDER INNOVATION AND PREDICTOR LINEAR FILTERS

As we recall, block line i=0,1,...,N of $\tilde{W}_N^{-1}$ represents the i^{th} order predictor filter representation of the ARMA(p,q) process. We want to study the linear dependence of each collection of u+1 consecutive nonzero elements in every block column of $\tilde{W}_N^{-1}$. A procedure that obtains asymptotically the coefficients of the transfer matrix numerator polynomial (see equation (5)) from the coefficients of increasing order prediction error filters is derived, together with the study of its rate of convergence. A dual, nonasymptotic result is provided, relating each collection of d+1 consecutive elements in every block column of $\tilde{W}_N$ and the coefficients of the system characteristc polynomial (see equation (4)).

Let A_N be the lower triangular band diagonal matrix of order (N+1)s, with block entries

$$(A_N)_{ij} \underset{i,j=0,\ldots,N}{=} \begin{cases} a_{i-j} I_s & 0 \leq i-j \leq d \\ \underline{0} & \text{otherwise} \end{cases} \tag{21}$$

where I_s is the identity matrix of order s. Premultiply the normalized innovation system (17) by A_N to get

$$A_N \, Y_o^N = (A_N \, \tilde{W}_N) \, \tilde{V}_o^N \tag{22}$$

and let

$$\beta_N \, \tilde{W}_N^{-1} = A_N \tag{23}$$

Theorem 1: β_N has the following structure

$$\beta_N = \begin{bmatrix} \Omega_0(0) & & & & & & & & \\ \Omega_1(1) & \Omega_0(1) & & & & & & \underline{0} & \\ \vdots & \vdots & \ddots & & & & & & \\ \Omega_u(u) & \Omega_{u-1}(u) & \cdots & \Omega_0(u) & & & & & \\ \vdots & & \ddots & & \ddots & & & & \\ \Omega_{d-1}(d-1) & \cdots & \Omega_u(d-1) & \Omega_{u-1}(d-1) & \cdots & \Omega_0(d-1) & & & \\ 0 & 0 \; \cdots & 0 & \Omega_u(d) & \cdots & \Omega_1(d) & \Omega_0(d) & & \\ & & & & \ddots & & \ddots & & \\ & \underline{0} & & & & & & & \\ & & & & & \Omega_u(N) & \cdots & \Omega_1(N) & \Omega_0(N) \end{bmatrix}$$

where

$$\Omega_m(i) = \{ N_m B_o^T - \sum_{r=m+d-u+1}^{d} a_r H F^{m-r} P_{i-m} H^T \} \, D_{i-m}^{-1/2} \qquad 0 \leq m \leq u \leq d, \quad i \geq d \tag{24}$$

is a square matrix of order s.

Proof of Theorem 1: From the matrix identity (23), the structure of A_N (21) and the Cayley-Hamilton theorem satisfied by any d+1 consecutive block entries on each column of $\tilde{W}_N$ below its main diagonal, β_N is a lower triangular band diagonal matrix, with block entries

$$(\beta_N)_{ij} \underset{i,j=0,1,\ldots,N}{=} \begin{cases} \underline{0} & j > i \\ \Omega_{i-j}(i) & 0 \leq i-j \leq d \\ \underline{0} & i-j > d \end{cases} \tag{25}$$

where,

$$\Omega_m(i) = \sum_{r=0}^{m} a_r H F^{m-1-r} K_{i-m} D_{i-m}^{-1/2} + (a_m/a_d) \; N_d B_o D_{i-m}^{-1/2} \qquad 0 \le m \le d, \; m \le i \qquad (26)$$

is obtained by direct calculation on (23). The proof then follows replacing (11) in (26) and using the results provided by Lemma 3 and the properties of the system Markov parameters (see (16a)-(16b)). All the details of the proof, concerning the scalar case are presented in [10].

♦

Corollary 1: $\lim \Omega_m(i) = N_m$ as i goes to infinity, $0 \le m \le u$.

Proof: The corollary follows from the asymptotic behavior of the solution of the Riccati equation, as given in Lemma 2, and the fact that D_i converges to $B_o B_o^T$ (see (12)).

♦

Represent by

$$\tilde{a}_j^i = (\tilde{W}_N^{-1})_{ij} \qquad i,j=0,\ldots,N \qquad (27)$$

the block entries of the normalized matrix $\tilde{W}_N^{-1}$. The structure of the matrix A_N and B_N in (23) lead to the following

Corollary 2: The elements $\Omega_m(i)$ satisfy the following linear recursions

$$\Omega_0(i) \, \tilde{a}_j^i + \Omega_1(i) \, \tilde{a}_j^{i-1} + \ldots + \Omega_u(i) \, \tilde{a}_j^{i-u} = 0 \quad d \le i \le N, \; 0 \le j \le i-d-1 \qquad (28a)$$

$$\Omega_0(i) \, \tilde{a}_j^i + \Omega_1(i) \, \tilde{a}_j^{i-1} + \ldots + \Omega_{i-j}(i) \, \tilde{a}_j^j = a_{i-j} I_s \quad d \le i \le N, \; i-d \le j \le i \qquad (28b)$$

where the leading coefficient $\Omega_0(i)=D_i^{1/2}$.

♦

Corollary 2 states that the set of u+1 block elements on each column of the normalized matrix $\tilde{W}_i^{-1}$ ($d \le i \le N$) satisfies the same linear relations defined by the $\Omega_m(i)$ coefficients; these relations are homogeneous for the first i-d block columns of $\tilde{W}_i^{-1}$ (see (28a)), and their left hand side is defined by the system characteristic polynomial coefficients for the last d+1 block columns of $\tilde{W}_i^{-1}$. For the reflection coefficient matrices, (the block elements of the first column of $\tilde{W}_i^{-1}$), this result is presented in [9]. Corollary 2 generalizes it to all normalized prediction error filter coefficients, giving, together with (24), a non-asymptotic relation between them.

Corollary 1 gives the asymptotic behavior of the nonzero $\Omega_m(i)$ as i goes to infinity. We study now the rate of this convergence. Let h_j, the j^{th} line of H be such that

$$T = [\ (h_j F^{-1})^T \mid (h_j F^{-2})^T \mid \ldots \mid (h_j F^{-d})^T\ ]^T \quad 1 \leq j \leq s \tag{29}$$

is a nonsingular matrix.

<u>Lemma 4</u>: Under the coordinate transformation T (29), the covariance matrix P_n becomes

$$\tilde{P}_n = \left[\begin{array}{c|c} \underline{0} & \underline{0} \\ \hline \underline{0} & Q_n \end{array}\right], \quad n \geq p-q \tag{30}$$

where the square matrix of order $u=d-(p-q)$, Q_n, satisfies the Riccati equation

$$Q_{n+1} = \tilde{F}_{22} Q_n \tilde{F}_{22}^T + \tilde{G}_{12} \tilde{G}_{12}^T - \tilde{K}_n \tilde{D}_n \tilde{K}_n^T, \quad n \geq p-q \tag{31}$$

and

$$\tilde{F} = TFT^{-1} = \left[\begin{array}{cccc|c} -a_1 & -a_2 & \cdots & & -a_d \\ \hline 1 & & & & \\ & \underline{0} & \ddots & & \underline{0} \\ & & & 1 & \end{array}\right] = \left[\begin{array}{c|c} \tilde{F}_{11} & \tilde{F}_{12} \\ \hline \tilde{F}_{21} & \tilde{F}_{22} \\ & (u \times u) \end{array}\right] \tag{32}$$

$$\tilde{G} = T\,G = [(h_j F^{-1} G)^T \ \ldots \ (h_j F^{-d} G)^T]^T = [\ \tilde{G}_{11}^T \mid \tilde{G}_{12}^T\ ]^T \tag{33}$$

$$\tilde{H} = H\,T^{-1} = [\ \tilde{H}_{11} \mid \tilde{H}_{12}\ ] \tag{34a}$$

$$\tilde{h}_j = [\ -a_1 \ -a_2 \ \ldots \ -a_d\ ] \tag{34b}$$

$$\tilde{K}_n = (\ \tilde{F}_{22}\, Q_n\, \tilde{H}_{12}^T + \tilde{G}_{12}\, B_o^T\ )\ \tilde{D}_n^{-1} \tag{35}$$

$$\tilde{D}_n = \tilde{H}_{12}\, Q_n\, \tilde{H}_{12}^T + B_o B_o^T \tag{36}$$

where $\tilde{h}_j$ is the j^{th} line of the matrix $\tilde{H}$.

Proof of Lemma 4: Under T, $\tilde{P}_n = TP_nT^T$. For p=q, equations (31)-(36) follow by direct evaluation of the Riccati equation (13) under T according to (32)-(36). For p>q, and by Lemma 3,

$$TP_n = [\ \underline{0}\ \mid\ (Q_n^1)^T\ ]^T \quad n \geq p-q \tag{37}$$

and

$$P_nT^T = [\ \underline{0}\ \mid\ Q_n^2\ ] \quad n \geq p-q \tag{38}$$

where Q_n^1 is uxd and Q_n^2 is dxu. Multiplying (37) by T^T on the right, and (38) by T on the left, equality of results leads to (30). Equation (31) then follows by the same argument used for p=q.

#

Having in mind the special structure of the covariance equation given by Lemma 4, the expression of the $\Omega_m(i)$ given by theorem 1 is rearranged in the following theorem.

Theorem 2: $$\Omega_m(i) = \{ N_mB_o^T - \sum_{r=m+d-u+1}^{d} a_r \Psi(r-m)_{12}\, Q_{i-m}\, H_{12}^T \}\, D_{i-m}^{-1/2} \tag{39}$$

$$i-m \geq d-u,\quad 0 \leq m < u \leq d$$

where $\Psi(r-m)_{12}$ is defined in (40).

Proof: Starting with equation (24), we note that HP_nH^T is coordinate invariant. By Lemma 4, it follows that

$$HP_nH^T = \tilde{H}_{12}Q_n\tilde{H}_{12}^T \qquad n \geq p-q$$

from which $\tilde{D}_n = D_n$ for $n \geq p-q$. Under the transformation T,

$$HF^{-(r-m)}P_{i-m}H^T = \tilde{H}\tilde{F}^{-(r-m)}\tilde{P}_{i-m}\tilde{H}^T \quad m+d-u+1 \leq r \leq d.$$

Defining

$$\tilde{h}\tilde{F}^{-(r-m)} = \Psi(r-m) = [\ \Psi(r-m)_{11} \mid \Psi(r-m)_{12}]\ , \tag{40}$$

and from (30) yields

$$\tilde{H}\tilde{F}^{-(r-m)}\tilde{P}_{i-m}\tilde{H}^T = \Psi(r-m)_{12}\, Q_{i-m}\tilde{H}_{12}^T$$

thus concluding the proof. For the scalar case, the line matrix $\Psi(r-m)$ has null entries except for its r-m position which is one [10]. On the multivariable case, the same structure occurs on line j of $\Psi(r-m)$, under the coordinate transformation T (29).

#

Theorem 2 shows that the rate of convergence of the $\Omega_m(i)$ to the N_m parameters is determined by the convergence to zero of Q_n. This depends on the rate by which

$$G_n = \Lambda^{-[n-(p-q)]} W \Lambda^{-[n-(p-q)]} \tag{41}$$

goes to zero, where $\Lambda = \mathrm{diag}(\lambda_i)$, λ_i being the nonstable eigenvalues of the Hamiltonian matrix H associated with (41),

$$H = \left[\begin{array}{c|c} (R^T)^{-1} & (R^T)^{-1} \tilde{H}_{12}^T (B_o B_o^T)^{-1} \tilde{H}_{12} \\ \hline 0 & R \end{array}\right] \tag{42}$$

$$R = \tilde{F}_{22} - \tilde{G}_{12} B_o^{-1} \tilde{H}_{12} \tag{43}$$

and W is a constant matrix [11],[12],[13].

#

Finally we focus on the asymptotics of the $\Omega_m(i)$.

<u>Theorem 3</u>: $\| \Omega_m(i) - N_m \| \sim O\left(\lambda_j^{2(i-(p-q))} \right)$ (44)

where O(.) means it goes to zero at least as the slowest argument and λ_j are the zeros of the original multivariable ARMA process.

<u>Proof</u>: The rate of convergence of the $\Omega_m(i)$'s depends on the rate by which $G_n \longrightarrow 0$, see (41); this convergence is a function of the nonstable eigenvalues of the Hamiltonian matrix H. The task is then to evaluate these eigenvalues. From (42), it follows

$$\Delta(\lambda) = \det(\lambda I - H) = \phi(\lambda)\, \phi(\lambda^{-1})$$

where $\phi(\lambda) = \det(\lambda I - R)$ and R is defined in (43) for the reduced order system $(\tilde{F}_{22}, \tilde{G}_{12}, \tilde{H}_{12}, B_o)$. It is shown in [9] that, in the case of no observation noise, the asymptotic closed loop poles are the zeros of the original open loop system, this being actually a dual result for singular control systems, see [12]. In the context of the fully correlated observation and input noise model being studied here, this says that

$$\phi(\lambda) = (\det B_o)^{-1}\det(\lambda I_u-\tilde{F}_{22})\ \det[\tilde{H}_{12}(\lambda I_u - \tilde{F}_{22})^{-1}\tilde{G}_{12} + B_o] \ . \tag{45}$$

It remains to show that the zeros of the reduced system coincide with the zeros of the original system (F,G,H,B_o). For the concept of the zeros of a multivaribale system see, e.g., [14]. For $p=q$, the triplet $(\tilde{F}_{22},\tilde{G}_{12},\tilde{H}_{12})$ coincides with $(\tilde{F},\tilde{G},\tilde{H})$ and the result directly holds from (45). For $p>q$, note that

$$\det[B_o+\tilde{H}(\lambda I_d-\tilde{F})^{-1}\tilde{G}]= \det B_o\ \det[\lambda I_d-(\tilde{F}-\tilde{G}B_o^{-1}\tilde{H})]/\det(\lambda I_d-\tilde{F}) \tag{46}$$

From (32)-(29) and the Markov parameters associated with (6)-(7), see (16a)-(16b), it follows that

$$\tilde{F}_{11}-\tilde{G}_{11}B_o^{-1}\tilde{H}_{11} = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \quad (d-u)x(d-u) \tag{47}$$

$$\tilde{F}_{12}-\tilde{G}_{11}B_o^{-1}\tilde{H}_{12} = \underline{0}$$

and thus

$$\det[\lambda I_d-(\tilde{F}-\tilde{G}B_o^{-1}\tilde{H})] = \lambda^{d-u}\ \det[\lambda I_u-(\tilde{F}_{22}-\tilde{G}_{12}B_o^{-1}\tilde{H}_{12})] \ . \tag{48}$$

By (45), equation (48) can be rewritten as

$$\det[\lambda I_d-(\tilde{F}-\tilde{G}B_o^{-1}\tilde{H})] = \lambda^{d-u}(\det B_o)^{-1}\det[B_o+\tilde{H}_{12}(\lambda I_u-\tilde{F}_{22})^{-1}\tilde{G}_{12}]\ \det(\lambda I_u-\tilde{F}_{22}) \ . \tag{49}$$

Replacing (49) in (46) yields

$$\det[B_o+\tilde{H}(\lambda I_d-\tilde{F})^{-1}\tilde{G}] = \lambda^{d}\ \det[B_o+\tilde{H}_{12}(\lambda I_u-\tilde{F}_{22})^{-1}\tilde{G}_{12}\ ]/\det(\lambda I_d-\tilde{F})$$

thus concluding the proof.

∎

Theorem 3 gives the very interesting result that the rate of convergence of the $\Omega_m(i)$ to the MA coefficients are determined by the zeros of the process itself. This means that when the zeros are near the unit circle (narrowband process) the convergence is slowed down; as the zeros get away from the unit circle (wideband), the convergence is fastened. Note, however, that the convergence goes with the second power of the zeros, which preludes a fast convergence in many practical situations.

The AR counterpart of corollary 2 can be obtained from the normalized innovation system representation (17) and the underlying structure of the matrices β_N as in theorem 1, A_N as in (22), and $\tilde{W}_N$ as in (18). From the matrix identity (23), obtain:

Corollary 3: The block coefficients of successively higher order linear predictors are related by:

$$\tilde{W}_j^i + a_1 \tilde{W}_j^{i-1} + a_2 \tilde{W}_j^{i-2} + \ldots + a_d \tilde{W}_j^{i-d} = 0, \quad d \leq i \leq N \quad 0 \leq j \leq i-u-1 \quad (50a)$$

$$\tilde{W}_j^i + a_1 \tilde{W}_j^{i-1} + a_2 \tilde{W}_j^{i-2} + \ldots + a_d \tilde{W}_j^{i-d} = \Omega_{i-j}(i) \quad d \leq i \leq N, \; i-u \leq j \leq i \quad (50b)$$

where $\tilde{W}_j^i = (\tilde{W}_N)_{ij}$ $(i,j=0,\ldots,N)$ is the block entry ij of the block triangular matrix W_N, which is zero for i>j.

#

Equation (50a) above means that the increasing order normalized innovation filter coefficients contained in the first i-u columns of $\tilde{W}_i$ satisfy a set of linear relations defined by the system characteristic polynomial coefficients.

Both (28a) and (50a) are used in section 4 to define an "overdetermined Yule-Walker" type algorithm which independently estimates the AR and MA components of a stationary multivariable ARMA process with known orders. Simulation results of such an algorithm that applies the Burg technique [4] to estimate the reflection coefficients directly from data are presented, for the scalar case, in Section 5. Some examples are also presen ted in [10] and will be extensively discussed in [15].

4. ARMA ESTIMATION ALGORITHM

In this section the estimation algorithm for both components of a multivariable ARMA(p,q) process is presented.

The $\{N_i\}$, $(i=0,\ldots,u)$ coefficients of the transfer matrix numerator polynomial are asymptotically determined from the linear dependence exhibited by the corresponding coefficients of successively higher order linear predictor filters. The $\{a_i\}$ coefficients of the system characteristic polynomial are obtained from the linear relations satisfied by the corresponding coefficients of the higher order linear innovation filters.

The MA and AR components are thus estimated by decoupled dualized procedures. Further, the $\{a_i\}$ coefficients may be obtained by a recursive scheme whose structure resembles that of an adaptive algorithm. The moving average counterpart result is asymptotic in nature.

4.1. AUTOREGRESSIVE ESTIMATION

The coefficients of the system characteristic polynomial $\{a_i\}$ are given as the solution of a set of linear (algebraic) equations, established from the linear, matricial relations presented in Corollary 3.

Let

$$\underline{A} = [\ a_1 I_s \ \ a_2 I_s \ \ \ldots \ a_d I_s \ ]^T \qquad (51)$$

be a block matrix defined by the coefficients a_i, $i=1,\ldots,d$, and rewrite (50a) as

$$[\ \tilde{W}_j^{i-1} \ \ \tilde{W}_j^{i-2} \ \ \ldots \ \ \tilde{W}_j^{i-d}\ ]\ \underline{A} = -\ \tilde{W}_j^{i} \qquad d\leq i\leq N \quad 0\leq j\leq i-u-1 \qquad (52)$$

where it is assumed that (52) is not verified when j belongs to an empty set. For a fixed pair (i,j), the equation (52) represents a system of s^2 linear, scalar relations defined by the corresponding entries of the block matrices $\tilde{W}_j^{i}$, ..., $\tilde{W}_j^{i-d}$; for each entry of those matrices, (52) is rewritten as

$$[(\tilde{W}_j^{i-1})_{km}\ (\tilde{W}_j^{i-2})_{km}\ \ldots\ (\tilde{W}_j^{i-d})_{km}]\ \underline{a} = -\ (\tilde{W}_j^{i})_{km} \qquad d\leq i\leq N \quad 0\leq j\leq i-u-1 \quad 1\leq k,m\leq s \qquad (53)$$

where $(\tilde{W}_j^{i})_{km}$ is the element on line k (k=1,...,s) and column m (m=1,...,s) of $\tilde{W}_j^{i}$, and

$$\underline{a} = [\, a_1 \;\; a_2 \; \ldots \;\; a_d \,]^T. \tag{54}$$

The number of equations in (53) is

$$s^2 \cdot \sum_{i=d}^{N} (i - u) = s^2 \, (N+1-d) \, (N+d-2u)/2 \tag{55}$$

which is much larger than the number d of unknown transfer matrix denominator polynomial coefficients.

If the coefficients of the normalized innovations filters were exactly known, the $\{a_i\}$ could be determined by solving jointly any d of the preceeding equations (53). However, in the presence of a finite sample of the observation process, the exact value of $\tilde{W}_N$ is not available. Accordingly, on what follows it is assumed that $\tilde{W}_N$ is replaced by a suitable estimate. It is then of interest to obtain a recursive scheme for the solution of (53) as N is increased.

<u>Definition</u>: Let $\hat{a}(m)$ be the least-squares solution of the system of all the linear equations (53) established for $d \leq i \leq m$.

A recursive solution of the oversized system of equations (53) is provided by the following theorem.

<u>Theorem 4</u>: The autoregressive coefficients are recursively estimated by

$$\hat{a}(N) = \hat{a}(N-1) + M(N)^{-1} \, C_N^T \, [\, f_N - C_N \, \hat{a}(N-1) \,] \tag{56}$$

where $\hat{a}(N)$ and $\hat{a}(N-1)$ were defined above,

$$C_N = \begin{bmatrix} \tilde{\underline{w}}_0^{N-1} & \cdots & \tilde{\underline{w}}_0^{N-d} \\ \vdots & & \vdots \\ \tilde{\underline{w}}_{N-u}^{N-1} & \cdots & \tilde{\underline{w}}_{N-d}^{N-d} \\ \vdots & \iddots & \underline{0} \\ \tilde{\underline{w}}_{N-u-1}^{N-1} \;\cdots & \tilde{\underline{w}}_{N-u-1}^{N-u-1} & \end{bmatrix} \tag{57}$$

$$\tilde{\underline{w}}_j^i = [\ (\tilde{W}_j^i)^T_{.1} \mid (\tilde{W}_j^i)^T_{.2} \mid \dots \mid (\tilde{W}_j^i)^T_{.s}]^T, \tag{58}$$

where $(\tilde{W}_j^i)_{.k}$ is the column k of the square matrix of order s, $\tilde{W}_j^i$, and

$$f_N = -\ [\ (\tilde{\underline{w}}_0^N)^T \ \ (\tilde{\underline{w}}_1^N)^T \ \ \dots\dots \ (\tilde{\underline{w}}_{N-u-1}^N)^T]^T. \tag{59}$$

The dxd matrix M(N) is defined by

$$M(N) = \mathcal{C}^T(N)\ \mathcal{C}(N) \tag{60}$$

and the block matrices $\mathcal{C}(N)$ and $\mathcal{f}(N)$ equal

$$\mathcal{C}(N) = [\ C_d^T\ C_{d+1}^T \ \ \dots \ \ C_{N-1}^T\ C_N^T\ ]^T \tag{61}$$

$$\mathcal{f}(N) = [\ f_d^T\ f_{d+1}^T \ \ \dots \ \ f_{N-1}^T\ f_N^T\ ]^T \tag{62}$$

the C_j and f_j being defined similary to (57), (59).

<u>Remark</u>: In the above theorem, the recursion stands for the normalized innovation filter order.

<u>Proof of Theorem 4</u>: Write the set of linear equations (53) in matrix format as

$$\mathcal{C}(N)\ \hat{a}(N) = \mathcal{f}(N) \tag{63}$$

Partitioning,

$$\mathcal{C}(N) = [\mathcal{C}(N-1)^T \mid \ C_N^T\ ]^T \tag{64}$$

$$\mathcal{f}(N) = [\mathcal{f}(N-1)^T \mid \ f_N\ ]^T \tag{65}$$

where $\mathcal{C}(N-1)$, $\mathcal{f}(N-1)$ accounts for the linear relations (53) using the innovation filters up to order N-1, and C_N, f_N (57)-(59) stands for the contribution of the N-th order innovation filter. Hence,

$$\mathcal{C}(N-1)\ \hat{a}(N-1) = \mathcal{f}(N-1). \tag{66}$$

Using the definition (60), the least-squares solution of (63) is

$$\hat{a}(N) = M(N)^{-1} \mathcal{C}^T(N)\, \mathcal{f}(N).$$

From the partition (64), it follows further

$$\hat{a}(N) = M(N)^{-1}[\mathcal{C}^T(N-1)\, \mathcal{f}(N-1) + C_N^T f_N] \;. \tag{67}$$

From (66),

$$\mathcal{C}^T(N-1)\, \mathcal{f}(N-1) = M(N-1)\, \hat{a}(N-1) \tag{68}$$

which substituted in (67) leads to

$$\hat{a}(N) = M(N)^{-1}[M(N-1)\, \hat{a}(N-1) + C_N^T f_N]. \tag{69}$$

Realizing that

$$M(N-1) = M(N) - C_N^T C_N$$

the autoregressive estimation "at time instant N" is

$$\hat{a}(N) = \hat{a}(N-1) + M(N)^{-1}[\, C_N^T f_N - C_N^T C_N\, \hat{a}(N-1)] \tag{70}$$

directly leading to (56).

The initial condition for the above recursion can be established for an arbitrary value of N,(N>d), using a nonrecursive scheme based on (53).

4.2. MOVING-AVERAGE ESTIMATION

The proposed moving-average estimation algorithm paralells the AR scheme presented in section 4.1.. There are however some differences that will be pointed out in this section.

The moving-average component $\{N_m\}$ is asymptotically evaluated from the $\{\Omega_m(i)\}$ coefficients. These are obtained as the solution of a set of linear equations established from (28a). Thus, the duality of both estimation schemes becomes apparent.

Let

$$\Omega(i) = [\ \Omega_1(i)\ \Omega_2(i)\ \dots\ \Omega_u(i)\ ] \tag{71}$$

which, together with $\Omega_0(i)$, converges the MA coefficients $\{N_m\}$ (see Corollary 1). Represent by

$$\Omega(i)\ [(\tilde{a}_j^{i-1})^T\ (\tilde{a}_j^{i-2})^T\ \dots\ (\tilde{a}_j^{i-u})^T]^T = -D_i^{1/2}\ \tilde{a}_j^{i} \qquad d \le i \le N,\ 0 \le j \le i-d-1 \tag{72}$$

the set of matricial equations (28a), satisfied by the time-varying coefficients $\{\Omega_m(i)\}$ $(0<m\le u)$, the leading coefficient $\Omega_0(i)$ being

$$\Omega_0(i) = D_i^{1/2}. \tag{73}$$

For $j=0$, equation (72) states that the normalized reflection matrix coefficient sequence $\tilde{a}_0^k$ satisfies, asymptotically, a difference equation determined by the numerator polynomial of the system transfer function. First presented in [9], this result is generalized by (72) to the first N-d block columns of the matrix $\tilde{W}_N^{-1}$.

The time-varying behavior of the $\Omega_m(i)$ coefficients prevents its estimation using linear relations established for different values of i. For $i \ge u+d$, one can use at least u from the i-d linear relations (72) to evaluate the $\Omega_m(i)$ coefficients. However, if the order N of the matrix $\tilde{W}_N^{-1}$ is large so that $\Omega_m(N)$ are sufficiently close to the corresponding N_m (convergence has been attained), a recursive scheme to update the MA parameter estimates may be derived. This parallels the recursive AR procedure presented in Theorem 4. Assuming that the convergence is attained for the prediction error filter of order N^*, the moving average component can be recursively estimated by

$$\hat{\underline{N}}(k) = \hat{\underline{N}}(k-1) + [\ \tilde{f}_k\ \tilde{C}_k^T - \hat{\underline{N}}(k-1)\ \tilde{C}_k\ \tilde{C}_k^T\ ]\ \tilde{M}(k)^{-1} \tag{74}$$

where $\hat{\underline{N}}(k)$ stands for the estimation of the block matrix

$$\underline{N} = [\, N_1 \;\; N_2 \;\; \dots \;\; N_u \,]$$

using all the linear relations (72) for $N^{*}\leq i \leq k$, and

$$\tilde{C}_k = \begin{bmatrix} \tilde{a}_0^{k-1} & \tilde{a}_1^{k-1} & \dots & \tilde{a}_{k-d-1}^{k-1} \\ \tilde{a}_0^{k-2} & \tilde{a}_1^{k-2} & \dots & \tilde{a}_{k-d-1}^{k-2} \\ \vdots & \vdots & & \vdots \\ \tilde{a}_0^{k-u} & \tilde{a}_1^{k-u} & \dots & \tilde{a}_{k-d-1}^{k-u} \end{bmatrix} \tag{75}$$

$$\tilde{f}_k = - D_k^{1/2} \, [\, \tilde{a}_0^k \;\; \tilde{a}_1^k \;\; \dots \;\; \tilde{a}_{k-d-1}^k \,].$$

The matrix $\tilde{M}(k)$ is defined in a similar way as in (60) using $\tilde{C}_k$.

5. SIMULATION RESULTS

In this section we present some simulation results concerning the dual estimation algorithm discussed in the previous section, for the particular case s=1, i.e. for a scalar process $[y_n]$.

The AR component estimation uses the recursive scheme (70), while the MA parameters are obtained as the least squares solution of the linear system of equations (72). Both component estimated values are based on a finite sample of the process with lenght T.

In the presence of a finite sample of the observation process, the exact values of $\tilde{W}_N$ and $\tilde{W}_N^{-1}$ are not available. Accordingly, the algorithm requires suitable estimates of both matrices. We use the Burg technique [4] to estimate the reflection coefficients directly from data, and the Levinson algorithm to recursively obtain the increasing order one-step ahead prediction error filter coefficients. This procedure obtains the estimated value of the matrix $\tilde{W}_N^{-1}$. A recursive inversion of this lower triangular matrix leads to the estimation of the normalized innovation filter coefficients, i.e. to an estimated value of $\tilde{W}_N$, for increasing values of N.

It is now evident that the use of an oversized system of equations to obtain the AR and MA component estimates (see (70) and (72)) has an important statistical relevance, compensating the errors in the estimation of the reflection coefficients and consequently in the overall elements of $\tilde{W}_N^{-1}$ and $\tilde{W}_N$.

An alternative procedure based on a sample covariance estimator could have been used to estimate R_N, followed by a Cholesky factorization and a recursive inversion of the factors.

#

All the simulation examples presented in this section are obtained for an ARMA(6,4) scalar process, with pole-zero location displayed in Table 1.

Poles		Zeros	
Ampl.	Angle	Ampl.	Angle
0.95	$\pm 90^\circ$	0.85	$\pm 67.5^\circ$
0.85	$\pm 45^\circ$	0.85	$\pm 112.5^\circ$
0.85	$\pm 135^\circ$		

Table 1

The orders p=6 and q=4 are assumed known, and b_o, the leading coefficient of the transfer function numerator polynomial is equal to one. The notation used for the real and estimated pole-zero pattern is the following:

x - real pole
- estimated pole
o - real zero
◇ - estimated zero

Figure 1 shows, for T=1000 data points, the evolution of the estimated pole-zero pattern and the estimated spectrum as the order of the filter N increases. From figure 1, one sees that, as the order N increases, the algorithm leads to better zero estimates, and consequently the deep valeys of the spectrum are better resolved. As we increase the sample size T, we obtain increased performance. This is shown is figure 2, for T=500, 1000 and 5000 and N=15.

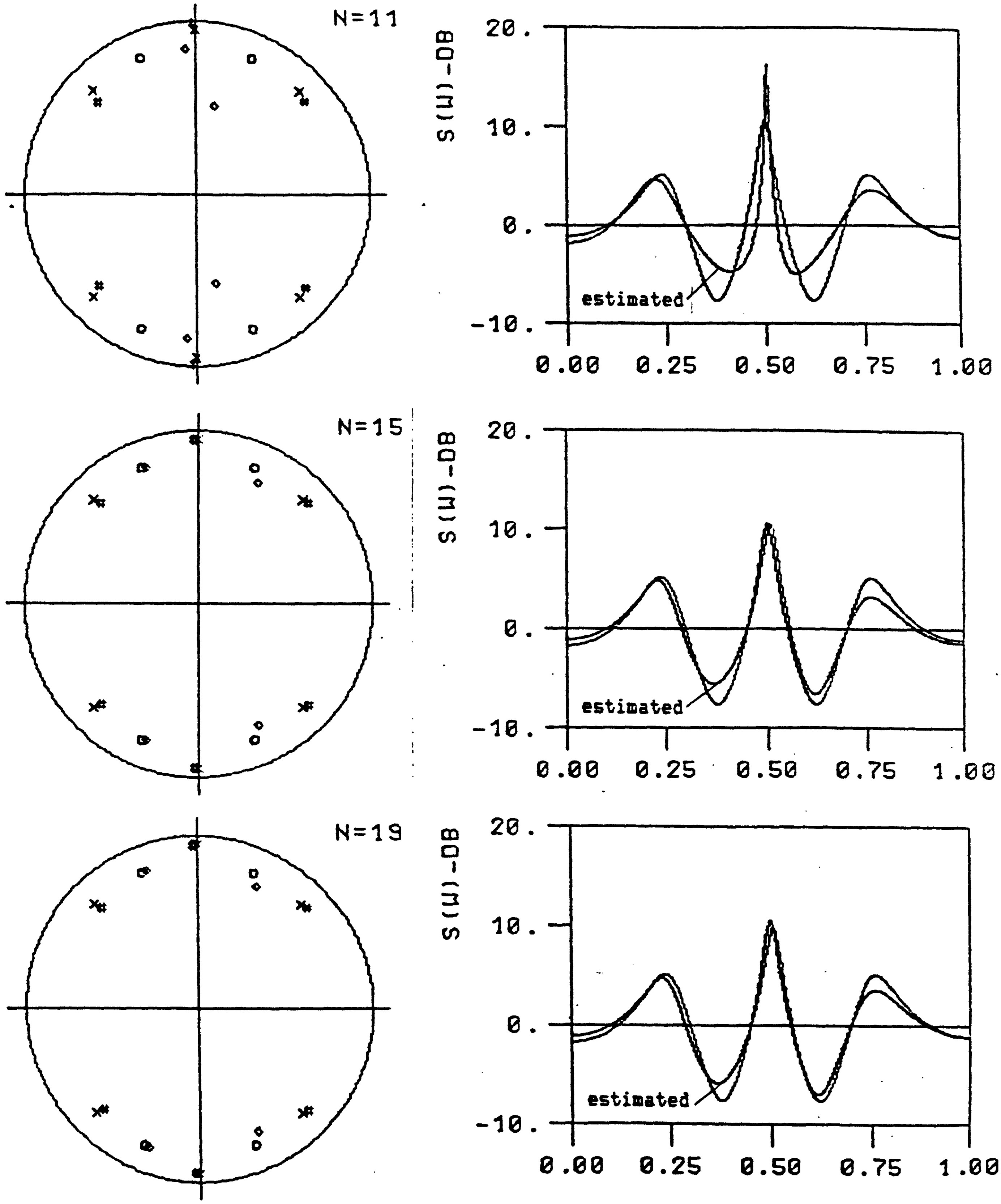

Fig.1 - Real and estimated values of the pole-zero pattern and spectrum, obtained with T=1000.

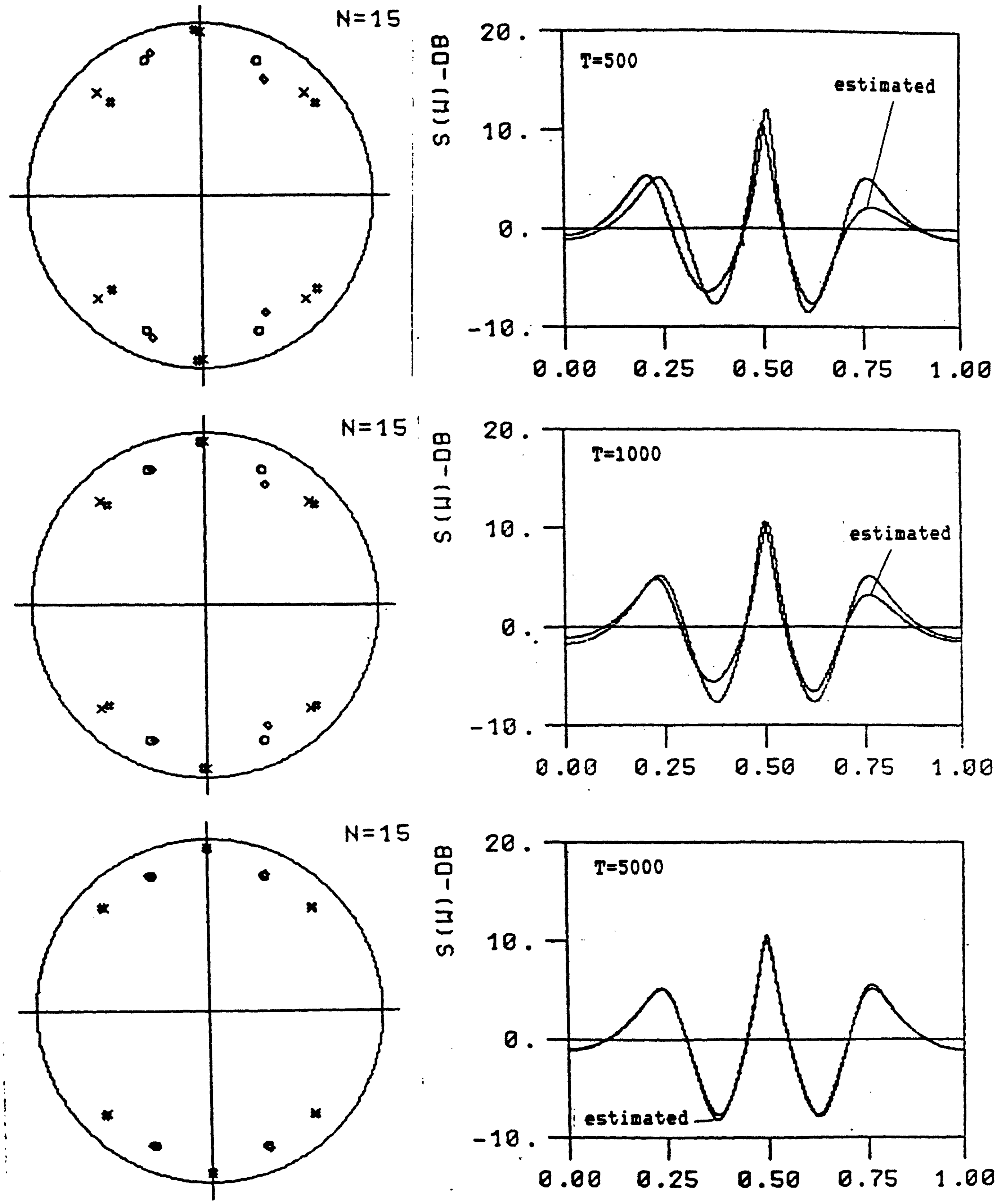

Fig.2 - Real and estimated values of the pole-zero pattern and spectrum, obtained for N=15 and T=500, 1000, 5000.

For small sample size, depending on the pole-zero pattern, the errors in the prediction error filter coefficients or innovation filter coefficients induce biasing errors in the pole-zero estimates. Following the analysis in [16] and [17], one can show that the bias on the coefficients $\{\hat{a}_i\}$, i=1,...,p, based on the estimated innovation filter coefficients up to order N is given by

$$E[\hat{\underline{a}}(N)]-\underline{a} = -\,(C_N^T C_N)^{-1}\,[C_N^T\,E[\delta C_N^*] + E(V_1) - C_N^T\,E(V_2)]\,\underline{a}^* + o(T^{-1}) \qquad (76)$$

where C_N is given in (54),

$$\underline{a}^* = [\,1 \quad a_1 \quad a_2 \quad \dots \quad a_p\,]^T \qquad (77a)$$

$$\delta C_N = C_N - \hat{C}_N \qquad (77b)$$

$$C_N^* = [\,-f_N \mid C_N\,] \qquad (77c)$$

with f_N as in (59), and

$$\delta C_N^* = C_N^* - \hat{C}_N^* \qquad (77d)$$

$$V_1 = \delta C_N^T\,(\,I - C_N(C_N^T\,C_N)^{-1}C_N^T\,)\,\delta C_N^* \qquad (77e)$$

$$V_2 = \delta C_N\,(C_N^T\,C_N)^{-1}\,C_N^T\,\delta C_N^*\,. \qquad (77f)$$

From [17] and [18], one can show that the bias and the error covariance on the elements of C_N^*, obtained through the Burg technique is given by

$$E\,[(\delta C_N^*)_{ij}\,] = T^{-1}\,Q_{ij} + o(T^{-1}) \qquad (78)$$

$$E\,[(\delta C_N^*)_{ij}\,(\delta C_N^*)_{kr}\,] = T^{-1}\,S_{ij,kr} + o(T^{-1}) \qquad (79)$$

with Q_{ij}, $S_{ij,kr}$ conveniently defined matrices. Consequently the bias on the autoregressive component estimation (76) is of order T^{-1}, meaning that the estimator is asymptotically unbiased. A dual analysis, leading to similar conclusions could have been presented for the MA component.

The bias effect is shown in the examples presented in figures 3 to 5 which represent the mean estimated spectrum obtained, for different values of N and T, with 100 Monte-Carlo runs.

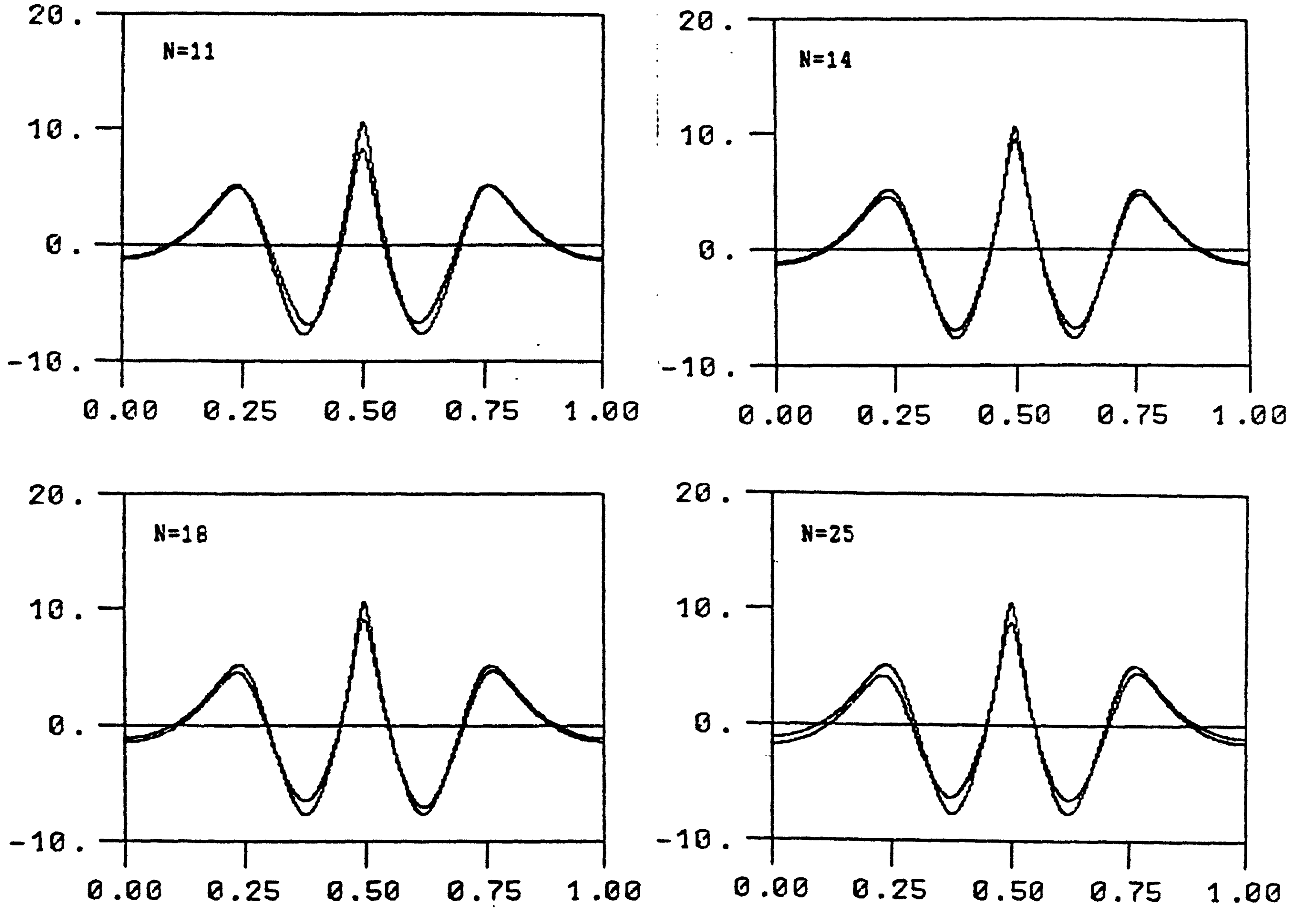

Fig.3 - Real and mean estimated spectrum obtained with 100 Monte-Carlo runs for T=500, N=11, 14, 18 and 25.

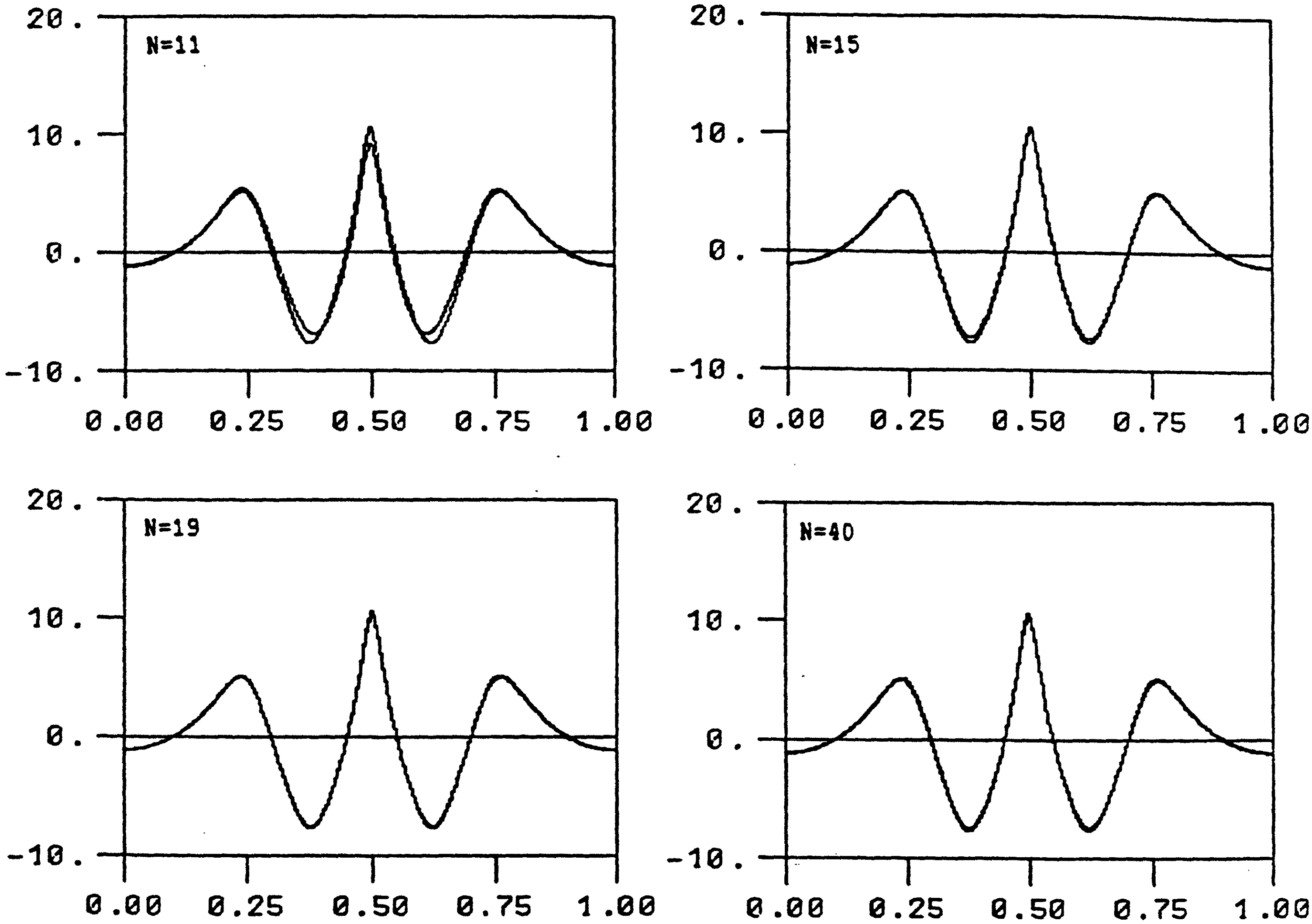

Fig.5 - Real and mean estimated spectrum obtained with 100 Monte-Carlo runs for T=5000, N=11, 15, 19 and 40.

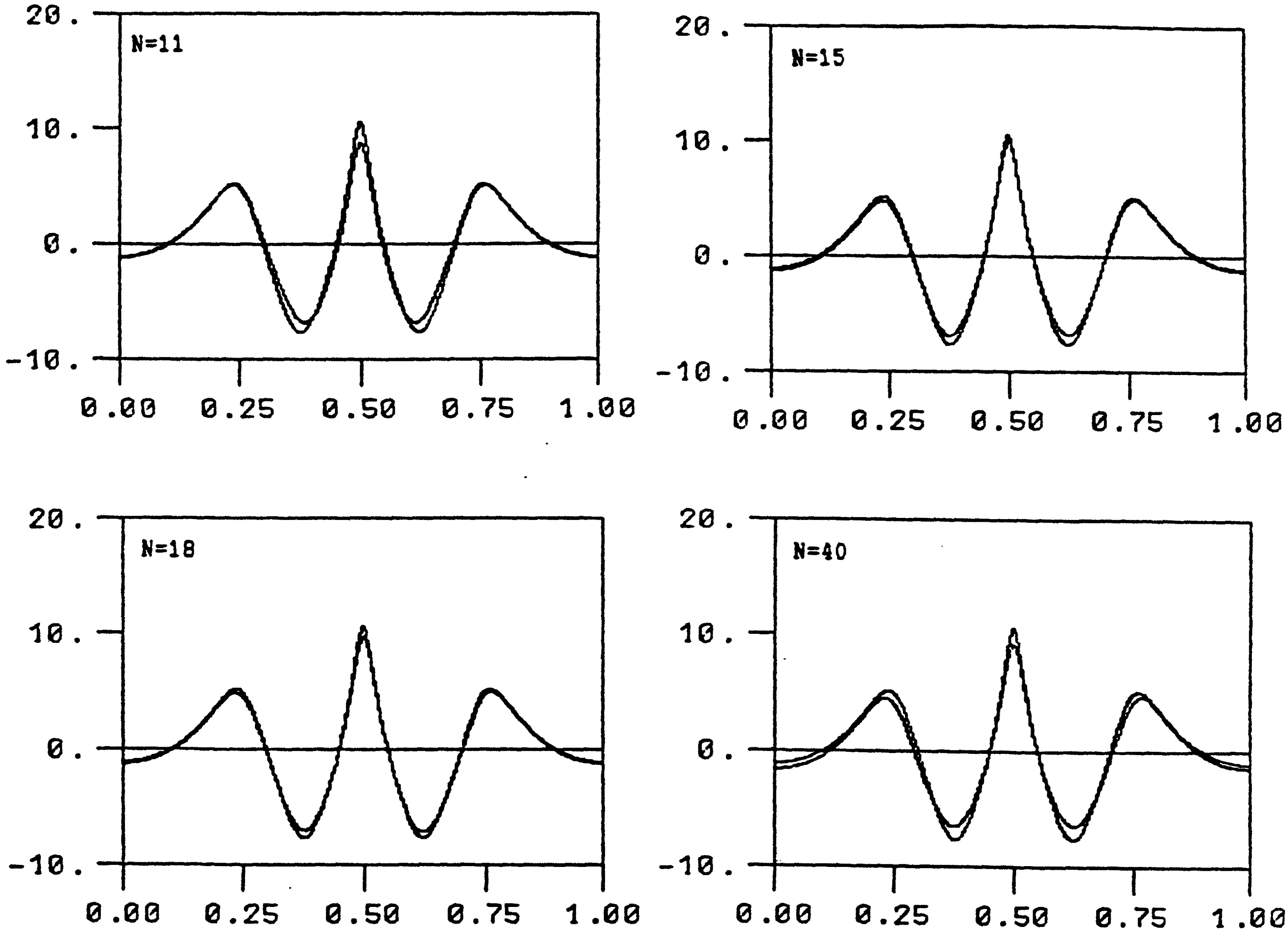

Fig.4 - Real and mean estimated spectrum obtained with 100 Monte-Carlo runs for T=1000, N=11, 15, 18 and 40.

Comparing the mean estimated spectrum obtained for N=11 in figures 3, 4 and 5, one sees that the bias decreases as the number of sample data points increases; the same is observed, for example, in figures 4 and 5 with N=15. This is in accordance with (76). However, for a fixed value of T, the increase of N does not always leads to a better performance.

In figure 3, and for N>14, the bias in the parameter estimates dominates the estimation error, thus leading to a worst performance for higher values of N. The same effect is evident in figure 4 for N>18. In figure 5, which is obtained for a higher value of T than the previous ones, the bias effect is not significant till N=40.

Following again the analysis in [17], the error covariance for the AR component estimation is also of order T^{-1}.

6. CONCLUSIONS

The work describes a method where the AR and MA components of a multivariable ARMA process are estimated by a dualized type algorithm. The estimation scheme provides a distinct Yule-Walker type equations for each component.

The estimation algorithm assumes that the orders p and q of the ARMA process are known. If they were unknown, the algorithm is still useful in fitting several classes of ARMA(p_i,q_i) to the data. A conveniently defined stopping rule picks up the "best" (in a certain sense) class. This is presently under experimentation.

In [19]-[20] an ARMA identification algorithm which is recursive on the orders is presented. Ours is different from the one in [19]-[20] in several regards. As mentioned before, the procedure studied here uses estimates of the coefficients of successively higher order linear predictors and innovation representations of the process, both of which can be obtained from the reflection coefficient sequence, avoiding the necessity of obtaining sample autocorrelations as in [19]-[20]. Also, the scheme dualizes the estimation of the AR part and of the MA part, without one interfering and degrading the other.

The AR coefficients are determined from the linear dependence exhibited by corresponding coefficients of successively higher order linear innovations filters. A recursive implementation of the AR estimation algorithm is also presented. In a pratical situation where the AR coefficients have to be estimated from a finite sample of the observation process, the algorithm requires suitable estimates of $\tilde{W}_N$ and $\tilde{W}_N^{-1}$. Because it simultaneously uses all the innovations filters coefficients, the numerical accuracy is then improved.

The MA coefficients are asymptotically determined from the linear dependence exhibited by corresponding coefficients of successively higher order linear predictor filters. The quality of the procedure is connected to the rate of this convergence, which is proved to go with the second power of the zeros of the original system.

Some simulation results together with a brief statistical and performance analysis are presented.

REFERENCES

[1] J.A.CADZOW, "Spectral Estimation: An Overdetermined Rational Model Approach", IEEE Proceedings, Vol.70, nº9, pp.907-939, Sept.1982.

[2] S.M.KAY, S.L.MARPLE, "Spectrum Analysis: A Modern Perspective", IEEE Proceedings, Vol.69, nº11, pp.1380-1427, Nov.1981.

[3] B.FRIEDLANDER, "Lattice Methods for Spectral Estimation", IEEE Proceedings, Vol.70, nº9, pp.990-1017, Sept.1982.

[4] J.P.BURG, "Maximum Entropy Spectral Analysis", Ph.D Thesis, Stanford University, 1975.

[5] H.WOLD, "A Study in the Analysis of Stationary Time Series", Uppsala, Almquist and Wiksells, 1938.

[6] J.L.DOOB, "Stochastic Processes", John Wiley, 1953.

[7] E.PARZEN, "Stochastic Processes", Holden-Day, 1967.

[8] R.E.KALMAN, R.S.BUCY, "New Results in Linear Filtering and Prediction Theory", ASME Journal of Basic Eng., pp.35-45, March 1961.

[9] R.S.BUCY, "Identification and Filtering", Math.Systems Theory, Vol.16, pp.307-317, Dec.1983.

[10] M.I.RIBEIRO, J.M.F.MOURA, "Dual Estimation of the Poles and Zeros of an ARMA(p,q) Process", CAPS, Technical Report 02/84, November 1984; Revised as LIDS Technical Report No. LIDS-P-1521, M.I.T., September 1985.

[11] R.S.BUCY, P.D.JOSEPH, "Filtering for Stochastic Processes whith Applications to Guidance", John Wiley and Sons, 1968.

[12] R.E.KALMAN, "When is a Linear Control System Optimal?", ASME Journal of Basic Eng., pp.51-60, March 1964.

[13] D.C.YOULA, "On the Factorization of Rational Matrices", IRE Trans. Information Theory, Vol.IT-7, pp.172-189, 1961.

[14] T.KAILATH, "Linear Systems", Prentice-Hall, 1980.

[15] M.I.RIBEIRO, "Estimação Paramétrica em Processos Autoregressivos de Média Móvel Vectoriais", forthcoming Ph.D. Thesis, Department of Electrical and Computer Engineering, Instituto Superior Técnico, Lisbon, Portugal, 1986.

[16] PORAT, B., FRIEDLANDER, B., "Asymptotic Analysis of the Bias of the Modified Yule-Walker Estimator", IEEE Transactions on Automatic Control, Vol.AC-30, No.8, pp.765-767, 1985.

[17] PORAT, B., FRIEDLANDER, B., "Asymptotic Accuracy of Arma Parameter Estimation Methods Based on Sample Covariances", IFAC on Identification and System Parameter Estimation, York, U.K., 1985.

[18] KAY, S., MAKHOUL, J., "On the Statistics of the Estimated Reflection Coefficients of an Autoregressive Process", IEEE Trans. ASSP, Vol.-31, No.6, December 1983.

[19] HANNAN, E. J., RISSANEN, J. "Recursive Estimation of Mixed Autoregressive - Moving Average Order", Biometrika, Vol.69, No.1, pp.81-94, 1982.

[20] FRANKE, J., "A Levinson-Durbin Recursion for ARMA-Processes", submitted for publication, 1984.