A SPECTRAL REPRESENTATION METHOD FOR CONTINUOUS-TIME STOCHASTIC SYSTEM ESTIMATION BASED ON ANALOG DATA RECORDS

P. Nurprasetio[1] and *S.D. Fassois*[2]

Report UM-MEAM-91-16

[1] Graduate Research Assistant, Mechanical Engineering and Applied Mechanics
[2] Assistant Professor, Mechanical Engineering and Applied Mechanics

engn
UMR 0955

1 INTRODUCTION

After two decades of almost complete dominance of discrete-time system identification and parameter estimation approaches in the engineering theory and practice, the relevance and importance of *continuous-time* methods based on *analog* data have been increasingly recognized (Young, 1981; Young and Jakeman, 1980; Unbehauen and Rao, 1987; Wahlberg, 1989; Neumann *et al.*, 1989; Sagara and Zhao, 1990; Unbehauen and Rao, 1990). A survey of the available literature, however, quickly reveals that the overwhelming majority of currently available continuous-time methods is restricted to the deterministic case, and is thus incapable of handling the more general and practically important class of *stochastic* systems[1], with which additional difficulties are associated (Young, 1981; Wahlberg, 1989).

One of the main difficulties in estimating continuous-time stochastic systems from analog data records is due to our inability to accurately compute time derivatives (of various orders) of the observed random signals (Young, 1981; Wahlberg, 1989; Neumann *et al.*, 1989). If that were possible, continuous-time versions of many discrete methods could be constructed; see, for instance, the works of Balakrishnan (1973), Bagchi (1975), and Pham-Dinh-Tuan (1977), who developed and analyzed the Maximum Likelihood method for this case and proved its asymptotic optimality, and also Unbehauen and Rao (1987), and, Priestley (1981), who discussed the autoregressive model case. The use of the so-called State Variable Filters (SVFs), that are extensively used in the deterministic case (Young and Jakeman, 1980; Young, 1981; Unbehauen and Rao, 1987) in alleviating this difficulty is neither trivial nor effective, as their proper selection is not obvious [and in fact it is known to optimally require a-priori system knowledge, in view of which adaptive SVFs have been also suggested (Neumann *et al.*, 1989)], their incorporation affects the stochastic part of the model, and, unlike the deterministic case, cannot prevent the introduction of large errors into the obtained random signal derivatives (Neumann *et al.*, 1989).

Attempts to overcome these problems through alternative approaches based on integral (instead of differential) operators have been also considered. Van Schuppen (1983) examined recursive such estimation algorithms for continuous-time autoregressive processes and proved their convergence, while Moore (1988) analyzed the convergence of the continuous-time version of the Recursive Extended Least Squares (RELS) estimation algorithm for AutoRegressive Moving Average with eXogeneous excitation (ARMAX) systems. In order to avoid numerical instability problems due to the pure integration of the observed signals, exponentially stable prefilters have to be, however, used, the selection of which is important for convergence and, also, requires some form of a-priori system information. Another related problem in this case is that of initial conditions, the effects of which

[1] A large number of engineering and physical systems are, indeed, stochastic in nature. Consider, for instance, the ambient vibrations of a building or structure, the vibrations of a machine tool during cutting, or those of traveling ground vehicles and aircrafts.

do not decay away and can cause significant estimation bias errors (Young, 1981; Sagara and Zhao, 1990).

It therefore comes as no surprise that, for almost all practical cases, the stochastic part of a given system is either *not* estimated at all, or, only a *discrete-time* representation of it is obtained based on *sampled* data (Young and Jakeman, 1980; Young, 1981; Neumann *et al.*, 1989; Sagara and Zhao, 1990; Lee and Fassois, 1991). The only, and arithmetically very few, currently available exceptions, appear to be based on an approach that postulates the estimation of continuous-time stochastic models through the estimation of a directly, and *approximately, discretized* representation by using, once again, *uniformly sampled* data records (Åström and Källström, 1973; Phadke and Wu, 1974; Pandit and Wu, 1975). The method developed by Pandit and Wu (Pandit, 1973; Wu, 1977; Pandit and Wu, 1983) is, in fact, a very systematic and comprehensive such procedure, that has been successfully used for the solution of a large number of production engineering, random vibration, and other types of problems. A number of additional such methods have been also, though independently, developed in the fields of statistics (Robinson, 1980; Jones, 1981; Solo, 1984) and econometrics (Sargan, 1974; Bergstrom, 1976; 1983), where the assumption of sampled data is practically reasonable.

The main advantage of these methods is in their ability to utilize the already available machinery for discrete-time system estimation (Priestley, 1981; Pandit and Wu, 1983; Ljung, 1987). Yet, and apart from being limited to *sampled* data records, a number of problems are known to be associated with them, namely: (a) The estimates of the continuous-time parameters are asymptotically biased; with the bias errors being due to the approximate nature of the employed discretization [Young (1981); also see the discussion in Ben Mrad and Fassois (1991), and, Lee and Fassois (1991)], and also analyzed in the econometrics literature [Wymer (1972); Sargan (1974); Bergstrom (1976)], (b) additional errors are introduced into the continuous-time parameter estimates due to sensitivity problems associated with the highly nonlinear discrete-to-continuous transformation (Fassois *et al.*, 1990); with the related question of sampling also being recognized as nontrivial (Fassois *et al.*, 1990; Unbehauen and Rao, 1990), (c) the loss of relative order, that is the difference between the degrees of the numerator and denominator polynomials in the system transfer functions, (d) the incorporation of a-priori system information into the estimation procedure is very difficult, if not completely impossible, due to the highly complicated and nonlinear nature of the discrete-to-continuous transformation (Wu, 1977; Wahlberg, 1989), (e) high computational complexity, which is further aggrevated by the aforementioned nonlinear transformation.

In this paper a novel *Maximum Likelihood type* method for the estimation of *continuous-time stochastic* systems from *analog* data, that utilizes the general Autoregressive Moving Average with eXogeneous excitation (ARMAX) stochastic model structure and block-pulse function (BPF) spectral representations, and overcomes many of the aforementioned difficulties and limitations, is introduced. The use of BPF signal representations has led to a number of developments in deterministic system theory and identification in recent years [see, for instance, Wang (1982), and, Jiang and Schaufelberger (1985)]; the present work, however, appears to be the first that, building upon them, utilizes BPF expansions in solving stochastic estimation problems.

Within the context of this paper the validity of BPF spectral representations for stochastic signals is, first, formally justified, and, then, based on them, as well as a set of linear operations motivated by recent developments in deterministic identification (Jiang and Schaufelberger, 1985), the problem of interest is shown to be transformed into that of estimating the parameters of an induced stochastic difference equation driven by the spectral representations of exogeneous (observable) and endogeneous (unobservable) excitations; the latter essentially being the spectral expansion of a Wiener process. The structural and probabilistic properties of this difference equation are studied, and the endogeneous excitation is shown to be amenable to a first-order Integrated Moving Average (IMA) representation driven by a stationary innovations sequence. Based on this, as well as the structural properties of the stochastic difference equation, an induced, special-form, discrete ARMAX system with parameters that are expressed as *linear* combinations of those of the original stochastic differential equation, is finally obtained.

The mapping between this discrete ARMAX system and the original stochastic differential equation is shown to be a *bijective transformation*, and, the discrete system *stationary* and *invertible*. These features lead to a Maximum Likelihood type estimation scheme that is based on the estimation of the discrete ARMAX system from which the parameters of the stochastic differential equation are subsequently obtained through a simple linear transformation. The linearity of this transformation, which results in reduced computational complexity while also circumventing sensitivity problems and allowing for the incorporation of a-priori information into the estimation procedure, along with the use of analog data without requiring estimates of signal derivatives, and the elimination of direct discretization procedures, are among the main features of this method, the performance characteristics of which are finally evaluated via a number of simulation experiments.

The remaining part of this paper is organized as follows: The exact problem statement is presented in Section 2, some preliminary considerations and the validity of BPF spectral expansions for random signals are discussed in Section 3, the induced stochastic difference equation is derived in Section 4, and its structural properties analyzed in 5. The stochastic modeling of the endogeneous signal spectral representation is discussed in Section 6, and the estimation method formulated in Section 7. Simulation results are presented in Section 8, and the conclusions are finally summarized in Section 9.

2 PROBLEM STATEMENT

Consider a continuous-time stochastic system with exogeneous (observable) excitation $\{u(t)\}$ and corresponding response $\{y(t)\}$ described by the ARMAX stochastic differential equation[2]:

$$\mathcal{S}_C: \qquad \sum_{i=0}^{n_a} a_i^\circ \frac{d^i y(t)}{dt^i} = \sum_{j=0}^{n_a-1} b_j^\circ \frac{d^j u(t)}{dt^j} + \sum_{k=0}^{n_a-1} c_k^\circ \frac{d^k w(t)}{dt^k} \tag{1}$$

in which the differentiation operations are to be interpreted in the mean-square sense (Jazwinski, 1970), $a_{n_a}^\circ \triangleq 1$, $c_{n_a-1}^\circ \triangleq 1$, and $\{w(t)\}$ represents an endogeneous (unobservable) zero-mean continuous-time innovations signal with autocovariance function:

$$E[w(s)w(t)] = (\sigma_w^\circ)^2 \cdot \delta(s-t) \tag{2}$$

where $\delta(\cdot)$ stands for the Dirac delta function. By using the mean-square differential operator D, the system $\mathcal{S}_C$ may be equivalently rewritten in the notational form:

$$\mathcal{S}_C: \qquad a^\circ(D)\cdot y(t) \;=\; b^\circ(D)\cdot u(t) \;+\; c^\circ(D)\cdot w(t) \tag{3}$$

with $a^\circ(D)$, $b^\circ(D)$, and $c^\circ(D)$ being the *Autoregressive* (AR), *Exogeneous* (X), and *Moving-Average* (MA) polynomials, respectively, which are of the forms:

$$a^\circ(D) \triangleq D^{n_a} + a_{n_a-1}^\circ D^{n_a-1} + \ldots + a_1^\circ D + a_0^\circ \tag{4a}$$

$$b^\circ(D) \triangleq b_{n_a-1}^\circ D^{n_a-1} + b_{n_a-2}^\circ D^{n_a-2} + \ldots + b_1^\circ D + b_0^\circ \tag{4b}$$

$$c^\circ(D) \triangleq D^{n_a-1} + c_{n_a-2}^\circ D^{n_a-2} + \ldots + c_1^\circ D + c_0^\circ \tag{4c}$$

The ARMAX system $\mathcal{S}_C$ is additionally assumed to satisfy the following standard assumptions:

A1. The polynomials $a^\circ(D)$, $b^\circ(D)$, and $c^\circ(D)$ are coprime (*irreducibility* assumption).

A2. The polynomials $a^\circ(D)$ and $c^\circ(D)$ are strictly minimum phase (*stationarity* and *invertibility* assumptions, respectively).

A3. The signals $\{w(t)\}$, $\{u(t)\}$, and $\{y(t)\}$ are Gaussian, with the latter two additionally being continuous in probability and having almost every sample path characterized by finite energy within the observation interval $[0,T)$, namely $\int_0^T [x(t)]^2\, dt < \infty$ almost surely (a.s.) (Jazwinski, 1970).

[2]The superscript $^\circ$ is used to indicate quantities associated with the actual system and distinguish them from those of corresponding candidate models.

As we will see the Gaussianity assumption is not critical for the development of the estimation method, but only the vehicle in casting the problem into a Maximum Likelihood framework. It may be thus relaxed, and the method formally interpreted within the broader Prediction Error estimation context (Ljung, 1987).

The estimation problem considered in this work may be then posed as follows:

"Given analog excitation $\{u(t)\}$ and response $\{y(t)\}$ data records generated by the system $\mathcal{S}_C$, subject to assumptions A1-A3, over a period of time $[0, T)$, estimate an ARMAX model of the form:

$$\mathcal{M}_C: \qquad a(D)\cdot y(t) = b(D)\cdot u(t) + c(D)\cdot w(t) \qquad E[w(s)w(t)] = \sigma_w^2 \cdot \delta(s-t) \tag{5}$$

that matches the system $\mathcal{S}_C$ as closely as possible."

3 PRELIMINARY CONSIDERATIONS

The ARMAX equation (1) provides a strictly formal representation of the underlying stochastic system, as it is well known that the continuous-time white noise signal is not second-order and its mean-square derivatives fail to exist. A more appropriate, and also convenient for our purposes, representation may be thus obtained by integrating the differential equation (1) n_a times, which, assuming, for simplicity, zero initial conditions, yields:

$$\mathcal{S}_{\bar{C}}: \qquad y(t) + a^{\circ}_{n_a-1}\int_0^t y(t')dt' + \cdots + a^{\circ}_0 \underbrace{\int_0^t \cdots \int_0^t}_{n_a-fold} y(t')\,dt' =$$

$$= b^{\circ}_{n_a-1}\int_0^t u(t')dt' + \cdots + b^{\circ}_0 \underbrace{\int_0^t \cdots \int_0^t}_{n_a-fold} u(t')dt' + \tilde{w}(t) + \cdots + c^{\circ}_0 \underbrace{\int_0^t \cdots \int_0^t}_{(n_a-1)-fold} \tilde{w}(t')dt' \tag{6}$$

In this expression the integration operations are to be interpreted in the mean-square sense as well (Jazwinski, 1970), while $\{\tilde{w}(t)\}$ represents the Wiener process:

$$\tilde{w}(t) \triangleq \int_0^t w(s)ds \tag{7}$$

which, due to the stated properties of $\{w(t)\}$, is Gaussian, continuous in probability, with $\int_0^T [\tilde{w}(t)]^2\,dt < \infty$ (a.s.), and also zero-mean and with autocovariance function (Jazwinski, 1970):

$$E[\tilde{w}(s)\tilde{w}(t)] = (\sigma^{\circ}_w)^2 \cdot \min(s,t) \tag{8}$$

For the development of the estimation method the m-th order block-pulse function (BPF) spectral representations (Kwong and Chen, 1981) of the signals $\{u(t)\}$, $\{\tilde{w}(t)\}$, and $\{y(t)\}$, are now used. Within the observation interval $[0,T)$ these signals may be then expressed as follows[3]:

$$u(t) \cong u_m(t) \triangleq \sum_{k=1}^{m} U_k\,\psi_k(t) = \mathbf{U}^T\mathbf{\Psi}(t) \tag{9a}$$

$$y(t) \cong y_m(t) \triangleq \sum_{k=1}^{m} Y_k\,\psi_k(t) = \mathbf{Y}^T\mathbf{\Psi}(t) \tag{9b}$$

$$\tilde{w}(t) \cong \tilde{w}_m(t) \triangleq \sum_{k=1}^{m} \tilde{W}_k\,\psi_k(t) = \tilde{\mathbf{W}}^T\mathbf{\Psi}(t) \tag{9c}$$

where $\{U_i\}$, $\{Y_i\}$ and $\{\tilde{W}_i\}$ represent the sequences of BPF expansion coefficients of $\{u(t)\}$, $\{y(t)\}$, and $\{\tilde{w}(t)\}$, respectively, $\mathbf{U}$, $\mathbf{Y}$, and $\tilde{\mathbf{W}}$ corresponding vector representations, and $\mathbf{\Psi}(t)$ an m-dimensional BPF vector of the form:

$$\mathbf{\Psi}(t) \triangleq [\psi_1(t)\ \psi_2(t)\ \cdots\ \psi_m(t)]^T \tag{10}$$

with $\psi_k(t)$ representing the k-th block pulse function of the m-th order BPF function set:

$$\psi_k(t) = \begin{cases} 1 & \text{for} \quad (k-1)\frac{T}{m} \le t < k\frac{T}{m} \\ 0 & \text{otherwise} \end{cases} \qquad (k = 1, 2, \ldots, m) \tag{11}$$

and T/m the block pulse function duration ("width"). For a given signal, say $\{y(t)\}$, the BPF expansion coefficients can be computed through the expression:

$$Y_k \triangleq \frac{m}{T}\int_0^T y(t)\,\psi_k(t)\,dt = \frac{m}{T}\int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} y(t)\,dt \qquad (k = 1, \cdots, m) \tag{12}$$

Such BPF expansions are frequently used for the representation of deterministic signals, and it is well-known, that, in the limit, as m approaches infinity, the set (11) is complete, and the signal representation $\{y_m(t)\}$ given by (9b) converges to $\{y(t)\}$ pointwise for any deterministic square-integrable signal defined in the interval $[0,T)$ (Kwong and Chen, 1981). Within the context of the present work the question of validity of such expansions for stochastic signals is an apparently legitimate and important one, and is therefore addressed in the following theorem:

Theorem 1: The BPF spectral expansions (9a)-(9c) of the stochastic signals $\{u(t)\}$, $\{y(t)\}$, and $\{\tilde{w}(t)\}$ are convergent in the following product measure sense (exemplified for the case of $\{y(t)\}$):

$$\lim_{m\to\infty} (P \times \lambda)\{(\omega, t) : |y(t) - y_m(t)| \ge \epsilon\} = 0 \qquad \forall\, \epsilon > 0 \tag{13}$$

with ω denoting an elementary event of the sample space, P the probability measure, and λ the Lebesgue measure on $[0,T)$. □

[3] Bold-face characters indicate vector/matrix quantities.

Proof:

The proof is a direct consequence of the completeness of the set (11) as $m \to \infty$, the continuity in probability and finite energy for almost every sample path, that is $\int_0^T [y(t)]^2\,dt < \infty$ (a.s.), properties of the random signals involved (see assumption A3 and the earlier discussion on $\{\tilde{w}(t)\}$), and Theorem 2 of Bharucha and Kadota (1970), which states that the expansion of a random process $\{y(t)\}$, with $t \in [0,T)$, in terms of an arbitrary basis in $\mathcal{L}_2(T)$, the space of square-integrable functions, is always convergent in the above product measure sense for signals satisfying the aforementioned conditions, regardless of the orthogonality of the basis used and the boundedness of the time interval T.

4 AN INDUCED MAPPING AND A STOCHASTIC DIFFERENCE EQUATION REPRESENTATION

By substituting the expansions (9) into the system expression (6), and using the operational matrix form of the BPF spectral representation for integration (see Appendix), we arrive at the following system representation:

$$\mathbf{Y}^T \sum_{i=0}^{n_a} a^{\circ}_{n_a-i}\mathbf{F}_i\Psi(t) = \mathbf{U}^T \sum_{i=0}^{n_a-1} b^{\circ}_{n_a-i-1}\mathbf{F}_{i+1}\Psi(t) + \tilde{\mathbf{W}}^T \sum_{i=0}^{n_a-1} c^{\circ}_{n_a-i-1}\mathbf{F}_i\Psi(t) \tag{14}$$

in which $\mathbf{F}_k$ represents the $m \times m$ operational matrix form for k-fold integration (see Appendix). By canceling the vector $\Psi(t)$ from both sides of this equation and performing the algebra, we obtain (n_a+1) equations $E_0, \ldots, E_{n_a}$ of the form:

$$E_k: \quad \sum_{i=0}^{n_a} \Big(\frac{(T/m)^i}{(i+1)!} a^{\circ}_{n_a-i} \sum_{j=1}^{l+k} Y_j f_{i,l+k+1-j}\Big) = \sum_{i=0}^{n_a-1} \Big(\frac{(T/m)^{i+1}}{(i+2)!} b^{\circ}_{n_a-1-i} \sum_{j=1}^{l+k} U_j f_{i+1,l+k+1-j}\Big)$$
$$+ \sum_{i=0}^{n_a-1} \Big(\frac{(T/m)^i}{(i+1)!} c^{\circ}_{n_a-1-i} \sum_{j=1}^{l+k} \tilde{W}_j f_{i,l+k+1-j}\Big) \qquad (k = 0, 1, \cdots, na) \tag{15}$$

with l representing an arbitrary positive integer and $f_{k,i}$ a quantity defined in the Appendix [Eq.(A.3)]. In a manner analogous to that used by Jiang and Schaufelberger (1985) for deterministic systems, by performing the linear operation:

$$\sum_{k=0}^{n_a} (-1)^k \binom{n_a}{k} E_{n_a-k} \tag{16}$$

on the set of equations $\{E_k\}$, we obtain the following stochastic difference equation:

$$\sum_{i=0}^{n_a} A^{\circ}_i \cdot Y_{l+i} = \sum_{i=0}^{n_a} B^{\circ}_i \cdot U_{l+i} + \sum_{i=0}^{n_a} \tilde{C}^{\circ}_i \cdot \tilde{W}_{l+i} \tag{17}$$

which is independent of any initial conditions, and, through appropriate reindexing, may be expressed as:

$$\mathcal{S}_{\tilde{\mathcal{D}}}: \qquad A^\circ(B)\cdot Y_k = B^\circ(B)\cdot U_k + \tilde{C}^\circ(B)\cdot \tilde{W}_k \qquad (k=1,2,\cdots,m) \tag{18}$$

In this equation $\{Y_k\}, \{U_k\}$ and $\{\tilde{W}_k\}$ represent the BPF expansions of the response, exogeneous excitation, and Wiener process, respectively, and $A^\circ(B), B^\circ(B)$, and $\tilde{C}^\circ(B)$, polynomials in the backshift operator B ($B\cdot Y_k \triangleq Y_{k-1}$) and of the respective forms:

$$A^\circ(B) \triangleq A^\circ_{n_a} + A^\circ_{n_a-1}\cdot B + \ldots + A^\circ_1\cdot B^{n_a-1} + A^\circ_0\cdot B^{n_a} \tag{19a}$$

$$B^\circ(B) \triangleq B^\circ_{n_a} + B^\circ_{n_a-1}\cdot B + \ldots + B^\circ_1\cdot B^{n_a-1} + B^\circ_0\cdot B^{n_a} \tag{19b}$$

$$\tilde{C}^\circ(B) \triangleq \tilde{C}^\circ_{n_a} + \tilde{C}^\circ_{n_a-1}\cdot B + \ldots + \tilde{C}^\circ_1\cdot B^{n_a-1} + \tilde{C}^\circ_0\cdot B^{n_a} \tag{19c}$$

The coefficients $\{A^\circ_i\}_{i=0}^{na}$ may be shown to be related to the coefficients of the original stochastic differential equation $\{a^\circ_i\}_{i=0}^{na-1}$ through the expressions:

$$A^\circ_0 = \sum_{i=0}^{na}(-1)^{n_a+i}\frac{(T/m)^i}{(i+1)!}a^\circ_{n_a-i} \tag{20a}$$

$$A^\circ_i = \sum_{j=0}^{n_a}\left(\frac{(T/m)^j}{(j+1)!}a^\circ_{n_a-j}\sum_{k=0}^{n_a-i}(-1)^k\begin{pmatrix} n_a \\ k \end{pmatrix} f_{j,n_a-k-i+1}\right) \qquad (i=1,2,\ldots,n_a) \tag{20b}$$

and similar expressions relate the rest of the coefficients in (19) with those of the original equation $\mathcal{S}_C$. By defining the continuous and discrete parameter vectors as:

$$\mathbf{a} \triangleq [a^\circ_{n_a}\ \cdots\ a^\circ_1\ a^\circ_0]^T \qquad [(n_a+1)\times 1] \tag{21a}$$

$$\mathbf{b} \triangleq [b^\circ_{n_a-1}\ \cdots\ b^\circ_1\ b^\circ_0]^T \qquad [n_a\times 1] \tag{21b}$$

$$\mathbf{c} \triangleq [c^\circ_{n_a-1}\ \cdots\ c^\circ_1\ c^\circ_0]^T \qquad [n_a\times 1] \tag{21c}$$

and:

$$\mathbf{A} \triangleq [A^\circ_{n_a}\ \cdots\ A^\circ_1\ A^\circ_0]^T \qquad [(n_a+1)\times 1] \tag{22a}$$

$$\mathbf{B} \triangleq [B^\circ_{n_a}\ \cdots\ B^\circ_1\ B^\circ_0]^T \qquad [(n_a+1)\times 1] \tag{22b}$$

$$\tilde{\mathbf{C}} \triangleq [\tilde{C}^{\circ}_{n_a} \; \ldots \; \tilde{C}^{\circ}_1 \; \tilde{C}^{\circ}_0]^T \qquad [(n_a+1) \times 1] \tag{22c}$$

respectively, the relations between the discrete and continuous system parameters may be compactly expressed as:

$$\mathbf{A} = \mathbf{D}_A \, \mathbf{a} \qquad \mathbf{B} = \mathbf{D}_B \, \mathbf{b} \qquad \tilde{\mathbf{C}} = \mathbf{D}_C \, \mathbf{c} \tag{23}$$

with $\mathbf{D}_A$ being a square matrix with elements determined from (20), and $\mathbf{D}_B$ and $\mathbf{D}_C$ submatrices of $\mathbf{D}_A$ formed by expressing $\mathbf{D}_A$ in terms of its column vectors as:

$$\mathbf{D}_A = \left[\; \mathbf{d}_A(1) \;\; | \;\; \ldots \;\; | \;\; \mathbf{d}_A(n_a+1) \; \right] \qquad [(n_a+1) \times (n_a+1)] \tag{24a}$$

and defining:

$$\mathbf{D}_B = \left[\; \mathbf{d}_A(2) \;\; | \;\; \ldots \;\; | \;\; \mathbf{d}_A(n_a+1) \; \right] \qquad [(n_a+1) \times n_a] \tag{24b}$$

$$\mathbf{D}_C = \left[\; \mathbf{d}_A(1) \;\; | \;\; \ldots \;\; | \;\; \mathbf{d}_A(n_a) \; \right] \qquad [(n_a+1) \times n_a] \tag{24c}$$

These mapping relationships between the continuous and corresponding discrete parameters are summarized in Table 1 for up to fourth-order systems.

Equations (23) define a mapping relationship $\mathcal{T}$ between the sets:

$$\mathcal{C} \triangleq \{(\mathbf{a},\mathbf{b},\mathbf{c}) \in \mathcal{R}^{n_a+1} \times \mathcal{R}^{n_a} \times \mathcal{R}^{n_a} \mid \text{with } \mathbf{a},\mathbf{b},\mathbf{c} \text{ of the form (21) with } a^{\circ}_{n_a} = c^{\circ}_{n_a-1} = 1\} \tag{25a}$$

and:

$$\tilde{\mathcal{D}} \triangleq \{(\mathbf{A},\mathbf{B},\tilde{\mathbf{C}}) \in \mathcal{R}^{n_a+1} \times \mathcal{R}^{n_a+1} \times \mathcal{R}^{n_a+1} \mid \mathbf{A} = \mathbf{D}_A \, \mathbf{a};\; \mathbf{B} = \mathbf{D}_B \, \mathbf{b};\; \tilde{\mathbf{C}} = \mathbf{D}_C \, \mathbf{c};\; \text{for all } (\mathbf{a},\mathbf{b},\tilde{\mathbf{c}}) \in \mathcal{C}\} \tag{25b}$$

The nature of this mapping is of particular importance for our developments and is therefore examined in the following lemma:

<u>Lemma 1</u>: The mapping $\mathcal{T} : \mathcal{C} \rightarrow \tilde{\mathcal{D}}$ is a bijective (one-to-one and onto) transformation.

<u>*Proof*</u>:

$\mathcal{T}$ is a transformation since for any triple $(\mathbf{a},\mathbf{b},\mathbf{c}) \in \mathcal{C}$ expressions (23) define a unique triple $(\mathbf{A},\mathbf{B},\tilde{\mathbf{C}}) \in \tilde{\mathcal{D}}$. $\mathcal{T}$ is also onto since by construction of $\tilde{\mathcal{D}}$ every triple $(\mathbf{A},\mathbf{B},\tilde{\mathbf{C}}) \in \tilde{\mathcal{D}}$ is the image of at least one triple $(\mathbf{a},\mathbf{b},\mathbf{c}) \in \mathcal{C}$. That $\mathcal{T}$ is one-to-one follows from the full rank property of the matrix $\mathbf{D}_A$; a fact that may be shown by using expressions developed by Kraus and Schaufelberger (1990) in such a way as to rewrite $\mathbf{D}_A$ as the product of two matrices that may be verified to be nonsingular. Therefore the (sub)matrices $\mathbf{D}_B$ and $\mathbf{D}_C$ are also full rank, and the one-to-one property follows.

5 STRUCTURAL PROPERTIES OF THE STOCHASTIC DIFFERENCE EQUATION $\mathcal{S}_{\tilde{\mathcal{D}}}$

Before the induced stochastic difference equation $\mathcal{S}_{\tilde{\mathcal{D}}}$ of Eq.(18) can be estimated, its structural and probabilistic properties need to be determined. Specifically, issues such as the stationarity of $\mathcal{S}_{\tilde{\mathcal{D}}}$, its identifiability, and the first and second order properties of the sequence $\{\tilde{W}_k\}$, are all essential for constructing a proper estimation algorithm. It is emphasized that the study of the last issue is indeed necessary, as the noise dynamics of $\mathcal{S}_{\tilde{\mathcal{D}}}$ are *not* completely determined by the polynomial $\tilde{C}^\circ(B)$ alone, but also depend upon the correlation structure of $\{\tilde{W}_k\}$.

With these ideas in mind we proceed to examine the structural properties of the stochastic difference equation $\mathcal{S}_{\tilde{\mathcal{D}}}$ first. The following theorem discusses the phase characteristics of the polynomials $A^\circ(B)$ and $\tilde{C}^\circ(B)$ of (18), and therefore the latter's stationarity and "partial" invertibility properties:

Theorem 2: Consider the continuous-time ARMAX system subject to assumptions A1 and A2, and its corresponding stochastic difference equation $\mathcal{S}_{\tilde{\mathcal{D}}}$ given by (18). Each one of the polynomials $A^\circ(B)$ and $\tilde{C}^\circ(B)$ of the latter will then be:

1. Strictly minimum phase, provided that its continuous-time counterpart is strictly minimum phase.
2. Minimum phase, provided that its continuous-time counterpart is minimum phase.
3. Nonminimum phase, provided that its continuous-time counterpart is nonminimum phase.

Proof:
(a) Let us examine the $A^\circ(B)$ polynomial first.

(a1) Consider the system $y(t) = g^\circ(D) \cdot u(t)$ with the strictly proper [see (4)] transfer function $g^\circ(D) \triangleq b^\circ(D)/a^\circ(D)$. For a strictly minimum phase $a^\circ(D)$ the system $g^\circ(D)$ is asymptotically stable (Chen, 1984), which implies that for every bounded excitation the response will be also bounded, that is:

$$\forall \{u(t)\}: \quad \|u(t)\|_\infty < \infty \quad \Longrightarrow \quad \|y(t)\|_\infty < \infty \tag{26}$$

where $\|u(t)\|_\infty \triangleq \sup_{t\in J_1} |u(t)|$ with $J_1 \triangleq [0,\infty)$. Now consider the corresponding discrete system $G^\circ(B) \triangleq B^\circ(B)/A^\circ(B)$ induced by the BPF expansions of fixed width T/m. Since $G^\circ(B)$ is (by Lemma 1) unique, its response $\{Y_k\}$ to any given excitation $\{U_k\}$ will be also unique. An arbitrary bounded excitation $\{U_k\}$ can be, however, constructed from a bounded $\{u(t)\}$ through (12); select for instance $u(t) = U_k$ for $(k-1)\frac{T}{m} \le t < k\frac{T}{m}$. From (26) the response $\{y(t)\}$ to the excitation $\{u(t)\}$ will be bounded, and, therefore, the (uniquely-determined) response $\{Y_k\}$ will be also bounded because:

$$\begin{aligned} \|Y_k\|_\infty &\le \frac{m}{T}\int_0^\infty \|y(t)\|_\infty \cdot \|\psi_k(t)\|_\infty \cdot dt \\ &= \frac{m}{T}\cdot \|y(t)\|_\infty \cdot \int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} dt \;=\; \|y(t)\|_\infty \;<\; \infty \end{aligned} \tag{27}$$

where $\|Y_k\|_\infty \triangleq \sup_{t\in J_2} |Y_k|$ with $J_2 \triangleq [1,\infty)$[4]. As a consequence, an arbitrary bounded excitation $\{U_k\}$ results in a bounded response $\{Y_k\}$, the system $Y_k = G^\circ(B)\cdot U_k$ hence is asymptotically stable, and the polynomial $A^\circ(B)$ strictly minimum phase (Chen, 1984).

(a2) For a minimum phase $a^\circ(D)$ the system $g^\circ(D) \triangleq b^\circ(D)/a^\circ(D)$ is stable in the sense of Lyapunov, and, therefore, its impulse response function $\{g(t)\}$ bounded:

$$\|g(t)\|_\infty < \infty \tag{28}$$

The corresponding discrete system $G^\circ(B) \triangleq B^\circ(B)/A^\circ(B)$ induced by the BPF expansions has impulse response function:

$$G_k = \frac{m}{T}\int_0^\infty g(t)\cdot\psi_k(t)\cdot dt \qquad (k=1,2,\cdots) \tag{29}$$

By taking the norms of both sides of (29), and using (28), we have:

$$\begin{aligned} \|G_k\|_\infty &\le \frac{m}{T}\int_0^\infty \|g(t)\|_\infty \cdot \|\psi_k(t)\|_\infty \cdot dt \\ &= \frac{m}{T}\,\|g(t)\|_\infty \cdot \int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} dt \\ &= \|g(t)\|_\infty < \infty \end{aligned} \tag{30}$$

which implies that $\{G_k\}$ is also bounded, and, thus, $G^\circ(B)$ stable in the sense of Lyapunov. As a consequence $A^\circ(B)$ is minimum phase (Chen, 1984).

(a3) Now assume that $a^\circ(D)$ is nonminimum phase. Then the system $g^\circ(D)$ is not stable in the sense of Lyapunov, and its impulse response function grows unbounded ($\|g(t)\|_\infty = \infty$). By using the continuity of $g(t)$ implied by the fact that $g^\circ(D)$ is strictly proper [see (4a), (4b)], the unboundedness of $\{g(t)\}$ implies that:

$$\begin{aligned} &\forall\, M>0 \quad \exists \quad \text{at least one interval } \Delta \text{ such that } \Delta \subseteq \left[(k-1)\frac{T}{m}, k\frac{T}{m}\right) \text{ for some value of } k, \\ &\qquad\qquad \text{and for which} \quad |g(t)| > M \quad \forall\, t \in \Delta \end{aligned} \tag{31}$$

[4]Notice that although $\|X_k\|_\infty$ and $\|x(t)\|_\infty$ designate different types of norms, the former may be interpreted within the context of the latter by defining $x'(t) = X_k$ for $(k-1)\frac{T}{m} \le t < k\frac{T}{m}$, so that: $\|X_k\|_\infty = \|x'(t)\|_\infty$.

The impulse response $\{G_k\}$ of the discrete system $G^o(B)$ induced by the BPF expansions will then be, for that particular value of k:

$$G_k = \frac{m}{T}\int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} g(t)\cdot dt \;>\; \frac{m}{T}\int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} M\cdot dt \;=\; M \qquad \Big(g(t)>0 \quad t\in\Delta\Big) \tag{32}$$

$$G_k = \frac{m}{T}\int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} g(t)\cdot dt \;<\; -\frac{m}{T}\int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} M\cdot dt \;=\; -M \qquad \Big(g(t)<0 \quad t\in\Delta\Big) \tag{33}$$

Based on this we conclude that:

$$\forall\, M>0 \quad \exists \quad \text{at least one } k \text{ such that: } \quad |G_k|>M \tag{34}$$

which implies that $\{G_k\}$ grows unbounded. As a consequence $G^o(B)$ is not stable in the sense of Lyapunov, and $A^o(B)$ is nonminimum phase.

(b) The fact that the above results hold for the polynomial $\tilde{C}^o(B)$ as well may be shown as follows: Rewrite the ARMAX system expression $\mathcal{S}_{\tilde{C}}$ [Eq.(6)] in terms of the Wiener process $\{\tilde{w}(t)\}$ in the following notational form:

$$\mathcal{S}_{\tilde{C}}: \qquad a^o(D)\cdot y(t) \;=\; b^o(D)\cdot u(t) \;+\; \tilde{c}^o(D)\cdot\tilde{w}(t) \tag{35}$$

with:

$$\tilde{c}^o(D) \triangleq D\cdot c^o(D) \;=\; c^o_{n_a-1}D^{n_a} + c^o_{n_a-2}D^{n_a-1} + \ldots + c^o_1 D^2 + c^o_0 D + 0 \tag{36}$$

By defining:

$$\tilde{\mathbf{c}} \triangleq \left[c^o_{n_a-1} \;\cdots c^o_1\; c^o_0 \mid 0\right]^T \tag{37}$$

one may readily show that the vectors $\tilde{\mathbf{c}}$ and $\tilde{\mathbf{C}}$ are related through the transformation expression:

$$\tilde{\mathbf{C}} = \mathbf{D}_A\cdot\tilde{\mathbf{c}} \tag{38}$$

which is of exactly the same form as the first of (23) that relates $\mathbf{A}$ and $\mathbf{a}$. As a consequence all previous results pertaining to $A^o(B)$ are fully applicable to the polynomial $\tilde{C}^o(B)$ as well.

Corollary 1: Lemma 1 and Theorem 2 imply that the *converse* of the latter holds for all autoregressive and moving average polynomials with coefficients in $\tilde{\mathcal{D}}$.

Corollary 2: An immediate consequence of Theorem 2 and assumption A2 pertaining to the strictly minimum phase nature of $a^o(D)$ is that the stochastic difference equation $\mathcal{S}_{\tilde{D}}$ is *stationary*.

The structure of the polynomial $\tilde{C}^o(B)$ is of particular importance in our developments, and therefore further discussed in the following theorem:

Theorem 3: The polynomial $\tilde{C}^o(B)$ of the stochastic difference equation (18) has a distinct root at $B = 1$ and can be factored as:

$$\tilde{C}^o(B) = (1 - B) \cdot C^o(B) \tag{39}$$

with $C^o(B)$ being minimum phase. □

Proof:

The fact that $\tilde{C}^o(B)$ has a root at $B = 1$ may be shown by using a known (Jiang and Schaufelberger, 1985) property stating that the sum of the A_i^o's is proportional to a_0^o, specifically: $\sum_{i=0}^{n_a} A_i^o = (T/m)^{n_a} a_0^o$. Because of (38), this is also applicable to $\tilde{C}^o(B)$, and therefore:

$$\sum_{i=0}^{n_a} \tilde{C}_i^o = \tilde{C}^o(B)|_{B=1} = (T/m)^{n_a} \tilde{c}_0^o = 0 \tag{40}$$

since $\tilde{c}_0^o = 0$ [see (37)]. The fact that the root $B = 1$ is distinct and $C^o(B)$ minimum phase, is a consequence of Theorem 2 which implies that $\tilde{C}^o(B)$ has to be minimum phase since $\tilde{c}^o(D)$ is such [based on the definition (36) and assumption A2]. □

6 STOCHASTIC MODELING OF THE DISCRETE ENDOGENEOUS EXCITATION SIGNAL $\{\tilde{W}_k\}$

For the development of an estimation method for the stochastic difference equation $\mathcal{S}_{\tilde{\mathcal{D}}}$ [Eq.(18)], the probabilistic properties of the endogeneous excitation signal $\{\tilde{W}_k\}$, defined as:

$$\tilde{W}_k \triangleq \frac{m}{T} \int_0^T \tilde{w}(t)\, \psi_k(t)\, dt \qquad k \in [1, m] \tag{41}$$

need to be also analyzed, and a proper stochastic representation of it developed. Due to the linearity of (41) and the Gaussianity assumption A3, $\{\tilde{W}_k\}$ will be Gaussian, and thus completely characterized by its first and second-order moments. The first-order moment of $\{\tilde{W}_k\}$ can be immediately verified to be zero:

$$E[\tilde{W}_k] = \frac{m}{T} \int_0^T E[\tilde{w}(t)] \cdot \psi_k(t) \cdot dt = 0 \qquad \forall\, k \in [1, m] \tag{42}$$

since the Wiener process is itself zero-mean. For the calculation of the second-order moment of $\{\tilde{W}_k\}$ we proceed as follows:

$$\begin{aligned} E[\tilde{W}_k \tilde{W}_l] &= \frac{m^2}{T^2} E\Big[\int_0^T \tilde{w}(s)\psi_k(s)ds \int_0^T \tilde{w}(t)\psi_l(t)dt\Big] \\ &= \frac{m^2(\sigma_w^o)^2}{T^2} \int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} \int_{(l-1)\frac{T}{m}}^{l\frac{T}{m}} \min(s,t) \cdot ds \cdot dt \end{aligned} \tag{43}$$

Consider the cases:

(a) Case $k = l$:

$$\begin{aligned} E[\tilde{W}_k^2] &= \frac{m^2(\sigma_w^\circ)^2}{T^2} \int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} \int_{(k-1)\frac{T}{m}}^{s} t \cdot dt \cdot ds + \frac{m^2(\sigma_w^\circ)^2}{T^2} \int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} \int_{(k-1)\frac{T}{m}}^{t} s \cdot ds \cdot dt \\ &\triangleq I + II \end{aligned} \tag{44}$$

where:

$$\begin{aligned} I &= \frac{m^2(\sigma_w^\circ)^2}{T^2} \int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} \left[\frac{1}{2}s^2 - \frac{1}{2}(k-1)^2\frac{T^2}{m^2}\right] ds \\ &= \frac{m^2(\sigma_w^\circ)^2}{T^2} \left[\frac{1}{6}k^3\frac{T^3}{m^3} - \frac{1}{2}(k-1)^2\frac{T^2}{m^2}k\frac{T}{m} - \frac{1}{6}(k-1)^3\frac{T^3}{m^3} + \frac{1}{2}(k-1)^2\frac{T^2}{m^2}(k-1)\frac{T}{m}\right] \\ &= \left[\frac{1}{2}k - \frac{1}{3}\right]\frac{T}{m}(\sigma_w^\circ)^2 \end{aligned} \tag{45}$$

By symmetry:

$$II = \left[\frac{1}{2}k - \frac{1}{3}\right]\frac{T}{m}(\sigma_w^\circ)^2 \tag{46}$$

and thus:

$$E[\tilde{W}_k^2] = \left(k - \frac{2}{3}\right) \cdot \frac{T}{m} \cdot (\sigma_w^\circ)^2 \tag{47}$$

(b) Case $k > l$:

$$\begin{aligned} E[\tilde{W}_k\tilde{W}_l] &= \frac{m^2(\sigma_w^\circ)^2}{T^2} \int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} \int_{(l-1)\frac{T}{m}}^{l\frac{T}{m}} t \cdot dt \cdot ds \\ &= \frac{m^2(\sigma_w^\circ)^2}{T^2} \left(\int_{(l-1)\frac{T}{m}}^{l\frac{T}{m}} tdt\right)\left(\int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} ds\right) \\ &= \left(l - \frac{1}{2}\right) \cdot \frac{T}{m} \cdot (\sigma_w^\circ)^2 \end{aligned} \tag{48}$$

(c) Case $k < l$:

$$\begin{aligned} E[\tilde{W}_k\tilde{W}_l] &= \frac{m^2(\sigma_w^\circ)^2}{T^2} \int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} \int_{(l-1)\frac{T}{m}}^{l\frac{T}{m}} s \cdot ds \cdot dt \\ &= \frac{m^2(\sigma_w^\circ)^2}{T^2} \left(\int_{(k-1)\frac{T}{m}}^{k\frac{T}{m}} sds\right)\left(\int_{(l-1)\frac{T}{m}}^{l\frac{T}{m}} dt\right) \\ &= \left(k - \frac{1}{2}\right) \cdot \frac{T}{m} \cdot (\sigma_w^\circ)^2 \end{aligned} \tag{49}$$

Based on (47)-(49) we then have the following expression for the autocovariance of $\{\tilde{W}_k\}$:

$$E[\tilde{W}_k\tilde{W}_l] = \begin{cases} (k - \frac{2}{3})\frac{T}{m}(\sigma_w^\circ)^2 & k = l \\ \left(\min(k,l) - \frac{1}{2}\right)\frac{T}{m}(\sigma_w^\circ)^2 & k \neq l \end{cases} \tag{50}$$

Clearly, since this autocovariance is not a function of the relative lag $k-l$, the sequence $\{\tilde{W}_k\}$ is nonstationary. In order to further investigate its structure we define a new sequence $\{Z_k\}$ as follows:

$$Z_k \triangleq \tilde{W}_k - \tilde{W}_{k-1} \tag{51}$$

$\{Z_k\}$ is, obviously, zero-mean as well, and with autocovariance $E[Z_k Z_l]$ that may be computed as follows:

(a) Case $k = l$:

$$\begin{aligned} E[Z_k^2] &= E[\tilde{W}_k^2] - 2E[\tilde{W}_k \tilde{W}_{k-1}] + E[\tilde{W}_{k-1}^2] \\ &= \left(k - \frac{2}{3}\right)\frac{T}{m}(\sigma_w^\circ)^2 - 2\left(k-1-\frac{1}{2}\right)\frac{T}{m}(\sigma_w^\circ)^2 + \left(k-1-\frac{2}{3}\right)\frac{T}{m}(\sigma_w^\circ)^2 = \frac{2}{3}\frac{T}{m}(\sigma_w^\circ)^2 \end{aligned} \tag{52}$$

(b) Case $k - l = 1$:

$$\begin{aligned} E[Z_k Z_{k-1}] &= E[\tilde{W}_k \tilde{W}_{k-1}] - E[\tilde{W}_{k-1}^2] - E[\tilde{W}_k \tilde{W}_{k-2}] + E[\tilde{W}_{k-1}\tilde{W}_{k-2}] \\ &= \left(k-1-\frac{1}{2}\right)\frac{T}{m}(\sigma_w^\circ)^2 - \left(k-1-\frac{2}{3}\right)\frac{T}{m}(\sigma_w^\circ)^2 - \left(k-2-\frac{1}{2}\right)\frac{T}{m}(\sigma_w^\circ)^2 + \left(k-2-\frac{1}{2}\right)\frac{T}{m}(\sigma_w^\circ)^2 \\ &= \frac{1}{6}\frac{T}{m}(\sigma_w^\circ)^2 \end{aligned} \tag{53}$$

(c) Case $k - l > 1$:

$$\begin{aligned} E[Z_k Z_l] &= E[\tilde{W}_k \tilde{W}_l] - E[\tilde{W}_{k-1}\tilde{W}_l] - E[\tilde{W}_k \tilde{W}_{l-1}] + E[\tilde{W}_{k-1}\tilde{W}_{l-1}] \\ &= \left(l-\frac{1}{2}\right)\frac{T}{m}(\sigma_w^\circ)^2 - \left(l-\frac{1}{2}\right)\frac{T}{m}(\sigma_w^\circ)^2 - \left(l-1-\frac{1}{2}\right)\frac{T}{m}(\sigma_w^\circ)^2 + \left(l-1-\frac{1}{2}\right)\frac{T}{m}(\sigma_w^\circ)^2 \\ &= 0 \end{aligned} \tag{54}$$

Observe that $E[Z_k Z_l]$ is not a function of k or l, but depends upon their difference $k-l$, and, by using the symmetry property of the autocovariance, we have:

$$E[Z_k Z_l] = \begin{cases} \frac{2}{3}\frac{T}{m}(\sigma_w^\circ)^2 & k-l=0 \\ \frac{1}{6}\frac{T}{m}(\sigma_w^\circ)^2 & |k-l|=1 \\ 0 & |k-l|\geq 2 \end{cases} \tag{55}$$

This implies that $\{Z_k\}$ is a stationary MA(1) process (Pandit and Wu, 1983), and may be thus modeled as:

$$Z_k = N_k + \theta_1^\circ N_{k-1} \tag{56}$$

with $\{N_k\}$ representing a discrete zero-mean Gaussian innovations (uncorrelated) sequence with variance $(\sigma_N^\circ)^2$. By comparing the autocovariance of the sequence $\{Z_k\}$ to that of a generic MA(1) process given as:

$$r_{zz}(k-l) = \begin{cases} (\sigma_N^\circ)^2[1+(\theta_1^\circ)^2] & k-l=0 \\ (\sigma_N^\circ)^2\theta_1^\circ & |k-l|=1 \\ 0 & |k-l|\geq 2 \end{cases} \tag{57}$$

we obtain the following parameters of an invertible MA(1) representation of $\{Z_k\}$:

$$\theta_1^\circ \cong 0.267949 \tag{58a}$$

$$(\sigma_N^\circ)^2 \cong 0.622008\frac{T}{m}(\sigma_w^\circ)^2 \tag{58b}$$

These results lead to the following important lemma:

<u>*Lemma 2:*</u> The BPF expansion series $\{\tilde{W}_k\}$ of the Wiener process $\{\tilde{w}(t)\}$ can be modeled as a nonstationary Integrated Moving Average IMA(1,1) process (Priestley, 1981) of the form:

$$(1-B)\cdot\tilde{W}_k = (1+\theta_1^\circ B)\cdot N_k \tag{59}$$

with θ_1° given by (58a) and $\{N_k\}$ being a Gaussian zero-mean innovations sequence with variance given by Eq.(58b).

7 THE ESTIMATION METHOD

By substituting the IMA(1,1) representation (59) of the discrete endogeneous excitation signal, and also the form of $\tilde{C}^\circ(B)$ given by (39), into the difference equation $\mathcal{S}_{\tilde{\mathcal{D}}}$ [Eq.(18)], we obtain the following stochastic difference equation:

$$\mathcal{S}_{\mathcal{D}}: \qquad A^\circ(B)\cdot Y_k = B^\circ(B)\cdot U_k + C^\circ(B)\cdot(1+\theta_1^\circ B)\cdot N_k \qquad N_k \sim \text{i.i.d.}\mathcal{N}(0,(\sigma_N^\circ)^2) \tag{60}$$

with i.i.d. standing for independently identically distributed. Since θ_1° is a-priori known, we may reexpress $\mathcal{S}_{\mathcal{D}}$ in terms of the filtered sequences:

$$U_k^F \triangleq (1+\theta_1^\circ B)^{-1}\cdot U_k \qquad Y_k^F \triangleq (1+\theta_1^\circ B)^{-1}\cdot Y_k \tag{61}$$

and by additionally normalizing the polynomials $A^\circ(B)$, $B^\circ(B)$, and $C^\circ(B)$ by dividing by $A_{n_a}^\circ$, we obtain the following normalized stochastic difference equation:

$$\mathcal{S}_{\mathcal{D}'}: \qquad A^{\circ'}(B)\cdot Y_k^F = B^{\circ'}(B)\cdot U_k^F + C^{\circ'}(B)\cdot N_k' \qquad N'_k \sim \text{i.i.d.}\mathcal{N}(0,(\sigma_{N'}^\circ)^2) \tag{62}$$

with:

$$A^{\circ'}(B) \triangleq 1 + A_1^{\circ'}B + \ldots + A_{n_a}^{\circ'}B^{n_a} \tag{63a}$$

$$\mathbf{A}' \triangleq [1\ A_1^{\circ'}\ \ldots\ A_{n_a}^{\circ'}]^T \triangleq \mathbf{A}/A_{n_a}^\circ \tag{63b}$$

$$B^{\circ'}(B) \triangleq B_0^{\circ'} + B_1^{\circ'}B + \ldots + B_{n_a}^{\circ'}B^{n_a} \tag{63c}$$

$$\mathbf{B}' \triangleq [B_0^{\circ'}\ B_1^{\circ'}\ \ldots\ B_{n_a}^{\circ'}]^T \triangleq \mathbf{B}/A_{n_a}^{\circ} \tag{63d}$$

$$C^{\circ'}(B) \triangleq 1 + C_1^{\circ'}B + \ldots + C_{n_a-1}^{\circ'}B^{n_a-1} \tag{63e}$$

$$\mathbf{C}' \triangleq [1\ C_1^{\circ'}\ \ldots\ C_{n_a-1}^{\circ'}]^T \triangleq \mathbf{C}/C_{n_a-1}^{\circ} \triangleq \mathbf{C}/\tilde{C}_{n_a}^{\circ} \tag{63f}$$

and:

$$N_k' \triangleq \frac{C_{n_a-1}^{\circ}}{A_{n_a}^{\circ}}N_k \triangleq \frac{\tilde{C}_{n_a}^{\circ}}{A_{n_a}^{\circ}}N_k \tag{64}$$

being a zero-mean innovations sequence with variance:

$$(\sigma_{N'}^{\circ})^2 = \left(C_{n_a-1}^{\circ}/A_{n_a}^{\circ}\right)^2 \cdot (\sigma_N^{\circ})^2 \tag{65}$$

The normalized stochastic difference equation $\mathcal{S}_{\mathcal{D}'}$ [Eq.(62)] with the stationary zero-mean and uncorrelated endogeneous excitation $\{N_k'\}$ can be now identified as a discrete-time ARMAX(n_a, n_a, n_a-1) model, which is:

1. Normalized by construction (the leading coefficients of the AR and MA polynomials are equal to unity),
2. Stationary, since $A^{\circ'}(B)$ is strictly minimum phase by virtue of corollary 2, which guarantees that $A^{\circ}(B)$ is strictly minimum phase,
3. Characterized by a minimum phase MA polynomial $C^{\circ'}(B)$ by virtue of assumption A2 and Theorem 3, which guarantee that $C^{\circ}(B)$ is minimum phase,

and is therefore *identifiable.* The proposed estimation method is based on both the identifiability of $\mathcal{S}_{\mathcal{D}'}$ and the bijective transformation nature of the mapping between the set of all systems of the form $\mathcal{S}_{\mathcal{D}'}$ and that of continuous-time ARMAX systems of the form $\mathcal{S}_{\mathcal{C}}$ [see Eq.(1)]. This latter property is formally given by the following theorem:

Theorem 4: The mapping $\mathcal{T}_a$ between the sets:

$$\mathcal{C}_a \triangleq \left\{(\mathbf{a}, \mathbf{b}, \mathbf{c}, (\sigma_w^{\circ})^2) \in \mathcal{R}^{n_a+1} \times \mathcal{R}^{n_a} \times \mathcal{R}^{n_a} \times \mathcal{R}^{+} \mid \text{with } \mathbf{a}, \mathbf{b}, \mathbf{c} \text{ of the form (21) with } a_{n_a}^{\circ} = c_{n_a-1}^{\circ} = 1\right\} \tag{66a}$$

and:

$$\mathcal{D}'_a \triangleq \Big\{ (\mathbf{A}',\mathbf{B}',\mathbf{C}',(\sigma^\circ_{N'})^2) \in \mathcal{R}^{n_a+1} \times \mathcal{R}^{n_a+1} \times \mathcal{R}^{n_a} \times \mathcal{R}^+ \mid \text{with } \mathbf{A}',\ \mathbf{B}',\ \mathbf{C}' \text{ generated through}$$

$$(23),\ (36),\ \text{and } (63)\ ,\ \text{and } (\sigma^\circ_{N'})^2 \text{ by } (58b),\ (65) \text{ by all } (\mathbf{a},\mathbf{b},\mathbf{c},(\sigma^\circ_w)^2) \in \mathcal{C}_a \Big\} \tag{66b}$$

is a bijective transformation. □

Proof:

First $\mathcal{T}_a$ is a transformation as for any 4-tuple $(\mathbf{a},\mathbf{b},\mathbf{c},(\sigma^\circ_w)^2) \in \mathcal{C}_a$ expressions (23), (36), (58), (61), (63), and (65) define a unique 4-tuple $(\mathbf{A}',\mathbf{B}',\mathbf{C}',(\sigma^\circ_{N'})^2) \in \mathcal{D}'_a$. $\mathcal{T}_a$ is also onto since, by construction of the set $\mathcal{D}'_a$, every 4-tuple $(\mathbf{A}',\mathbf{B}',\mathbf{C}',(\sigma^\circ_{N'})^2) \in \mathcal{D}'_a$ is the image of at least one 4-tuple $(\mathbf{a},\mathbf{b},\mathbf{c},(\sigma^\circ_w)^2) \in \mathcal{C}_a$.

The fact that $\mathcal{T}_a$ is one-to-one may be shown as follows: Assume a 4-tuple $(\mathbf{A}',\mathbf{B}',\mathbf{C}',(\sigma^\circ_{N'})^2) \in \mathcal{D}'_a$. A corresponding 4-tuple $(\mathbf{a},\mathbf{b},\mathbf{c},(\sigma^\circ_w)^2) \in \mathcal{C}_a$ may be then computed by the following sequence of operations:

$$\underline{\mathbf{a}} \triangleq \mathbf{D_A}^{-1} \cdot \mathbf{A}' \tag{67a}$$

$$\mathbf{a} = \underline{\mathbf{a}}/\underline{a}^\circ_{n_a} \tag{67b}$$

$$\mathbf{A} = \mathbf{D_A} \cdot \mathbf{a} \tag{68a}$$

$$\mathbf{b} = \boldsymbol{\Delta}_B^{-1} \cdot \mathbf{B}'_r \cdot A^\circ_{n_a} \tag{68b}$$

$$\tilde{C}^{\circ\prime}(B) = (1-B) \cdot C^{\circ\prime}(B) \tag{69a}$$

$$\underline{\mathbf{c}} \triangleq \boldsymbol{\Delta}_\mathbf{C}^{-1} \cdot \tilde{\mathbf{C}}'_r \tag{69b}$$

$$\mathbf{c} = \underline{\mathbf{c}}/\underline{c}^\circ_{n_a-1} \tag{69c}$$

$$\tilde{\mathbf{C}} = \mathbf{D_C} \cdot \mathbf{c} \tag{70a}$$

$$(\sigma^\circ_w)^2 = \frac{m}{T} \times \frac{1}{0.622008} \times \left(\frac{A^\circ_{n_a}}{\tilde{C}^\circ_{n_a}}\right)^2 \times (\sigma^\circ_{N'})^2 \tag{70b}$$

where $\mathbf{B}'_r$ and $\tilde{\mathbf{C}}'_r$ represent arbitrary n_a-dimensional subvectors of $\mathbf{B}'$ and $\tilde{\mathbf{C}}'$, respectively, $\Delta_{\mathbf{B}}$, $\Delta_{\mathbf{C}}$, appropriate $n_a \times n_a$ submatrices of $\mathbf{D_B}$ and $\mathbf{D_C}$, respectively, and $\underline{\mathrm{a}} \triangleq [\underline{\mathrm{a}}^{\circ}_{n_a} \cdots \underline{\mathrm{a}}^{\circ}_0]^T$, $\underline{\mathrm{c}} \triangleq [\underline{\mathrm{c}}^{\circ}_{n_a-1} \cdots \underline{\mathrm{c}}^{\circ}_0]^T$ intermediate (unnormalized) parameter vectors. Due to the nature of these expressions and the full rank property of $\mathbf{D_A}$, $\mathbf{D_B}$, and $\mathbf{D_C}$ (and thus of $\Delta_{\mathbf{B}}$ and $\Delta_{\mathbf{C}}$), the 4-tuple $(\mathbf{a}, \mathbf{b}, \mathbf{c}, (\sigma^{\circ}_w)^2)$ thus determined is indeed unique. □

The proposed estimation method is thus composed of two main stages: In the first stage a discrete-time ARMAX$(n_a, n_a, n_a - 1)$ model of the form [compare with $\mathcal{S}_{\mathcal{D}'}$ given by Eq.(62)]:

$$\mathcal{M}_{\mathcal{D}'}: \qquad A'(B) \cdot Y_k^F = B'(B) \cdot U_k^F + C'(B) \cdot E'_k \tag{71}$$

with $\{E'_k\}$ representing the model's one-step-ahead prediction error sequence, is estimated based on the available sequences $\{U_k^F\}_{k=1}^m$ and $\{Y_k^F\}_{k=1}^m$, whereas, in the second, the parameter estimates of the continuous-time ARMAX system $\mathcal{S}_{\mathcal{C}}$ are obtained through expressions similar to (67)-(70). Due to inevitable estimation errors, however, the actually obtained model $\mathcal{M}_{\mathcal{D}'}$ may not belong to the set $\mathcal{D}'_a$ ($\mathcal{M}_{\mathcal{D}'} \notin \mathcal{D}'_a$), in which case it will have no image within the set $\mathcal{C}'_a$ of continuous-time ARMAX systems. This problem may be dealt with by assigning to $\mathcal{M}_{\mathcal{D}'}$ that continuous-time ARMAX model $\mathcal{M}_{\mathcal{C}} \in \mathcal{C}'_a$ whose image in $\mathcal{D}'_a$ is, in some appropriate sense, closest to that of the estimated $\mathcal{M}_{\mathcal{D}'}$. Within the context of this work we have chosen to achieve that by using the Moore-Penrose pseudo-inverse, and the parameters of the continuous-time ARMAX process may be then estimated as follows [compare with (67)-(70)]:

<u>For the estimation of the AR parameter vector $\mathbf{a}^{\circ}$:</u>

$$\hat{\underline{\mathrm{a}}} \triangleq \mathbf{D}_A^{-1}\, \widehat{\mathbf{A}'} \tag{72a}$$

$$\hat{\mathbf{a}} = \hat{\underline{\mathrm{a}}} / \hat{a}_{n_a} \tag{72b}$$

where $\widehat{\mathbf{A}'}$ represents the estimate of $\mathbf{A}^{\circ\prime}$ [the coefficients of the polynomial $A^{\circ\prime}(B)$], $\hat{\underline{\mathrm{a}}} \triangleq [\hat{\underline{\mathrm{a}}}_{n_a} \dots \hat{\underline{\mathrm{a}}}_1 \hat{\underline{\mathrm{a}}}_0]^T$ an intermediate parameter vector, and $\hat{\mathbf{a}} \triangleq [\hat{a}_{n_a} \dots \hat{a}_1 \hat{a}_0]^T$ the vector of the final AR parameter estimates [the estimates of the coefficients of the polynomial $a^{\circ}(D)$]. □

<u>For the estimation of the X parameter vector $\mathbf{b}^{\circ}$:</u>

$$\hat{\mathbf{A}} \triangleq \mathbf{D}_A\, \hat{\mathbf{a}} \tag{73a}$$

$$\hat{\mathbf{b}} = (\mathbf{D}_B^T \mathbf{D}_B)^{-1}\, \mathbf{D}_B^T \widehat{\mathbf{B}'}\, \hat{A}_{n_a} \tag{73b}$$

where $\widehat{\mathbf{A}}$ is the vector of the estimated coefficients of $A^o(B)$, $\widehat{\mathbf{B'}}$ the vector of the estimated coefficients of $B^{o\prime}(B)$, and $\widehat{\mathbf{b}}$ the vector of the final X parameter estimates [the estimates of the coefficients of the polynomial $b^o(D)$]. □

For the estimation of the MA parameter vector $\mathbf{c}^o$:

$$\widehat{\tilde{C}'}(B) \triangleq (1-B)\cdot\widehat{C'}(B) \tag{74a}$$

$$\underline{\hat{\mathbf{c}}} \triangleq (\mathbf{D}_C^T\mathbf{D}_C)^{-1}\,\mathbf{D}_C^T\widehat{\tilde{\mathbf{C}}'} \tag{74b}$$

$$\hat{\mathbf{c}} = \underline{\hat{\mathbf{c}}}/\underline{\hat{c}}_{n_a-1} \tag{74c}$$

where $\widehat{\tilde{C}'}(B)$ is the estimate of $\tilde{C}(B)/\tilde{C}_{n_a}$, $\widehat{\tilde{\mathbf{C}}'}$ the vector of the coefficients of $\widehat{\tilde{C}'}(B)$, $\widehat{C'}(B)$ the estimate of $C'(B)$, $\underline{\hat{\mathbf{c}}}$ an intermediate estimate, and $\hat{\mathbf{c}} = [\hat{c}_{n_c}\ldots\hat{c}_1\hat{c}_0]^T$ the vector of the final MA parameter estimates [the estimates of the coefficients of the polynomial $c^o(D)$].

For the estimation of spectral height $(\sigma_w^o)^2$:

$$\widehat{\tilde{\mathbf{C}}} = \mathbf{D}_C\hat{\mathbf{c}} \tag{75a}$$

$$\widehat{\sigma}_w^2 = \frac{m}{T}\frac{1}{0.622008}\left(\frac{\widehat{A}_{n_a}}{\widehat{\tilde{C}}_{n_a}}\right)^2\left(\hat{\sigma}_{N'}\right)^2 \tag{75b}$$

Summary of the Estimation Method:

Based on the above, the proposed estimation method may be summarized as follows:

Step 1: Obtain the BPF spectral representations $\{U_k\}_{k=1}^m$ and $\{Y_k\}_{k=1}^m$ of the exogeneous excitation $\{u(t)\}$ and corresponding response $\{y(t)\}$ signals, respectively, by using the operation (12).
Step 2: Obtain the filtered representations $\{U_k^F\}$ and $\{Y_k^F\}$ by using expressions (61).
Step 3: Fit discrete ARMAX(n_a, n_a, n_a-1) models of the form (71) to the above filtered representations for successively higher values of n_a, by using Maximum Likelihood estimation. In each case compute the Bayesian Information Criterion (Akaike, 1977):

$$BIC(n_\theta) = \log(\hat{\sigma}_{N'})^2 + \frac{n_\theta}{2}\frac{\log m}{m} \tag{76}$$

where n_θ denotes the total number of estimated parameters, $(\hat{\sigma}_{N'})^2$ the estimated variance of the discrete innovations $\{N_k'\}$, and m the length of the BPF spectral representations used. The model that yields the smallest BIC is selected as best.

Step 4: Obtain the estimates of the continuous-time ARMAX system parameters through expressions (72)-(75).

Remarks:

The following remarks are finally in order:

(a) *Existence and Uniqueness of the Estimates* $\widehat{\mathbf{a}}$, $\widehat{\mathbf{b}}$, $\widehat{\mathbf{c}}$, and $\widehat{\sigma}_w^2$.

The existence and uniqueness of the continuous-time ARMAX system parameter estimates is a consequence of Theorem 4 and the full rank property of $\mathbf{D}_B$ and $\mathbf{D}_C$, which, in turn, guarantees (Barnett, 1990) the existence and uniqueness of their Moore-Penrose pseudo-inverses used in Eqs.(73b) and (74b), respectively.

(b) *Consistency of the Parameter Estimates.*

As is well-known, the Maximum Likelihood, or, in fact, the Prediction Error, estimator of the parameters of a discrete ARMAX system is, under mild assumptions, consistent (Ljung, 1987), and, therefore, such is the estimator of $\mathcal{S}_{\mathcal{D}'}$ [Eq.(62)]. Hence, asymptotically ($m \rightarrow \infty$), the estimated model $\mathcal{M}_{\mathcal{D}'} \in \mathcal{D}'_a$ (in probability), and since the parameters of the continuous-time ARMAX system are obtained as rational functions of the former [see Eqs.(72)-(75)], the application of Slutsky's lemma (Cramer, 1946) implies that the estimator:

$$\hat{\theta} \triangleq [\ \hat{\mathbf{a}}^T\ \hat{\mathbf{b}}^T\ \hat{\mathbf{c}}^T\ \hat{\sigma_w^2}\ ]^T \tag{77}$$

will be also consistent, that is:

$$\hat{\theta} \xrightarrow{prob} \theta^\circ \quad \text{as} \quad m \rightarrow \infty \tag{78}$$

where θ° represents the true parameter vector and m the length of the BPF spectral representations used. □

8 SIMULATION RESULTS AND DISCUSSION

In this section the performance characteristics of the proposed method are evaluated via numerical simulations. For a given continuous-time ARMAX system of the form (3) and specified exogeneous excitation $\{u(t)\}$, the response signal $\{y(t)\}$ is calculated by integrating each one of the following differential equations:

$$a^\circ(D) \cdot y_1(t) = b^\circ(D) \cdot u(t) \tag{79a}$$

$$a^\circ(D) \cdot y_2(t) = c^\circ(D) \cdot w(t) \tag{79b}$$

and superimposing their solutions:

$$y(t) = y_1(t) + y_2(t) \tag{79c}$$

For the faithful computation of continuous-time responses the integration step Δt is selected such that $20 \leq \tau/\Delta t \leq 200$, with τ representing the smallest system period or time constant. Additional care, is, however, needed for the simulation of the stochastic differential equation (79b). Indeed, the continuous-time innovations $\{w(t)\}$ is approximated by discrete white noise, and the differential equation is integrated by using a fourth-order Runge-Kutta method. In order for the second-order characteristics of the resulting response to match, at the beginning and end of each integration step, those of the theoretical solution, the variance of the discrete white noise is selected equal to:

$$\sigma_D^2 = 3.6 \cdot (\sigma_w^\circ)^2/\Delta t \tag{80}$$

with $(\sigma_w^\circ)^2$ representing the desired spectral height of the continuous-time white noise (Riggs and Phillips, 1987).

In all cases examined, the exogeneous excitation signals are composed of trains of pulses with amplitudes forming a sequence of Gaussian independently identically distributed (i.i.d.) random variables with zero mean and unit variance, and the BPF expansions are computed from (12) by using Simpson's composite rule (Burden and Faires, 1985). The estimation of a discrete ARMAX model of the form (71) is then based on the computed BPF spectral records, and the estimated model is validated by examining its predictive ability and the characteristics of its residual (the one-step-ahead prediction error) sequence computed from Eq.(71) for the estimated parameter values. For a good model the latter must be uncorrelated, and this is judged by examining whether its normalized sample autocorrelation lies within the 95% confidence interval of $\pm 1.96/\sqrt{m}$ (Pandit and Wu, 1983).

Once an estimated discrete-time model is succesfully validated and accepted as an accurate system representation, the continuous-time ARMAX parameters are obtained through the expressions (72)-(75). Estimation accuracy is finally judged in terms of parametric error indices of the form:

$$E_p = \frac{\| \widehat{\theta} - \theta^\circ \|}{\| \theta^\circ \|} \times 100\% \tag{81}$$

with θ representing a selected parameter vector and $\| \cdot \|$ Euclidean norm.

Estimation Results

The estimation of the underdamped ARMAX(2,1,1) system (System A):

$$(D^2 + 2D + 16) \cdot y(t) = (D + 10) \cdot x(t) + (D + 9) \cdot w(t) \tag{82}$$

is considered first, based on data records that are 100 *secs* long ($T = 100$ *secs*) and generated with white noise sequences having spectral heights $(\sigma_w^\circ)^2 = 0.005$ and $(\sigma_w^\circ)^2 = 0.01$. The integration step and BPF width were selected equal to $T/m = 10\Delta t$ and $\Delta t = 0.01$ *secs*, respectively. In both of the considered cases discrete ARMAX(2,2,1) models were estimated as statistically adequate, and, as the results of Figure 1, that depicts the normalized sample autocorrelation of the residuals lying within the 95% confidence interval of $\pm 1.96/\sqrt{m}$, indicate, were successfully validated. From these models the continuous-time ARMAX system parameters were obtained, and the final estimation results are summarized in Table 2, with Figure 2 also comparing the estimated frequency response characteristics of both the $b(D)/a(D)$ and $c(D)/a(D)$ transfer functions to their theoretical counterparts. As these results demonstrate excellent accuracy is achieved, with parametric percentage errors confined to reasonably small values. Similar remarks may be made from Table 3, that presents the results of a Monte Carlo analysis of the method based on twenty data records generated by different seed numbers and $(\sigma_w^\circ)^2 = 0.005$.

Next, the estimation of the overdamped ARMAX(2,1,1) system (System B) described by the equation:

$$(D^2 + 3D + 2) \cdot y(t) = x(t) + (D + 10) \cdot w(t) \tag{83}$$

is considered based on data records that are 100 *secs* long and generated with white noise sequences having spectral heights $(\sigma_w^\circ)^2 = 0.00005$ and $(\sigma_w^\circ)^2 = 0.0001$. The integration step and BPF width were selected as in the previous example, that is $T/m = 10\Delta t$ and $\Delta t = 0.01$ *secs*. In both of the considered cases discrete ARMAX(2,2,1) models were estimated and successfully validated (Figure 3), from which the final estimation results presented in Table 4 were obtained. The achievable accuracy is very good, and the estimated frequency response characteristics nicely match the corresponding theoretical curves (Figure 4). A Monte Carlo analysis based on twenty data records and $(\sigma_w^\circ)^2 = 0.00005$ (Table 5) further confirms the excellent characteristics of the method.

In this final case the ARMAX(3,2,2) system (System C):

$$(D^3 + 11D^2 + 424D + 1200) \cdot y(t) = (2D^2 + 60D + 800) \cdot x(t) + (D^2 + 12D + 225) \cdot w(t) \tag{84}$$

is considered based on data records that are 37.5 *secs* long and generated with white noise sequences having spectral heights $(\sigma_w^\circ)^2 = 0.005$ and $(\sigma_w^\circ)^2 = 0.01$. The integration step and BPF width were selected as $T/m = 8\Delta t$ and $\Delta t = 3.125 \times 10^{-3}$ *secs*, respectively. Discrete ARMAX(3,3,2) models were estimated as adequate (Table 6) and successfully validated (Figure 5). The final estimation results, corresponding frequency response characteristics, and Monte Carlo analysis of the method [$(\sigma_w^\circ)^2 = 0.005$], are presented in Table 7, Figure 6, and Table 8, respectively, from which very good accuracy is, once again, observed.

9 CONCLUSION

In this paper a novel and effective Maximum Likelihood type method for the estimation of continuous-time stochastic ARMAX systems from analog data records was introduced. This method is based on block-pulse function spectral representations, through which the problem is transformed into that of estimating the parameters of an induced stochastic difference equation subject to endogeneous and exogeneous excitations. The study of the structural and probabilistic properties of this equation was shown to further reduce the problem into that of estimating a special-form and identifiable discrete ARMAX system from spectral data. The method was then based on a number of key properties that this discrete ARMAX system was shown to possess, including stationarity, invertibility, and the bijective transformation nature of its mapping relationship with the original continuous-time system.

Among the unique features and advantages of the proposed estimation method, that make it especially attractive for applications, are: (a) The fact that neither estimates of signal derivatives, nor direct discretizations that lead to asymptotic bias errors, are used, (b) no prefilters or a-priori information regarding the system dynamics is required, (c) the data are allowed to be in *analog* form; a fact that, apart from its obvious significance, also implies that non-uniformly sampled and/or missing data can be also accomodated if necessary, and, (d) the relationship between the discrete and the original continuous-time system parameters is *linear*, so that sensitivity problems associated with highly nonlinear transformations are circumvented, the computational complexity is alleviated, and a-priori system knowledge can be readily incorporated into the estimation procedure.

The effectiveness and excellent performance characteristics of the proposed method were finally verified via a number of numerical simulations.

ACKNOWLEDGEMENT

The financial support of the Indonesian Second University Development Project and the Whirlpool Corporation is gratefully acknowledged.

REFERENCES

Akaike, H., 1977, "On Entropy Maximization Principle", in *Applications of Statistics*, edited by Krishnaiah, P.R., North-Holland Publishing Company, Amsterdam, pp.27-41.

Åström, K.J., and Källström, C.G., 1973, "Application of System Identification Techniques to the Determination of Ship Dynamics", in the *Proceedings of the 3rd IFAC/IFORS Symposium on Identification and System Parameter Estimation*, Eykhoff, P., ed., North Holland, pp.415-424.

Bagchi, A., 1975, "Continuous Time System Identification with Unknown Noise Covariance" *Automatica*, Vol. 11, pp.533-536.

Balakrishnan, A.V., 1973, *Stochastic Differential Systems*, Lecture Notes in Economics and Mathematical Systems, Vol. 84, Springer-Verlag.

Barnett, S., 1990, *Matrices: Methods and Applications*, Clarendon Press, Oxford.

Ben Mrad, R., and Fassois, S.D., 1991, "Recursive Identification of Vibrating Structures from Noise-Corrupted Observations, - Part 1: Identification Approaches", *ASME Journal of Vibration and Acoustics*, in press.

Bergstrom, A.R., (ed.), 1976, *Statistical Inference in Continuous Time Economic Models*, North-Holland/American Elsevier.

Bergstrom, A.R., 1983, "Gaussian Estimation of Structural Parameters in Higher Order Continuous Time Dynamic Models," *Econometrica*, Vol. 51, pp.117-152.

Bharucha, B.H., and Kadota, T.T., 1970, "On the Representation of Continuous Parameter Processes by a Sequence of Random Variables," *IEEE Transactions on Information Theory*, Vol. IT-16, pp.139-141.

Burden, R.L., and Faires, J.D., 1985, *Numerical Analysis*, 3rd ed., PWS Publishers, Boston.

Chen, C.T., 1984, *Linear System Theory and Design*, Holt, Rinehart and Winston, New York.

Cramer, H., 1946, *Mathematical Methods of Statistics*, Princeton University Press.

Fassois, S.D., Eman, K.F., and Wu, S.M., 1990, "Sensitivity Analysis of the Discrete-to-Continuous Dynamic System Transformation," *ASME Journal of Dynamic Systems, Measurement, and Control*, Vol. 112, pp.1-9.

Jazwinski, A.H., 1970, *Stochastic Processes and Filtering Theory*, Academic Press.

Jiang, Z.H., and Schaufelberger, W., 1985, "A New Algorithm for Single-Input Single-Output System Identification via Block-Pulse Functions," *International Journal of Systems Science*, Vol. 16, pp.1559-1571.

Jones, R.H., 1981, "Fitting a Continuous Time Autoregression to Discrete Data," in *Applied Time Series Analysis II*, D.F. Findley, ed., Academic Press, New York, pp.651-682.

Kraus, F., and Schaufelberger, W., 1990, "Identification with Block Pulse Functions, Modulating Functions, and Differential Operators," *International Journal of Control*, Vol. 51, pp.931-942.

Kwong, C.P., and Chen, C.F., 1981, "The Convergence Properties of Block-Pulse Series," *International Journal of Systems Science*, Vol. 12, pp.745-751.

Lee, J.E., and Fassois, S.D., 1991, "Suboptimum Maximum Likelihood of Structural Parameters from Multiple Excitation Vibration Data", *ASME Journal of Vibration and Acoustics*, in press.

Ljung, L., 1987, *System Identification - Theory for the User*, Prentice Hall, Englewood Cliffs, N.J.

Moore, J.B., 1988, "Convergence of Continuous Time Stochastic ELS Parameter Estimation," *Stochastic Processes and Their Applications*, Vol. 27, pp.195-215.

Neumann, D., Isermann, R., and Nold, S., 1989, "Comparisons of Some Parameter Estimation Methods for Continuous-Time Models," *Proceedings of the 8th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, H.F. Chen, ed., Pergamon Press, pp.851-856.

Pandit, S.M., 1973, "Data-Dependent Systems: Modeling, Analysis, and Optimal Control via Time-Series", Ph.D. Dissertation, University of Wisconsin at Madison.

Pandit, S.M., and Wu, S.M., 1975, "Unique Estimates of the Parameters of a Continuous Stationary Stochastic Process," *Biometrika*, Vol. 62, pp.497-501.

Pandit, S.M., and Wu, S.M., 1983, *Time Series and System Analysis with Applications*, John Wiley and Sons, New York.

Phadke, M.S., and Wu, S.M., 1974, "Modeling of Continuous Stochastic Processes from Discrete Observations with Applications to Sunspot Data," *Journal of the American Statistical Association*, Vol. 69, pp.325-329.

Pham-Dinh-Tuan, 1977, "Estimation of Parameters of A Continuous Time Gaussian Stationary Process with Rational Spectral Density," *Biometrika*, Vol. 64., pp.385-399.

Priestley, M.B., 1981, *Spectral Analysis and Time Series*, Academic Press.

Riggs, T.L., Jr., and Phillips, C.L., 1987, "Modeling Continuous Noise Sources in Digital Simulations," *Simulation*, Vol. 48, pp.11-18.

Robinson, P.M., 1980, "Continuous Model Fitting from Discrete Data," in *Directions in Time Series*, D. Brillinger, G.C. Tiao, eds., Institute of Mathematical Statistics.

Sagara, S., and Zhao, Z.Y., 1990, "Numerical Integration Approach to On-line Identification of Continuous-Time Systems," *Automatica*, Vol. 26, pp.63-74.

Sargan, J.D., 1974, "Some Discrete Approximations to Continuous Time Stochastic Models," *Journal of the Royal Statistical Society*, Series B, Vol. 36, pp.74-90.

Solo, V., 1984, "Some Aspects of Continuous-Discrete Time Series Modelling," in *Time Series Analysis of Irregularly Observed Data*, E. Parzen, ed., Springer Verlag, pp.325-345.

Unbehauen, H., and Rao, G.P., 1987, *Identification of Continuous Systems*, North-Holland, Amsterdam.

Unbehauen, H., and Rao, G.P., 1990, "Continuous-time Approaches to System Identification -

A Survey," *Automatica*, Vol. 26, pp.23-35.

Van Schuppen, J.H., 1983, "Convergence Results for Continuous-Time Adaptive Stochastic Filtering Algorithms," *Journal of Mathematical Analysis and Applications*, Vol. 96, pp.209-225.

Wahlberg, B., 1989, "On the Identification of Continous Time Dynamical Systems," *Proceedings of the 8th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, H.F. Chen, ed., Pergamon Press, pp.435-440.

Wang, C.H., 1982, "Generalized Block-Pulse Operational Matrices and Their Applications to Operational Calculus," *International Journal of Control*, Vol. 36, pp.67-76.

Wu, S.M., 1977, "Dynamic Data System: A New Modeling Technique", *ASME Journal of Engineering for Industry*, Vol. 99, pp.708-714.

Wymer, C.R., 1972, "Econometric Estimation of Stochastic Differential Equation Systems," *Econometrica*, Vol. 40, pp.565-577.

Young, P., 1981, "Parameter Estimation for Continuous-Time Models - A Survey," *Automatica*, Vol. 17, pp.23-39.

Young, P., and Jakeman, A., 1980, "Refined Instrumental Variable Methods of Recursive Time-Series Analysis - Part III: Extensions," *International Journal of Control*, Vol. 31, pp.741-764.

APPENDIX: The operational matrix for integration of signals in the BPF representation

Consider the m-th order BPF expansion of the signal $\{y(t)\}$, as given by expression (9b). The k-fold integral of $\{y(t)\}$ may be then expressed as:

$$\underbrace{\int_0^t \cdots \int_0^t}_{k-fold} y(t') \cdot dt' \cong \mathbf{X}^T \mathbf{F}_k \Psi(t) \qquad \text{(A.1)}$$

where:

$$\mathbf{F}_k \triangleq \left(\frac{T}{m}\right)^k \frac{1}{(k+1)!} \begin{bmatrix} f_{k,1} & f_{k,2} & f_{k,3} & \cdots & f_{k,m} \\ 0 & f_{k,1} & f_{k,2} & \cdots & f_{k,m-1} \\ 0 & 0 & f_{k,1} & \cdots & f_{k,m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f_{k,1} \end{bmatrix} \qquad [m \times m] \qquad \text{(A.2)}$$

and:

$$\begin{aligned} f_{k,1} &= 1 \qquad (\forall k) \\ f_{k,i} &= i^{k+1} - 2\,(i-1)^{k+1} + (i-2)^{k+1} \qquad (i = 2, 3, \ldots, m) \end{aligned} \qquad \text{(A.3)}$$

The matrix $\mathbf{F}_k$ is called the k-th order operational matrix for integration (Wang, 1982). Finally note that the approximation in (A.1) is due to the truncation error associated with the m-th order BPF signal representation, and, due to Theorem 1, converges to zero as $m \to \infty$.

Table 1: Relationships between the coefficients of $\mathcal{S}_C$ [Eq.(1)] and $\mathcal{S}_{\bar{D}}$ [Eq.(18)].

1st-order system	$A_1 = a_1 + \frac{1}{2}\frac{T}{m}a_0$ $A_0 = -a_1 + \frac{1}{2}\frac{T}{m}a_0$
	$B_1 = \frac{1}{2}\frac{T}{m}b_0$ $B_0 = \frac{1}{2}\frac{T}{m}b_0$
	$\tilde{C}_1 = c_0$ $\tilde{C}_0 = -c_0$
2nd-order system	$A_2 = a_2 + \frac{1}{2}\frac{T}{m}a_1 + \frac{1}{6}(\frac{T}{m})^2 a_0$ $A_1 = -2a_2 + \frac{2}{3}(\frac{T}{m})^2 a_0$ $A_0 = a_2 - \frac{1}{2}\frac{T}{m}a_1 + \frac{1}{6}(\frac{T}{m})^2 a_0$
	$B_2 = \frac{1}{2}\frac{T}{m}b_1 + \frac{1}{6}(\frac{T}{m})^2 b_0$ $B_1 = \frac{2}{3}(\frac{T}{m})^2 b_0$ $B_0 = -\frac{1}{2}\frac{T}{m}b_1 + \frac{1}{6}(\frac{T}{m})^2 b_0$
	$\tilde{C}_2 = c_1 + \frac{1}{2}\frac{T}{m}c_0$ $\tilde{C}_1 = -2c_1$ $\tilde{C}_0 = c_1 - \frac{1}{2}\frac{T}{m}c_0$
3rd-order system	$A_3 = a_3 + \frac{1}{2}\frac{T}{m}a_2 + \frac{1}{6}(\frac{T}{m})^2 a_1 + \frac{1}{24}(\frac{T}{m})^3 a_0$ $A_2 = -3a_3 - \frac{1}{2}\frac{T}{m}a_2 + \frac{1}{2}(\frac{T}{m})^2 a_1 + \frac{11}{24}(\frac{T}{m})^3 a_0$ $A_1 = 3a_3 - \frac{1}{2}\frac{T}{m}a_2 - \frac{1}{2}(\frac{T}{m})^2 a_1 + \frac{11}{24}(\frac{T}{m})^3 a_0$ $A_0 = -a_3 + \frac{1}{2}\frac{T}{m}a_2 - \frac{1}{6}(\frac{T}{m})^2 a_1 + \frac{1}{24}(\frac{T}{m})^3 a_0$
	$B_3 = \frac{1}{2}\frac{T}{m}b_2 + \frac{1}{6}(\frac{T}{m})^2 b_1 + \frac{1}{24}(\frac{T}{m})^3 b_0$ $B_2 = -\frac{1}{2}\frac{T}{m}b_2 + \frac{1}{2}(\frac{T}{m})^2 b_1 + \frac{11}{24}(\frac{T}{m})^3 b_0$ $B_1 = -\frac{1}{2}\frac{T}{m}b_2 - \frac{1}{2}(\frac{T}{m})^2 b_1 + \frac{11}{24}(\frac{T}{m})^3 b_0$ $B_0 = \frac{1}{2}\frac{T}{m}b_2 - \frac{1}{6}(\frac{T}{m})^2 b_1 + \frac{1}{24}(\frac{T}{m})^3 b_0$
	$\tilde{C}_3 = c_2 + \frac{1}{2}\frac{T}{m}c_1 + \frac{1}{6}(\frac{T}{m})^2 c_0$ $\tilde{C}_2 = -3c_2 - \frac{1}{2}\frac{T}{m}c_1 + \frac{1}{2}(\frac{T}{m})^2 c_0$ $\tilde{C}_1 = 3c_2 - \frac{1}{2}\frac{T}{m}c_1 - \frac{1}{2}(\frac{T}{m})^2 c_0$ $\tilde{C}_0 = -c_2 + \frac{1}{2}\frac{T}{m}c_1 - \frac{1}{6}(\frac{T}{m})^2 c_0$
4th-order system	$A_4 = a_4 + \frac{1}{2}\frac{T}{m}a_3 + \frac{1}{6}(\frac{T}{m})^2 a_2 + \frac{1}{24}(\frac{T}{m})^3 a_1 + \frac{1}{120}(\frac{T}{m})^4 a_0$ $A_3 = -4a_4 - \frac{T}{m}a_3 + \frac{1}{3}(\frac{T}{m})^2 a_2 + \frac{5}{12}(\frac{T}{m})^3 a_1 + \frac{13}{60}(\frac{T}{m})^4 a_0$ $A_2 = 6a_4 - (\frac{T}{m})^2 a_2 + \frac{11}{20}(\frac{T}{m})^4 a_0$ $A_1 = -4a_4 + \frac{T}{m}a_3 + \frac{1}{3}(\frac{T}{m})^2 a_2 - \frac{5}{12}(\frac{T}{m})^3 a_1 + \frac{13}{60}(\frac{T}{m})^4 a_0$ $A_0 = a_4 - \frac{1}{2}\frac{T}{m}a_3 + \frac{1}{6}(\frac{T}{m})^2 a_2 - \frac{1}{24}(\frac{T}{m})^3 a_1 + \frac{1}{120}(\frac{T}{m})^4 a_0$
	$B_4 = \frac{1}{2}\frac{T}{m}b_3 + \frac{1}{6}(\frac{T}{m})^2 b_2 + \frac{1}{24}T^3 b_1 + \frac{1}{120}(\frac{T}{m})^4 b_0$ $B_3 = -\frac{T}{m}b_3 + \frac{1}{3}(\frac{T}{m})^2 b_2 + \frac{5}{12}T^3 b_1 + \frac{13}{60}(\frac{T}{m})^4 b_0$ $B_2 = -(\frac{T}{m})^2 b_2 + \frac{11}{20}(\frac{T}{m})^4 b_0$ $B_1 = +\frac{T}{m}b_3 + \frac{1}{3}(\frac{T}{m})^2 b_2 - \frac{5}{12}T^3 b_1 + \frac{13}{60}(\frac{T}{m})^4 b_0$ $B_0 = -\frac{1}{2}\frac{T}{m}b_3 + \frac{1}{6}(\frac{T}{m})^2 b_2 - \frac{1}{24}T^3 b_1 + \frac{1}{120}(\frac{T}{m})^4 b_0$
	$\tilde{C}_4 = c_3 + \frac{1}{2}\frac{T}{m}c_2 + \frac{1}{6}(\frac{T}{m})^2 c_1 + \frac{1}{24}(\frac{T}{m})^3 c_0$ $\tilde{C}_3 = -4c_3 - \frac{T}{m}c_2 + \frac{1}{3}(\frac{T}{m})^2 c_1 + \frac{5}{12}(\frac{T}{m})^3 c_0$ $\tilde{C}_2 = 6c_3 - (\frac{T}{m})^2 c_1$ $\tilde{C}_1 = -4c_3 + \frac{T}{m}c_2 + \frac{1}{3}(\frac{T}{m})^2 c_1 - \frac{5}{12}(\frac{T}{m})^3 c_0$ $\tilde{C}_0 = c_3 - \frac{1}{2}\frac{T}{m}c_2 + \frac{1}{6}(\frac{T}{m})^2 c_1 - \frac{1}{24}(\frac{T}{m})^3 c_0$

Table 2: Estimation results for System A at two different noise powers.

	Process Parameters	Estimated Parameters [a]	
		$\sigma_w^2 = 0.005$	$\sigma_w^2 = 0.01$
a_2	1	1.0000	1.0000
a_1	2	2.0406	2.0256
a_0	16	16.1395	16.0939
b_1	1	1.0537	1.0623
b_0	10	10.1251	10.1007
c_1	1	1.0000	1.0000
c_0	9	8.7383	8.7138
$E_p^A(\%)$	—	0.8992	0.6026
$E_p^B(\%)$	—	1.3544	1.1783
$E_p^C(\%)$	—	2.8901	3.1605
σ_w^2	—	0.0112	0.0223

[a] For the simulation $\tau/\Delta t \cong 157$.

Table 3: Monte Carlo results for System A.

	Process Parameters	Estimated Parameters	
		Mean Value	Standard Deviation
a_1	2	2.0249	0.0833
a_0	16	16.0771	0.2961
b_1	1	1.0364	0.0293
b_0	10	10.0765	0.2234
c_0	9	8.4713	0.5563
$E_p^A(\%)$	—	0.5025	—
$E_p^B(\%)$	—	0.8426	—
$E_p^C(\%)$	—	5.8746	—
σ_w^2	0.005	0.01144	0.00066

Table 4: Estimation results for System B at two different noise powers.

	Process Parameters	Estimated Parameters [a]	
		$\sigma_w^2 = 0.00005$	$\sigma_w^2 = 0.0001$
a_2	1	1.0000	1.0000
a_1	3	3.0570	3.0618
a_0	2	2.0181	2.0194
b_1	0	0.0014	0.0021
b_0	1	1.0203	1.0217
c_1	1	1.0000	1.0000
c_0	10	9.6879	9.7009
$E_p^A(\%)$	—	1.5995	1.7324
$E_p^B(\%)$	—	2.0381	2.1774
$E_p^C(\%)$	—	3.1054	2.9757
σ_w^2	—	1.1084×10^{-4}	2.2154×10^{-4}

[a] For the simulation $\tau/\Delta t = 50$.

Table 5: Monte Carlo results for System B.

	Process Parameters	Estimated Parameters	
		Mean Value	Standard Deviation
a_1	3	2.94573	0.12459
a_0	2	2.00387	0.09451
b_1	0	0.00029	0.00312
b_0	1	1.00906	0.01792
c_0	10	9.24614	0.62962
$E_p^A(\%)$	—	1.50911	—
$E_p^B(\%)$	—	0.90632	—
$E_p^C(\%)$	—	7.53856	—
σ_w^2	5×10^{-5}	1.14245×10^{-4}	6.92103×10^{-6}

Table 6: Order determination results for System C ($\sigma_w^2 = 0.005$).

System order n_a	BIC
2	-8.16820
3	-8.77149
4	-8.77039

Table 7: Estimation results for System C at two different noise powers.

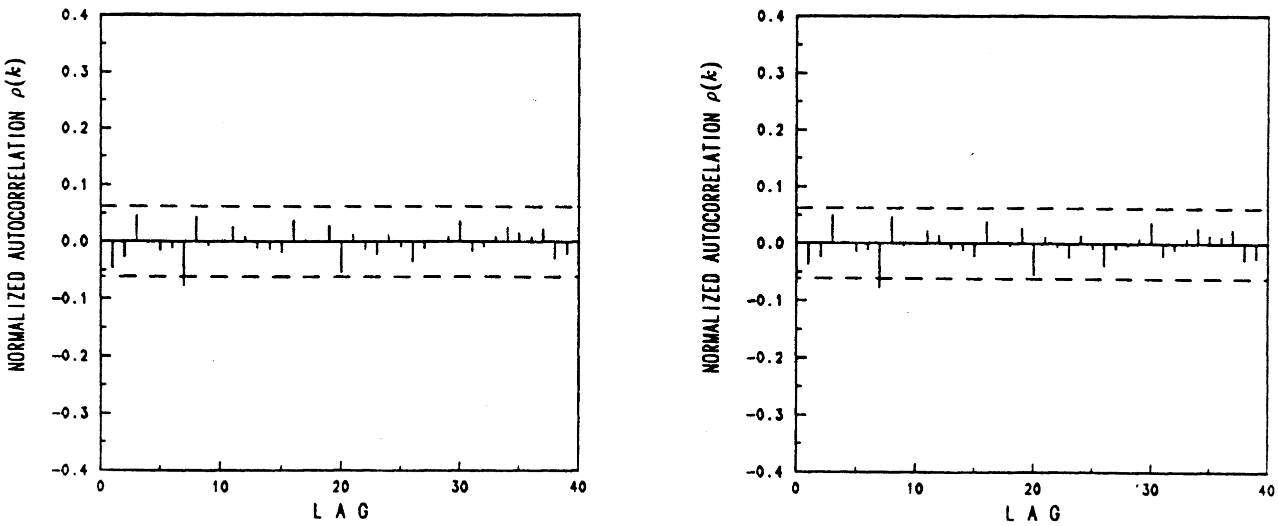
	Process Parameters	Estimated Parameters [a]	
		$\sigma_w^2 = 0.005$	$\sigma_w^2 = 0.01$
a_3	1	1.0000	1.0000
a_2	11	11.7901	11.9525
a_1	424	433.4398	434.9380
a_0	1200	1268.5162	1304.6710
b_2	2	2.0402	2.0266
b_1	60	64.3514	65.2344
b_0	800	822.3662	827.7895
c_0	1	1.0000	1.0000
c_1	12	12.7478	12.8618
c_0	225	223.1065	224.7139
$E_p^A(\%)$	—	5.4345	8.2691
$E_p^B(\%)$	—	2.8402	3.5249
$E_p^C(\%)$	—	0.9035	0.4030
σ_w^2	—	0.0103	0.0205

[a] For the simulation $\tau/\Delta t \cong 100$.

Table 8: Monte Carlo results for System C.

	Process Parameters	Estimated Parameters	
		Mean Value	Standard Deviation
a_2	11	11.3989	0.2347
a_1	424	431.7177	4.7444
a_0	1200	1225.5114	67.7825
b_2	2	2.0780	0.0281
b_1	60	62.0788	1.5323
b_0	800	824.2760	23.5585
c_1	12	12.6146	1.1445
c_0	225	232.4035	13.7799
$E_p^A(\%)$	—	2.0944	—
$E_p^B(\%)$	—	3.0371	—
$E_p^C(\%)$	—	3.2970	—
σ_w^2	0.005	0.01023	0.00045

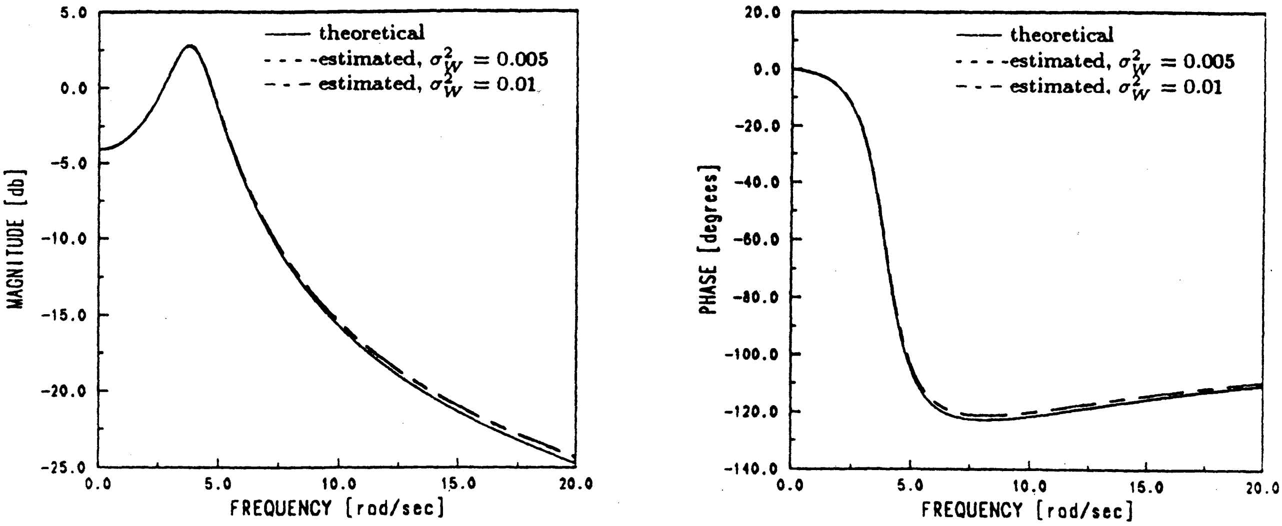
LIST OF FIGURES



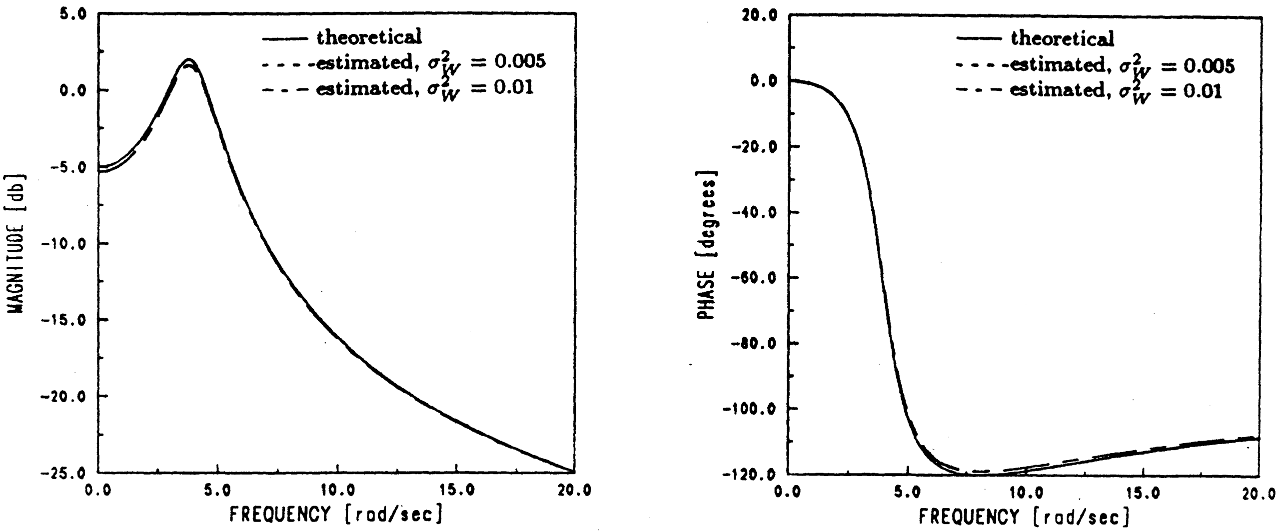
(a) $\sigma_W^2 = 0.005$

(b) $\sigma_W^2 = 0.01$

Figure 1: The normalized sample autocorrelation of the discrete residual sequence (System A).

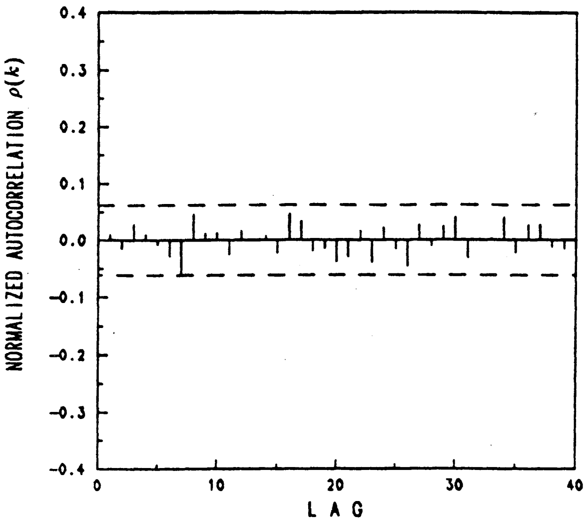


(a) Transfer Function $b(D)/a(D)$

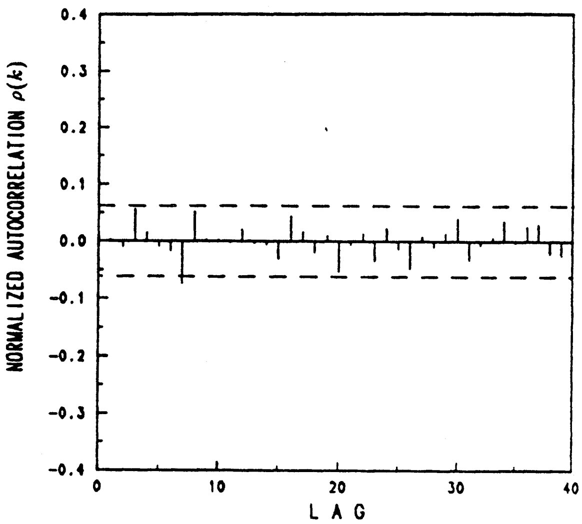


(b) Transfer Function $c(D)/a(D)$

Figure 2: Frequency response curves of the estimated continuous-time system (System A).

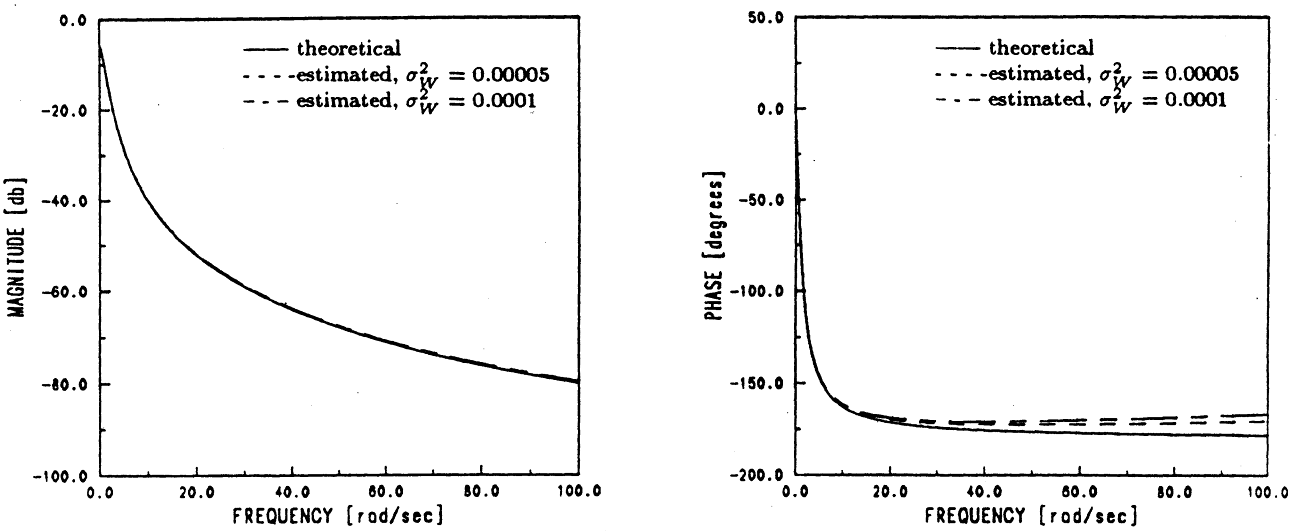


(a) $\sigma_W^2 = 0.00005$

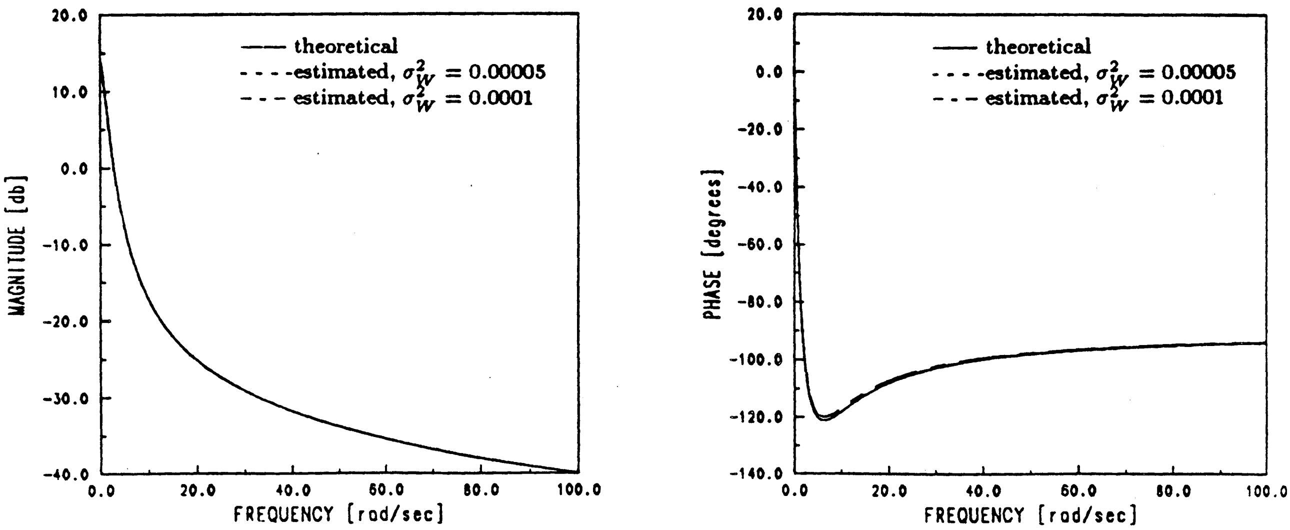


(b) $\sigma_W^2 = 0.0001$

Figure 3: The normalized sample autocorrelation of the discrete residual sequence (System B).

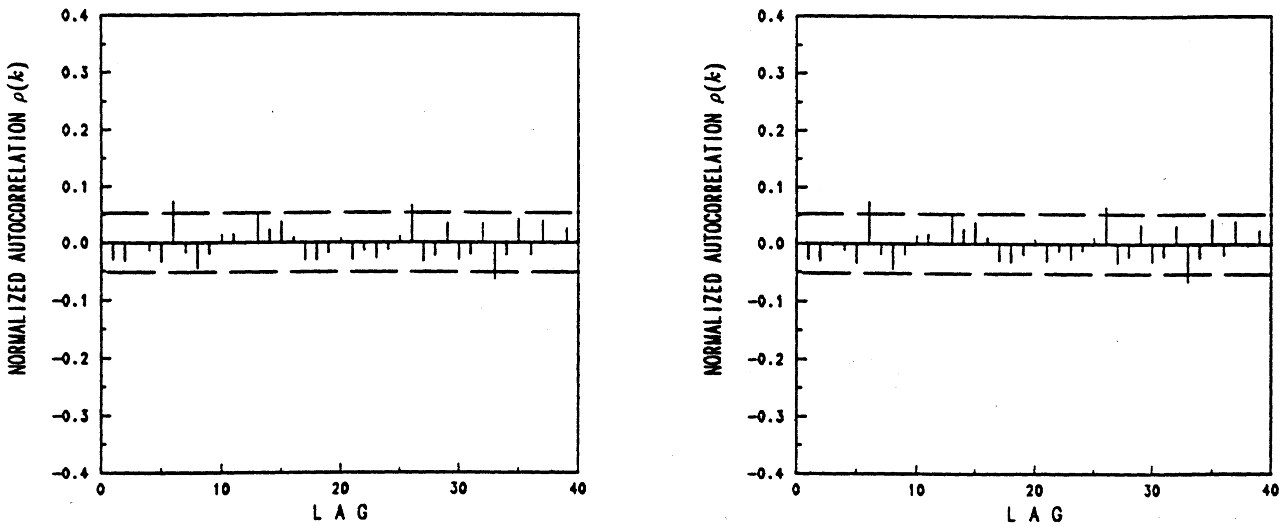


(a) Transfer Function $b(D)/a(D)$



(b) Transfer Function $c(D)/a(D)$

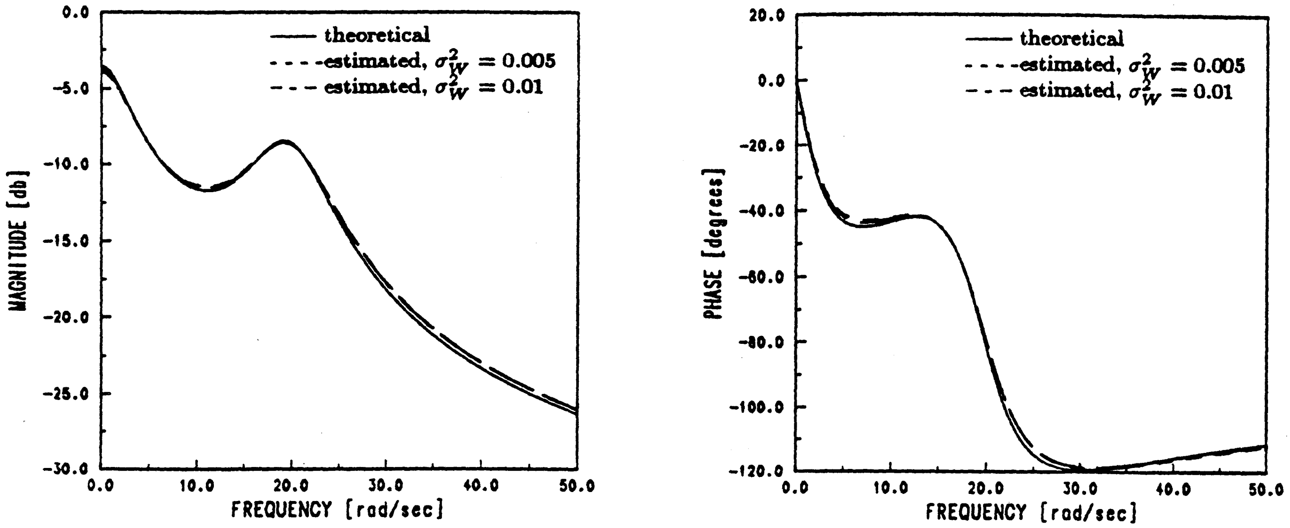
Figure 4: Frequency response curves of the estimated continuous-time system (System B).



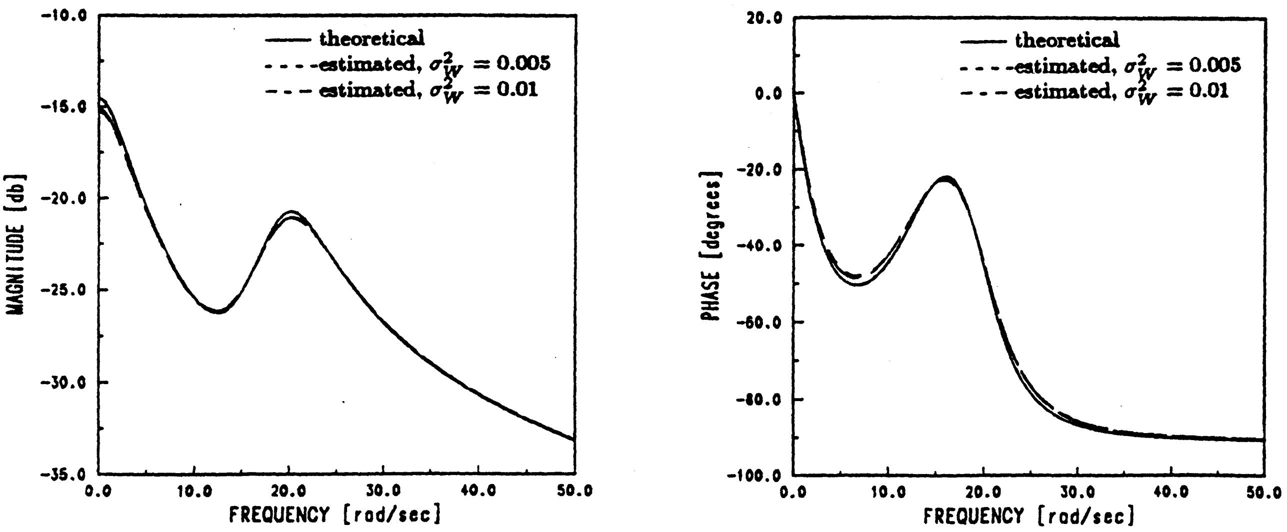
(a) $\sigma_W^2 = 0.005$

(b) $\sigma_W^2 = 0.01$

Figure 5: The normalized sample autocorrelation of the discrete residual sequence (System C).



(a) Transfer Function $b(D)/a(D)$



(b) Transfer Function $c(D)/a(D)$

Figure 6: Frequency response curves of the estimated continuous-time system (System C).