Journal of Sound and Vibration (1993) **167**(3), 481–509

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

P. NURPRASETIO AND S. D. FASSOIS

Department of Mechanical Engineering and Applied Mechanics, The University of Michigan, Ann Arbor, Michigan 48109-2121, *U.S.A.*

(*Received* 10 *July* 1991, *and in final form* 18 *May* 1992)

In this paper a novel and effective maximum likelihood type method for the estimation of physically meaningful *continuous-time stochastic* systems from *analog* data records is introduced. The method utilizes the ARMAX canonical form and block-pulse function spectral representations, through which the problem is shown to be transformed into that of estimating an induced and special-form discrete stochastic system from spectral data. The proposed method is based on a number of key structural and probabilistic properties that this discrete system is shown to possess, including stationarity, invertibility, and the bijective transformation nature of its mapping relationship with the original continuous-time system.

Unlike previous schemes, the proposed method utilizes *analog* data without depending upon estimates of signal derivatives or prefilters, avoids errors due to direct discretizations associated with instantaneous sampling, and is characterized by a *linear* transformation relationship between the discrete and the original continuous-time system parameters. This leads to additional important advantages, such as the elimination of sensitivity problems associated with highly non-linear mappings, the capability of incorporating *a priori* system information, and reduced computational complexity. The effectiveness of the method is verified via numerical experiments with a number of stochastic systems.

1. INTRODUCTION

After two decades of almost complete dominance of discrete-time system identification and parameter estimation approaches in engineering theory and practice, the relevance and importance of *continuous-time* methods based on *analog* data have been increasingly recognized [1–7]. A survey of the available literature, however, reveals that the overwhelming majority of currently available continuous-time methods are restricted to the deterministic case, and are inappropriate for the more general and practically important class of *stochastic* systems† with which additional difficulties are associated [1, 4].

One of the main difficulties in estimating continuous-time stochastic systems from analog data records is due to our inability accurately to compute time derivatives (of various orders) of the observed random signals [1, 4, 5]. If that were possible, continuous-time versions of many discrete methods could be constructed; see, for instance, the work of Balakrishnan [8], Bagchi [9] and Pham-Dinh-Tuan [10], who developed and analyzed

† 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 traveling ground vehicles and aircrafts.

0022–460X/93/240481 + 29 $08.00/0

the Maximum Likelihood method for this case and proved its asymptotic optimality. Also see Unbehauen and Rao [3] and Priestley [11], who discussed the autoregressive model case. The use of the so-called State Variable Filters (SVF's), that are extensively used in the deterministic case [1–3] for alleviating this difficulty, is neither trivial nor effective, and their proper selection is not obvious,† their incorporation affects the stochastic part of the model and, unlike the deterministic case, they cannot prevent the introduction of large errors into the obtained random signal derivatives [5].

Attempts to overcome these problems through alternative approaches based on integral (instead of differential) operators have been also considered. Van Schuppen [12] examined recursive forms of such estimation algorithms for continuous-time autoregressive processes and proved their convergence. Moore [13] analyzed the convergence of the continuous-time version of the Recursive Extended Least Squares (RELS) estimation algorithm for AutoRegressive Moving Average with eXogenous excitation (ARMAX) systems. However, in order to avoid numerical instability problems due to the pure integration of the observed signals, exponentially stable prefilters have to be 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 do not decay and can cause significant estimation bias errors [1, 6].

It therefore comes as no surprise that in almost all practical cases and applications 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 [1, 2, 5, 6, 14].

The only, and arithmetically very few, exceptions appear to be based on an approach that postulates the estimation of continuous-time stochastic models through the estimation of directly, and *approximately*, discretized representations associated with instantaneous data sampling. However, this approach is also restricted to *uniformly sampled* data records [15–17]. Moreover, very fast sampling cannot be used due to well-known limitations [18]. The method developed by Pandit and Wu [19–21] is a systematic and comprehensive such procedure that has been successfully used for the solution of a number of production engineering, random vibration, and other types of problems. Some similar methods have also, although independently, been developed in the fields of statistics [22–24] and econometrics [25–27], where the assumption of sampled data is practically reasonable.

The main advantage of all these methods is in their ability to utilize the already available machinery for discrete-time system estimation [11, 21, 28]. Yet, they are known to be characterized by a number of drawbacks and limitations, as follows. (a) They are restricted to operate on *uniformly sampled* data records and at *reasonably slow* sampling rates. (b) The estimates of the continuous-time parameters are asymptotically biased. The bias errors are due to the approximate nature of the instantaneous discretization employed ([1]; also see the discussion in Ben Mrad and Fassois [29] and Lee and Fassois [14]), and are also analyzed in the econometrics literature [25, 26, 30]. (c) Additional errors are introduced into the continuous-time parameter estimates due to sensitivity problems associated with the highly non-linear discrete-to-continuous transformation [18]; with the related question of sampling also being recognized as non-trivial [7, 18]. (d) The incorporation of *a priori* system information (that is often available in engineering applications) into the estimation procedure is very difficult, if not completely impossible, due to the highly complicated and non-linear nature of the discrete-to-continuous transformation [4, 20].

† It is in fact known that an optimal selection requires *a priori* system knowledge. To cope with that, adaptive SVF's have been also suggested [5].

(e) The required computational complexity is finally high, and further aggravated by the aforementioned non-linear transformation.

In this paper a novel *Maximum Likelihood type* method that overcomes the foregoing difficulties and limitations and estimates *continuous-time stochastic* systems from *analog* data records is introduced. The proposed method utilizes the general Autoregressive Moving Average with eXogenous excitation (ARMAX) stochastic model structure and block-pulse function (BPF) spectral representations. 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 [31] and Jiang and Schaufelberger [32]); the present work, however, appears to be the first that utilizes BPF expansions in solving stochastic estimation problems.

Within the context of this paper, the validity of BPF spectral representations for stochastic signals is formally justified first. Based upon them, as well as a set of linear operations motivated by recent developments in deterministic identification [32], 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 representation of exogenous (observable) and endogenous (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 endogenous 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 and 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 to be *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 transformation. The linearity of this transformation not only results in reduced computational complexity, but is also instrumental in circumventing sensitivity problems and allowing for the incorporation of frequently available *a priori* information into the estimation procedure. These characteristics, along with the use of analog data without requiring estimates of signal derivatives, and the elimination of direct discretization procedures associated with instantaneous sampling, are among the main features of the proposed method.

The remaining portion 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 section 5. The stochastic modeling of the endogenous 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. STATEMENT OF THE PROBLEM

Consider a continuous-time stochastic system with exogenous (observable) excitation $\{u(t)\}$, endogenous (unobservable) excitation $\{w(t)\}$ and response $\{y(t)\}$, described by the ARMAX stochastic differential equation:†

$$\mathscr{S}_{\mathscr{C}}: \quad \sum_{i=0}^{n_a} a_i^\circ \frac{\mathrm{d}^i y(t)}{\mathrm{d}t^i} = \sum_{j=0}^{n_a-1} b_j^\circ \frac{\mathrm{d}^j u(t)}{\mathrm{d}t^j} + \sum_{k=0}^{n_a-1} c_k^\circ \frac{\mathrm{d}^k w(t)}{\mathrm{d}t^k}, \tag{1}$$

in which the differentiation operations are to be interpreted in the mean-square sense [33] and $a_{n_a}^\circ = c_{n_a-1}^\circ = 1$. The endogenous excitation $\{w(t)\}$ is a zero-mean continuous-time innovations signal with autocovariance function:

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

where $\delta(\cdot)$ denotes the Dirac delta function.

By using the mean square differential operator D, the system $\mathscr{S}_{\mathscr{C}}$ may be equivalently written in notational form

$$\mathscr{S}_{\mathscr{C}}: \quad a^\circ(D)y(t) = b^\circ(D)u(t) + c^\circ(D)w(t), \tag{3}$$

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

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

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

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

The ARMAX system $\mathscr{S}_{\mathscr{C}}$ 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

$$\int_0^T [x(t)]^2 \,\mathrm{d}t < \infty$$

almost surely (a.s.) [33].

As we will see, the Gaussianity assumption is not crucial for our developments, but simply the vehicle for casting the problem into a Maximum Likelihood framework. If relaxed, the method may be still formally interpreted within the Prediction Error estimation context [28].

The estimation problem of interest may be then posed as follows: "Given analog excitation $\{u(t)\}$ and response $\{y(t)\}$ data records over a period of time $[0, T)$, generated by the system $\mathscr{S}_{\mathscr{C}}$ subject to assumptions A1–A3, estimate an ARMAX model of the form

$$\mathscr{M}_{\mathscr{C}}: \quad a(D)y(t) = b(D)u(t) + c(D)w(t), \qquad E[w(s)w(t)] = \sigma_w^2 \delta(s-t), \tag{5}$$

that matches $\mathscr{S}_{\mathscr{C}}$ as closely as possible."

† The superscript ° is used to indicate quantities associated with the actual system and to distinguish them from those of corresponding candidate models.

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 more convenient for our purposes, representation may be obtained by integrating the differential equation (1) n_a times which, assuming zero initial conditions, yields:

$$\mathcal{S}_{\bar{\mathscr{C}}}:\quad y(t)+a^{\circ}_{n_a-1}\int_0^t y(t')\,\mathrm{d}t'+\cdots+a^{\circ}_0\underbrace{\int_0^t\cdots\int_0^t}_{n_a\text{-fold}} y(t')\,\mathrm{d}t'$$

$$=b^{\circ}_{n_a-1}\int_0^t u(t')\,\mathrm{d}t'+\cdots+b^{\circ}_0\underbrace{\int_0^t\cdots\int_0^t}_{n_a\text{-fold}} u(t')\,\mathrm{d}t'$$

$$+\,\tilde{w}(t)+\cdots+c^{\circ}_0\underbrace{\int_0^t\cdots\int_0^t}_{(n_a-1)\text{-fold}}\tilde{w}(t')\,\mathrm{d}t'. \quad (6)$$

In this expression the integration operations are to be interpreted in the mean-square sense as well [33], while $\{w(t)\}$ represents the Wiener process

$$\tilde{w}(t)\triangleq\int_0^t w(s)\,\mathrm{d}s. \quad (7)$$

Due to the stated properties of $\{w(t)\}$, $\{\tilde{w}(t)\}$ is Gaussian, continuous in probability, with

$$\int_0^T[\tilde{w}(t)]^2\,\mathrm{d}t<\infty \qquad \text{(a.s.)},$$

and also zero-mean, and with autocovariance function [33]

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

For the development of the estimation method, the mth order block-pulse function (BPF) spectral representations [34] of the signals $\{u(t)\}$, $\{\tilde{w}(t)\}$ and $\{y(t)\}$ are needed. Based on them, the signals may be expressed as follows within the observation interval $[0, T)$:†

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

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

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

† Bold-face characters indicate vector/matrix quantities.

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 $\tilde{\mathbf{W}}$ are the corresponding vector representations, and $\mathbf{\Psi}(t)$ is an m-dimensional BPF vector of the form

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

with $\psi_k(t)$ being the kth block-pulse function of the mth order BPF function set,

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

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

$$Y_k \triangleq \frac{m}{T} \int_0^T y(t)\psi_k(t)\,\mathrm{d}t = \frac{m}{T} \int_{(k-1)(T/m)}^{k(T/m)} y(t)\,\mathrm{d}t, \qquad k = 1, \ldots, m. \tag{12}$$

BPF expansions are frequently used for the representation of deterministic signals, and it is well known that in the limit, as $m \to \infty$, the set (11) is complete and the signal representation $\{y_m(t)\}$ given by equation (9b) converges to $\{y(t)\}$ pointwise for any deterministic square-integrable signal defined in the interval $[0, T)$ [34]. Within the context of the present work the question of validity of such expansions for the case of stochastic signals is fundamental and, therefore, addressed first 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 $\{y(t)\}$):

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

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

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

$$\int_0^T [y(t)]^2\,\mathrm{d}t < \infty \qquad \text{(a.s.)},$$

properties of the random signals involved (see assumption A3 and the earlier discussion on $\{\tilde{w}(t)\}$), and Theorem 2 of Bharucha and Kadota [35]. This theorem states that the expansion of a random process $\{y(t)\}$, with $t \in [0, T)$, in terms of an arbitrary basis in $\mathscr{L}_2(T)$, 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:

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

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

$$
\begin{aligned}
E_k:\quad & \sum_{i=0}^{n_a} \left(\frac{(T/m)^i}{(i+1)!} a^\circ_{n_a - i} \sum_{j=1}^{l+k} Y_j f_{i,l+k+1-j} \right) \\
&= \sum_{i=0}^{n_a - 1} \left(\frac{(T/m)^{i+1}}{(i+2)!} b^\circ_{n_a - 1 - i} \sum_{j=1}^{l+k} U_j f_{i+1,l+k+1-j} \right) \\
&+ \sum_{i=0}^{n_a - 1} \left(\frac{(T/m)^i}{(i+1)!} c^\circ_{n_a - 1 - i} \sum_{j=1}^{l+k} \tilde{W}_j f_{i,l+k+1-j} \right), \qquad k = 0, 1, \ldots, na,
\end{aligned}
\tag{15}
$$

with l representing an arbitrary positive integer and $f_{k,i}$ a quantity defined in the Appendix [equation (A3)]. In a manner analogous to that used by Jiang and Schaufelberger [32] for deterministic systems, 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 $\{E_k\}$ leads to the following stochastic difference equation:

$$
\sum_{i=0}^{n_a} A_i^\circ Y_{l+i} = \sum_{i=0}^{n_a} B_i^\circ U_{l+i} + \sum_{i=0}^{n_a} \tilde{C}_i^\circ \tilde{W}_{l+i}, \tag{17}
$$

which is independent of any initial conditions. Through appropriate reindexing equation (17) may be expressed as:

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

In equation (18) $\{Y_k\}$, $\{U_k\}$ and $\{\tilde{W}_k\}$ represent the BPF expansions of the response, exogenous excitation, and Wiener process, respectively, and $A^\circ(B)$, $B^\circ(B)$ and $\tilde{C}^\circ(B)$ are polynomials in the backshift operator B ($BY_k \triangleq Y_{k-1}$) and of the following respective forms:

$$
A^\circ(B) \triangleq A^\circ_{n_a} + A^\circ_{n_a - 1} B + \cdots + A_1^\circ B^{n_a - 1} + A_0^\circ B^{n_a}, \tag{19a}
$$

$$
B^\circ(B) \triangleq B^\circ_{n_a} + B^\circ_{n_a - 1} B + \cdots + B_1^\circ B^{n_a - 1} + B_0^\circ B^{n_a}, \tag{19b}
$$

$$
\tilde{C}^\circ(B) \triangleq \tilde{C}^\circ_{n_a} + \tilde{C}^\circ_{n_a - 1} B + \cdots + \tilde{C}_1^\circ B^{n_a - 1} + \tilde{C}_0^\circ B^{n_a}. \tag{19c}
$$

The coefficients $\{A_i^\circ\}_{i=0}^{na}$ may be shown to be related to the coefficients of the original stochastic differential equation $\{a_i^\circ\}_{i=0}^{na-1}$ through the mapping expressions

$$
A_0^\circ = \sum_{i=0}^{na} (-1)^{n_a + i} \frac{(T/m)^i}{(i+1)!} a^\circ_{n_a - i}, \tag{20a}
$$

$$
A_i^\circ = \sum_{j=0}^{n_a} \left(\frac{(T/m)^j}{(j+1)!} a^\circ_{n_a - j} \sum_{k=0}^{n_a - i} (-1)^k \binom{n_a}{k} f_{j, n_a - k - i + 1} \right), \qquad i = 1, 2, \ldots, n_a. \tag{20b}
$$

Similar expressions relate the rest of the coefficients in equation (19) with those of the original equation $\mathscr{S}_{\mathscr{C}}$. By defining the continuous and discrete parameter vectors as

$$
\mathbf{a} \triangleq [a^\circ_{n_a} \quad \cdots \quad a_1^\circ \quad a_0^\circ]^{\mathrm{T}}, \qquad [(n_a + 1) \times 1]; \tag{21a}
$$

$$
\mathbf{b} \triangleq [b^\circ_{n_a - 1} \quad \cdots \quad b_1^\circ \quad b_0^\circ]^{\mathrm{T}}, \qquad [n_a \times 1]; \tag{21b}
$$

$$
\mathbf{c} \triangleq [c^\circ_{n_a - 1} \quad \cdots \quad c_1^\circ \quad c_0^\circ]^{\mathrm{T}}, \qquad [n_a \times 1]; \tag{21c}
$$

and

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

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

$$\tilde{\mathbf{C}} \triangleq [\tilde{C}^\circ_{n_a} \quad \cdots \quad \tilde{C}^\circ_1 \quad \tilde{C}^\circ_0]^{\mathrm{T}}, \qquad [(n_a+1)\times 1]; \tag{22c}$$

respectively, the relations between the discrete and continuous system parameters may be compactly rewritten 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 equation (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

$$\mathbf{D}_A = [\mathbf{d}_A(1) \mid \cdots \mid \mathbf{d}_A(n_a+1)], \qquad [(n_a+1)\times(n_a+1)]; \tag{24a}$$

and defining

$$\mathbf{D}_B = [\mathbf{d}_A(2) \mid \cdots \mid \mathbf{d}_A(n_a+1)], \qquad [(n_a+1)\times n_a]; \tag{24b}$$

$$\mathbf{D}_C = [\mathbf{d}_A(1) \mid \cdots \mid \mathbf{d}_A(n_a)], \qquad [(n_a+1)\times n_a]. \tag{24c}$$

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 $\mathscr{T}$ between the sets

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

TABLE 1

Relationships between the coefficients of $\mathscr{S}_{\mathscr{C}}$ [equation (1)] and $\mathscr{S}_{\tilde{\mathscr{D}}}$ [equation (18)]

First order system	Second order system
$A_1 = a_1 + \frac{1}{2}\frac{T}{m}a_0$	$A_2 = a_2 + \frac{1}{2}\frac{T}{m}a_1 + \frac{1}{6}\left(\frac{T}{m}\right)^2 a_0$
$A_0 = -a_1 + \frac{1}{2}\frac{T}{m}a_0$	$A_1 = -2a_2 + \frac{2}{3}\left(\frac{T}{m}\right)^2 a_0$
$B_1 = \frac{1}{2}\frac{T}{m}b_0$	$A_0 = a_2 - \frac{1}{2}\frac{T}{m}a_1 + \frac{1}{6}\left(\frac{T}{m}\right)^2 a_0$
$B_0 = \frac{1}{2}\frac{T}{m}b_0$	$B_2 = \frac{1}{2}\frac{T}{m}b_1 + \frac{1}{6}\left(\frac{T}{m}\right)^2 b_0$
$\tilde{C}_1 = c_0$	$B_1 = \frac{2}{3}\left(\frac{T}{m}\right)^2 b_0$
$\tilde{C}_0 = -c_0$	$B_0 = -\frac{1}{2}\frac{T}{m}b_1 + \frac{1}{6}\left(\frac{T}{m}\right)^2 b_0$
	$\tilde{C}_2 = c_1 + \frac{1}{2}\frac{T}{m}c_0$
	$\tilde{C}_1 = -2c_1$
	$\tilde{C}_0 = c_1 - \frac{1}{2}\frac{T}{m}c_0$

continued

TABLE 1—*continued*

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

and

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

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

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

Proof. $\mathscr{T}$ is a transformation since for any triple $(\mathbf{a}, \mathbf{b}, \mathbf{c}) \in \mathscr{C}$ expressions (23) define a unique triple $(\mathbf{A}, \mathbf{B}, \tilde{\mathbf{C}}) \in \tilde{\mathscr{D}}$. $\mathscr{T}$ is also onto since by construction of $\tilde{\mathscr{D}}$ every triple $(\mathbf{A}, \mathbf{B}, \tilde{\mathbf{C}}) \in \tilde{\mathscr{D}}$ is the image of at least one triple $(\mathbf{a}, \mathbf{b}, \mathbf{c}) \in \mathscr{C}$. That $\mathscr{T}$ 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 [36] in such a way as to rewrite $\mathbf{D}_A$ as the product of two matrices that may be verified to be non-singular. 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 $\mathscr{S}_{\tilde{\mathscr{D}}}$

Before the induced stochastic difference equation $\mathscr{S}_{\tilde{\mathscr{D}}}$ of equation (18) can be estimated, its structural and probabilistic properties need to be determined. Specifically, issues such as the stationarity of $\mathscr{S}_{\tilde{\mathscr{D}}}$, 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 very important as the noise dynamics of $\mathscr{S}_{\tilde{\mathscr{D}}}$ are *not* completely determined by the polynomial $\tilde{C}^\circ(B)$ alone, but also depend upon the correlation structure of $\{\tilde{W}_k\}$ itself.

With these ideas in mind, we proceed to examine the structural properties of the stochastic difference equation $\mathscr{S}_{\tilde{\mathscr{D}}}$ 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 $\mathscr{S}_{\tilde{\mathscr{D}}}$ given by equation (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) non-minimum phase, provided that its continuous-time counterpart is non-minimum phase.

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

(a1) Consider the system $y(t) = g^\circ(D)u(t)$ with the strictly proper [see equation (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 [37], which implies that for every bounded excitation the response will be also bounded; that is,

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

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, however, be always constructed from a bounded $\{u(t)\}$ through equation (12); for instance, select $u(t) = U_k$ for $(k-1)(T/m) \leqslant t < k(T/m)$. From equation (26), the response $\{y(t)\}$ to the excitation $\{u(t)\}$ will be bounded and, therefore, the (uniquely determined) response $\{Y_k\}$ will also be bounded because

$$\| Y_k \|_\infty \leqslant \frac{m}{T} \int_0^\infty \| y(t) \|_\infty \| \psi_k(t) \|_\infty \, \mathrm{d}t$$

$$= \frac{m}{T} \| y(t) \|_\infty \int_{(k-1)(T/m)}^{k(T/m)} \mathrm{d}t = \| y(t) \|_\infty < \infty, \tag{27}$$

where $\| Y_k \|_\infty \triangleq \sup_{t \in J_2} | Y_k |$ with $J_2 \triangleq [1, \infty)$.† As a consequence an arbitrary bounded excitation $\{U_k\}$ results in a bounded response $\{Y_k\}$, and the system $Y_k = G^\circ(B) U_k$ is asymptotically stable. The polynomial $A^\circ(B)$ is thus strictly minimum phase [37].

(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 the impulse response function:

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

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

$$\| G_k \|_\infty \leqslant \frac{m}{T} \int_0^\infty \| g(t) \|_\infty \| \psi_k(t) \|_\infty \, \mathrm{d}t$$

$$= \frac{m}{T} \| g(t) \|_\infty \int_{(k-1)(T/m)}^{k(T/m)} \mathrm{d}t$$

$$= \| g(t) \|_\infty < \infty, \tag{30}$$

which implies that $\{G_k\}$ is also bounded, and thus $G^\circ(B)$ is stable in the sense of Lyapunov. As a consequence, $A^\circ(B)$ is minimum phase [37].

(a3) Now assume that $a^\circ(D)$ is non-minimum 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 equations (4a) and (4b)], the unboundedness of $\{g(t)\}$ implies that:

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

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

$$G_k = \frac{m}{T} \int_{(k-1)(T/m)}^{k(T/m)} g(t) \, \mathrm{d}t > \frac{m}{T} \int_{(k-1)(T/m)}^{k(T/m)} M \, \mathrm{d}t = M, \qquad (g(t) > 0, \quad t \in \Delta,) \tag{32}$$

$$G_k = \frac{m}{T} \int_{(k-1)(T/m)}^{k(T/m)} g(t) \, \mathrm{d}t < -\frac{m}{T} \int_{(k-1)(T/m)}^{k(T/m)} M \, \mathrm{d}t = -M, \qquad (g(t) < 0, \quad t \in \Delta.) \tag{33}$$

Based on this we conclude that:

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

† 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 $(k-1)(T/m) \leqslant t < k(T/m)$, so that: $\| X_k \|_\infty = \| x'(t) \|_\infty$.

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

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

$$\mathscr{S}_{\tilde{\mathscr{C}}}: \quad a^\circ(D)y(t) = b^\circ(D)u(t) + \tilde{c}^\circ(D)\tilde{w}(t), \tag{35}$$

with

$$\tilde{c}^\circ(D) \triangleq Dc^\circ(D) = c^\circ_{n_a-1}D^{n_a} + c^\circ_{n_a-2}D^{n_a-1} + \quad \cdots \quad + c^\circ_1 D^2 + c^\circ_0 D + 0. \tag{36}$$

By defining:

$$\tilde{\mathbf{c}} \triangleq [c^\circ_{n_a-1} \quad \cdots \quad c^\circ_1 \quad c^\circ_0 \,|\, 0]^{\mathrm{T}}, \tag{37}$$

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

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

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

Corollary 2. An immediate consequence of Theorem 2 and assumption A2 pertaining to the strictly minimum phase nature of $a^\circ(D)$ is that the stochastic difference equation $\mathscr{S}_{\tilde{\mathscr{D}}}$ of equation (18) is *stationary*.

The structure of the polynomial $\tilde{C}^\circ(B)$ is of particular importance in the development of the estimation method and is thus further discussed in the following theorem:

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

$$\tilde{C}^\circ(B) = (1 - B)C^\circ(B), \tag{39}$$

with $C^\circ(B)$ being minimum phase.

Proof. The fact that $\tilde{C}^\circ(B)$ has a root at $B = 1$ may be shown by using a known property [32] stating that the sum of the A°_i's is proportional to a°_0; specifically,

$$\sum_{i=0}^{n_a} A^\circ_i = (T/m)^{n_a} a^\circ_0.$$

Because of equation (38), this is also applicable to $\tilde{C}^\circ(B)$, and therefore

$$\sum_{i=0}^{n_a} \tilde{C}^\circ_i = \tilde{C}^\circ(B)|_{B=1} = (T/m)^{n_a}\tilde{c}^\circ_0 = 0, \tag{40}$$

since $\tilde{c}^\circ_0 = 0$ [see equation (37)]. The fact that the root $B = 1$ is distinct, and $C^\circ(B)$ minimum phase, is a consequence of Theorem 2, which implies that $\tilde{C}^\circ(B)$ has to be minimum phase since $\tilde{c}^\circ(D)$ is such (based on the definition (36) and assumption A2). □

6. STOCHASTIC MODELING OF THE DISCRETE ENDOGENOUS EXCITATION SIGNAL $\{W_k\}$

For the development of an estimation method for the stochastic difference equation $\mathscr{S}_{\mathfrak{F}}$ [equation (18)], the probabilistic properties of the endogenous excitation signal $\{\tilde{W}_k\}$, defined as

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

also need to be analyzed, and a proper stochastic representation developed.

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

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

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

$$\begin{aligned} E[\tilde{W}_k\,\tilde{W}_l] &= \frac{m^2}{T^2}E\left[\int_0^T \tilde{w}(s)\psi_k(s)\,\mathrm{d}s \int_0^T \tilde{w}(t)\psi_l(t)\,\mathrm{d}t\right] \\ &= \frac{m^2(\sigma_w^\circ)^2}{T^2}\int_{(k-1)(T/m)}^{k(T/m)}\int_{(l-1)(T/m)}^{l(T/m)} \min(s, t)\,\mathrm{d}s\,\mathrm{d}t. \end{aligned} \tag{43}$$

Consider the following cases.

(a) Case $k = l$:

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

where

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

By symmetry,

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

and thus

$$E[\tilde{W}_k^2] = (k - \tfrac{2}{3})(T/m)(\sigma_w^\circ)^2. \tag{47}$$

(b) Case $k > l$:

$$E[\tilde{W}_k \tilde{W}_l] = \frac{m^2(\sigma_w^\circ)^2}{T^2} \int_{(k-1)(T/m)}^{k(T/m)} \int_{(l-1)(T/m)}^{l(T/m)} t \, \mathrm{d}t \, \mathrm{d}s$$

$$= \frac{m^2(\sigma_w^\circ)^2}{T^2} \left(\int_{(l-1)(T/m)}^{l(T/m)} t \, \mathrm{d}t \right) \left(\int_{(k-1)(T/m)}^{k(T/m)} \mathrm{d}s \right)$$

$$= (l - \tfrac{1}{2}) \frac{T}{m} (\sigma_w^\circ)^2. \tag{48}$$

(c) Case $k < l$:

$$E[\tilde{W}_k \tilde{W}_l] = \frac{m^2(\sigma_w^\circ)^2}{T^2} \int_{(k-1)(T/m)}^{k(T/m)} \int_{(l-1)(T/m)}^{l(T/m)} s \, \mathrm{d}s \, \mathrm{d}t$$

$$= \frac{m^2(\sigma_w^\circ)^2}{T^2} \left(\int_{(k-1)(T/m)}^{k(T/m)} s \, \mathrm{d}s \right) \left(\int_{(l-1)(T/m)}^{l(T/m)} \mathrm{d}t \right)$$

$$= (k - \tfrac{1}{2}) \frac{T}{m} (\sigma_w^\circ)^2. \tag{49}$$

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

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

Evidently, this autocovariance is not a function of the relative lag $k - l$, and the sequence $\{\tilde{W}_k\}$ is therefore non-stationary. In order further to investigate its structure we define a new sequence $\{Z_k\}$ as

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

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

(a) Case $k = l$:

$$E[Z_k^2] = E[\tilde{W}_k^2] - 2E[\tilde{W}_k \tilde{W}_{k-1}] + E[\tilde{W}_{k-1}^2]$$

$$= (k - \tfrac{2}{3}) \frac{T}{m} (\sigma_w^\circ)^2 - 2(k - 1 - \tfrac{1}{2}) \frac{T}{m} (\sigma_w^\circ)^2 + (k - 1 - \tfrac{2}{3}) \frac{T}{m} (\sigma_w^\circ)^2 = \tfrac{2}{3} \frac{T}{m} (\sigma_w^\circ)^2. \tag{52}$$

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

$$E[Z_k Z_{k-1}] = E[\tilde{W}_k \tilde{W}_{k-1}] - E[\tilde{W}_{k-1}^2] - E[\tilde{W}_k \tilde{W}_{k-2}] + E[\tilde{W}_{k-1} \tilde{W}_{k-2}]$$

$$= (k - 1 - \tfrac{1}{2}) \frac{T}{m} (\sigma_w^\circ)^2 - (k - 1 - \tfrac{2}{3}) \frac{T}{m} (\sigma_w^\circ)^2$$

$$- (k - 2 - \tfrac{1}{2}) \frac{T}{m} (\sigma_w^\circ)^2 + (k - 2 - \tfrac{1}{2}) \frac{T}{m} (\sigma_w^\circ)^2$$

$$= \tfrac{1}{6} \frac{T}{m} (\sigma_w^\circ)^2. \tag{53}$$

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

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

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

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

This implies that $\{Z_k\}$ is a stationary MA(1) process [21], and thus amenable to the representation

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

with $\{N_k\}$ being 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

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

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

$$\theta_1^\circ \simeq 0{\cdot}267949, \qquad (\sigma_N^\circ)^2 \simeq 0{\cdot}622008 \frac{T}{m} (\sigma_w^\circ)^2. \tag{58a, b}$$

These results lead to the following important lemma:

Lemma 2. The BPF expansion series $\{\tilde{W}_k\}$ of the Wiener process $\{\tilde{w}(t)\}$ can be modeled as a non-stationary Integrated Moving Average IMA(1,1) process [11] of the form:

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

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

7. THE ESTIMATION METHOD

By substituting the IMA(1,1) representation (59) of the discrete endogenous excitation signal and the form of $\tilde{C}^\circ(B)$ given by equation (39) into the difference equation $\mathscr{S}_{\tilde{\mathscr{D}}}$

[equation (18)], the following stochastic difference equation is obtained:

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

In this expression, i.i.d. stands for independently identically distributed. Since θ_1° is *a priori* known we may re-express $\mathscr{S}_{\mathscr{D}}$ in terms of the filtered sequences:

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

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

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

with

$$A^{\circ\prime}(B) \triangleq 1 + A_1^{\circ\prime}B + \cdots + A_{n_a}^{\circ\prime}B^{n_a}, \quad (63a)$$

$$\mathbf{A}' \triangleq [1 \quad A_1^{\circ\prime} \quad \cdots \quad A_{n_a}^{\circ\prime}]^{\mathrm{T}} \triangleq \mathbf{A}/A_{n_a}^{\circ}, \quad (63b)$$

$$B^{\circ\prime}(B) \triangleq B_0^{\circ\prime} + B_1^{\circ\prime}B + \cdots + B_{n_a}^{\circ\prime}B^{n_a}, \quad (63c)$$

$$\mathbf{B}' \triangleq [B_0^{\circ\prime} \quad B_1^{\circ\prime} \quad \cdots \quad B_{n_a}^{\circ\prime}]^{\mathrm{T}} \triangleq \mathbf{B}/A_{n_a}^{\circ}, \quad (63d)$$

$$C^{\circ\prime}(B) \triangleq 1 + C_1^{\circ\prime}B + \cdots + C_{n_a-1}^{\circ\prime}B^{n_a-1}, \quad (63e)$$

$$\mathbf{C}' \triangleq [1 \quad C_1^{\circ\prime} \quad \cdots \quad C_{n_a-1}^{\circ\prime}]^{\mathrm{T}} \triangleq \mathbf{C}/C_{n_a-1}^{\circ} \triangleq \mathbf{C}/\tilde{C}_{n_a}^{\circ}, \quad (63f)$$

and

$$N_k' \triangleq \frac{C_{n_a-1}^{\circ}}{A_{n_a}^{\circ}} N_k \triangleq \frac{\tilde{C}_{n_a}^{\circ}}{A_{n_a}^{\circ}} N_k. \quad (64)$$

Based upon equation (64), $\{N_k'\}$ is a zero-mean innovations sequence with variance

$$(\sigma_{N'}^{\circ})^2 = (C_{n_a-1}^{\circ}/A_{n_a}^{\circ})^2(\sigma_N^{\circ})^2. \quad (65)$$

The normalized stochastic difference equation $\mathscr{S}_{\mathscr{D}'}$ [equation (62)] with the stationary zero-mean and uncorrelated endogenous 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 $A^{\circ\prime}(B)$ is strictly minimum phase by virtue of Corollary 2, which guarantees the strictly minimum phase nature of $A^{\circ}(B)$; and (3) characterized by a minimum phase MA polynomial $C^{\circ\prime}(B)$ by virtue of assumption A2 and Theorem 3, which guarantee that $C^{\circ}(B)$ is minimum phase. $\mathscr{S}_{\mathscr{D}'}$ is therefore *identifiable.*

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

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

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

and

$$\mathscr{D}'_a \triangleq \{(\mathbf{A}', \mathbf{B}', \mathbf{C}', (\sigma^\circ_{N'})^2) \in \mathscr{R}^{n_a+1} \times \mathscr{R}^{n_a+1} \times \mathscr{R}^{n_a} \times \mathscr{R}^+ | \text{ with } \mathbf{A}', \mathbf{B}', \mathbf{C}' \text{ generated through equations (23), (36) and (63), and } (\sigma^\circ_{N'})^2 \text{ by equations (58b), (65) by all } (\mathbf{a}, \mathbf{b}, \mathbf{c}, (\sigma^\circ_w)^2) \in \mathscr{C}_a\} \tag{66b}$$

is a bijective transformation.

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

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

$$\underline{\mathbf{a}} \triangleq \mathbf{D}_A^{-1}\mathbf{A}', \qquad \mathbf{a} = \underline{\mathbf{a}}/\underline{a}^\circ_{n_a}, \tag{67a, b}$$

$$\mathbf{A} = \mathbf{D}_A\mathbf{a}, \qquad \mathbf{b} = \Delta_B^{-1}\mathbf{B}'_r A^\circ_{n_a}, \tag{68a, b}$$

$$\tilde{C}^{\circ\prime}(B) = (1-B)C^{\circ\prime}(B), \qquad \underline{\mathbf{c}} \triangleq \Delta_C^{-1}\tilde{\mathbf{C}}'_r, \qquad \mathbf{c} = \underline{\mathbf{c}}/\underline{c}^\circ_{n_a-1}, \tag{69a–c}$$

$$\tilde{\mathbf{C}} = \mathbf{D}_C\mathbf{c}, \qquad (\sigma^\circ_w)^2 = \frac{m}{T} \times \frac{1}{0{\cdot}622008} \times \left(\frac{A^\circ_{n_a}}{\tilde{C}^\circ_{n_a}}\right)^2 \times (\sigma^\circ_{N'})^2, \tag{70a, b}$$

where $\mathbf{B}'_r$ and $\tilde{\mathbf{C}}'_r$ represent arbitrary n_a-dimensional subvectors of $\mathbf{B}'$ and $\tilde{\mathbf{C}}'$, respectively, Δ_B, Δ_C, appropriate $n_a \times n_a$ submatrices of $\mathbf{D}_B$ and $\mathbf{D}_C$, respectively, and $\underline{\mathbf{a}} \triangleq [\underline{a}^\circ_{n_a} \cdots \underline{a}^\circ_0]^T$, $\underline{\mathbf{c}} \triangleq [\underline{c}^\circ_{n_a-1} \cdots \underline{c}^\circ_0]^T$ 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 Δ_B and Δ_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 $\mathscr{S}_{\mathscr{D}'}$ given by equation (62)]

$$\mathscr{M}_{\mathscr{D}'}: \quad A'(B)Y^F_k = B'(B)U^F_k + C'(B)E'_k, \tag{71}$$

with $\{E'_k\}$ representing the model's one-step-ahead prediction error sequence, is estimated based on the available sequences $\{U^F_k\}^m_{k=1}$ and $\{Y^F_k\}^m_{k=1}$. In the second stage, the parameter estimates of the continuous-time ARMAX system $\mathscr{S}_{\mathscr{C}}$ [equation (1)] are obtained through expressions similar to equations (67)–(70).

Due to inevitable estimation errors, however, the model actually obtained, $\mathscr{M}_{\mathscr{D}'}$, may not belong to the set $\mathscr{D}'_a$ ($\mathscr{M}_{\mathscr{D}'} \notin \mathscr{D}'_a$), in which case the former will have no image within the set $\mathscr{C}'_a$ of continuous-time ARMAX systems. This problem may be dealt with by assigning to $\mathscr{M}_{\mathscr{D}'}$ that continuous-time ARMAX model $\mathscr{M}_{\mathscr{C}} \in \mathscr{C}'_a$ the image in $\mathscr{D}'_a$ of which is, in some appropriate sense, "closest" to that of the estimated $\mathscr{M}_{\mathscr{D}'}$. 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 equations (67)–(70)].

(1) For the estimation of the AR parameter vector $\mathbf{a}^\circ$:

$$\hat{\underline{\mathbf{a}}} \triangleq \mathbf{D}_A^{-1}\widehat{\mathbf{A}'}, \qquad \hat{\mathbf{a}} = \hat{\underline{\mathbf{a}}}/\hat{\underline{a}}_{n_a}. \tag{72a, b}$$

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

(2) For the estimation of the X parameter vector $\mathbf{b}^{\circ}$:

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

In these expressions, $\hat{\mathbf{A}}$ is the vector of the estimated coefficients of $A^{\circ}(B)$, $\widehat{\mathbf{B}'}$ the vector of the estimated coefficients of $B^{\circ\prime}(B)$, and $\hat{\mathbf{b}}$ the vector of the final X parameter estimates (the estimates of the coefficients of the polynomial $b^{\circ}(D)$).

(3) For the estimation of the MA parameter vector $\mathbf{c}^{\circ}$:

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

$$\underset{\sim}{\hat{\mathbf{c}}} \triangleq (\mathbf{D}_C^{\mathrm{T}} \mathbf{D}_C)^{-1} \mathbf{D}_C^{\mathrm{T}} \widehat{\tilde{\mathbf{C}}'}, \qquad \hat{\mathbf{c}} = \underset{\sim}{\hat{\mathbf{c}}} / \underset{\sim}{\hat{c}}_{n_a - 1}. \tag{74b, c}$$

In these expressions, $\widehat{\tilde{C}'}(B)$ is the estimate of $\tilde{C}(B)/\tilde{C}_{n_a}$, $\widehat{\tilde{\mathbf{C}}'}$ is the vector of the coefficients of $\tilde{C}'(B)$, $\widehat{C'}(B)$ is the estimate of $C'(B)$, $\underset{\sim}{\hat{\mathbf{c}}}$ is an intermediate estimate, and $\hat{\mathbf{c}} = [1 \hat{c}_{n_a - 2} \cdots \hat{c}_1 \; \hat{c}_0]^{\mathrm{T}}$ is the vector of the final MA parameter estimates (the estimates of the coefficients of the polynomial $c^{\circ}(D)$).

(4) For the estimation of the spectral height $(\sigma_w^{\circ})^2$:

$$\hat{\tilde{\mathbf{C}}} = \mathbf{D}_C \hat{\mathbf{c}}, \qquad \hat{\sigma}_w^2 = \frac{m}{T} \frac{1}{0{\cdot}622008} \left(\frac{\hat{A}_{n_a}}{\hat{\tilde{C}}_{n_a}} \right)^2 (\hat{\sigma}_{N'})^2. \tag{75a, b}$$

7.1. SUMMARY OF THE ESTIMATION METHOD

The proposed estimation method may be thus summarized as follows:

Step 1: obtain the BPF spectral representations $\{U_k\}_{k=1}^m$ and $\{Y_k\}_{k=1}^m$ of the exogenous excitation $\{u(t)\}$ and 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 successive values of n_a by using Maximum Likelihood estimation. In each case, compute the Bayesian Information Criterion [38]:

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

where n_θ denotes the total number of estimated parameters, $(\hat{\sigma}_{N'})^2$ is the estimated variance of the discrete innovations $\{N_k'\}$, and m is 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).

7.2. REMARKS

The following remarks are in order.

(a) *Existence and uniqueness of the estimates* $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$, $\hat{\mathbf{c}}$ *and* $\hat{\sigma}_w^2$. The existence and uniqueness of the continuous-time ARMAX system parameter estimates is a consequence of Theorem 4 and of the full rank property of $\mathbf{D}_B$ and $\mathbf{D}_C$. This in turn, guarantees [39] the existence and uniqueness of their Moore–Penrose pseudo-inverses used in equations (73b) and (74b).

(b) *Consistency of the parameter estimates.* As is well known, the Maximum Likelihood or Prediction Error estimator of the parameters of a discrete ARMAX system are, under mild assumptions, consistent [28], and such therefore is the estimator of $\mathscr{S}_{\mathscr{D}'}$ [equation (62)]. Hence, asympotically $(m \to \infty)$, the estimated model $\mathscr{M}_{\mathscr{D}'} \in \mathscr{D}_a'$ (in probability), and since the parameters of the continuous-time ARMAX system are obtained as rational functions of the former [due to normalization, see equations (72)–(75)], the application of Slutsky's lemma [40] implies that the estimator

$$\hat{\boldsymbol{\theta}} \triangleq [\hat{\mathbf{a}}^{\mathrm{T}} \quad \hat{\mathbf{b}}^{\mathrm{T}} \quad \hat{\mathbf{c}}^{\mathrm{T}} \quad \hat{\sigma}_w^2]^{\mathrm{T}} \tag{77}$$

will also be consistent; that is,

$$\hat{\boldsymbol{\theta}} \xrightarrow{\text{prob}} \boldsymbol{\theta}^{\circ}, \qquad \text{as } m \to \infty, \tag{78}$$

where $\boldsymbol{\theta}^{\circ}$ represents the true parameter vector and m is the length of the BPF spectral representations used.

8. SIMULATION RESULTS AND DISCUSSION

The performance characteristics of the proposed method are now evaluated via numerical simulations using a digital implementation. For a given continuous-time ARMAX system of the form (3) and specified exogenous $\{u(t)\}$ and endogenous $\{w(t)\}$ excitations, the response signal $\{y(t)\}$ is calculated by integrating (using a fourth order Runge–Kutta method) each of the following differential equations:

$$a^{\circ}(D)y_1(t) = b^{\circ}(D)u(t), \qquad a^{\circ}(D)y_2(t) = c^{\circ}(D)w(t), \tag{79a, b}$$

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 $20 \leqslant \tau/\Delta t \leqslant 200$, with τ representing the smallest system period or time constant.

In all cases examined, the selected exogenous excitation signals are composed of trains of pulses of duration equal to the BPF width T/m and amplitudes forming a sequence of Gaussian independently identically distributed (i.i.d.) random variables with zero mean and unit variance. The BPF expansions are computed from equation (12) by using Simpson's composite rule [41]. The estimation of discrete ARMAX models of the form (71) is then based on the computed BPF spectral records.

The finally selected estimated model is validated by examining its predictive ability and the characteristics of its residual (the one-step-ahead prediction error) sequence computed from equation (71) for the estimated parameter values. For a good model, the residual sequence must be uncorrelated, and this is judged by examining whether its normalized sample autocorrelation lies within the 95% confidence interval of $\pm 1{\cdot}96/\sqrt{m}$ [21].

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

$$E_p = \frac{\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{\circ} \|}{\| \boldsymbol{\theta}^{\circ} \|} \times 100\%, \tag{80}$$

with $\boldsymbol{\theta}$ representing a selected parameter vector and $\| \cdot \|$ the Euclidean norm.

8.1. ESTIMATION RESULTS

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

$$(D^2 + 2D + 16)y(t) = (D + 10)u(t) + (D + 9)w(t)$$

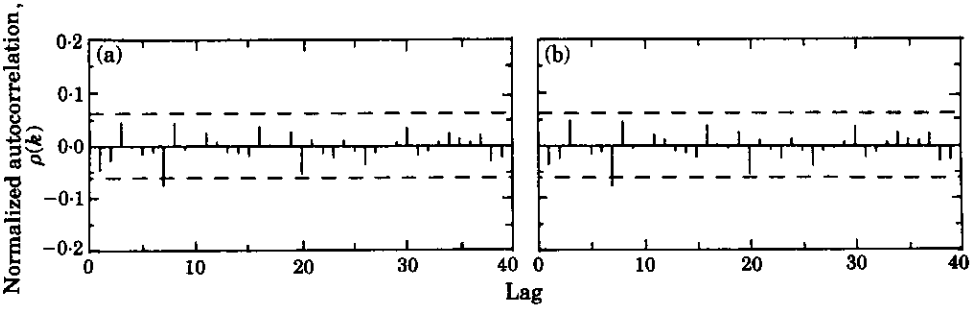

Figure 1. The normalized sample autocorrelation of the discrete residual sequence for the estimated system A: (a) $\sigma_w^2 = 0{\cdot}005$; (b) $\sigma_w^2 = 0{\cdot}01$.

is considered first, based on data records that are 100 s long ($T = 100$ s) and generated with white noise sequences having spectral heights $(\sigma_w^\circ)^2 = 0{\cdot}005$ and $(\sigma_w^\circ)^2 = 0{\cdot}01$ (two cases). The integration step and BPF width were selected equal to $\Delta t = 0{\cdot}01$ s and $T/m = 10\Delta t$, 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 depicting the normalized sample autocorrelation of the residuals lying within the 95% confidence interval of $\pm 1{\cdot}96/\sqrt{m}$ indicate, were successfully validated. From these models, the continuous-time ARMAX system parameters that are summarized in Table 2 are obtained. The estimated frequency response characteristics of both the $b(D)/a(D)$ and $c(D)/a(D)$ transfer functions are also compared to their theoretical counterparts in Figure 2. As the results demonstrate, excellent accuracy is achieved and the parametric percentage errors are confined to reasonably small values. The discrepancy in the estimated innovations spectral height, which is roughly constant in all cases, is attributed to the way in which the stochastic differential equation (79b) was digitally simulated. Similar remarks may be also made from Table 3, in which the results of a Monte Carlo analysis of the method

TABLE 2

Estimation results for system A at two different noise powers

	Process parameters	Estimated parameters†	
		$\sigma_w^2 = 0{\cdot}005$	$\sigma_w^2 = 0{\cdot}01$
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

† For the simulation $\tau/\Delta t \simeq 157$.

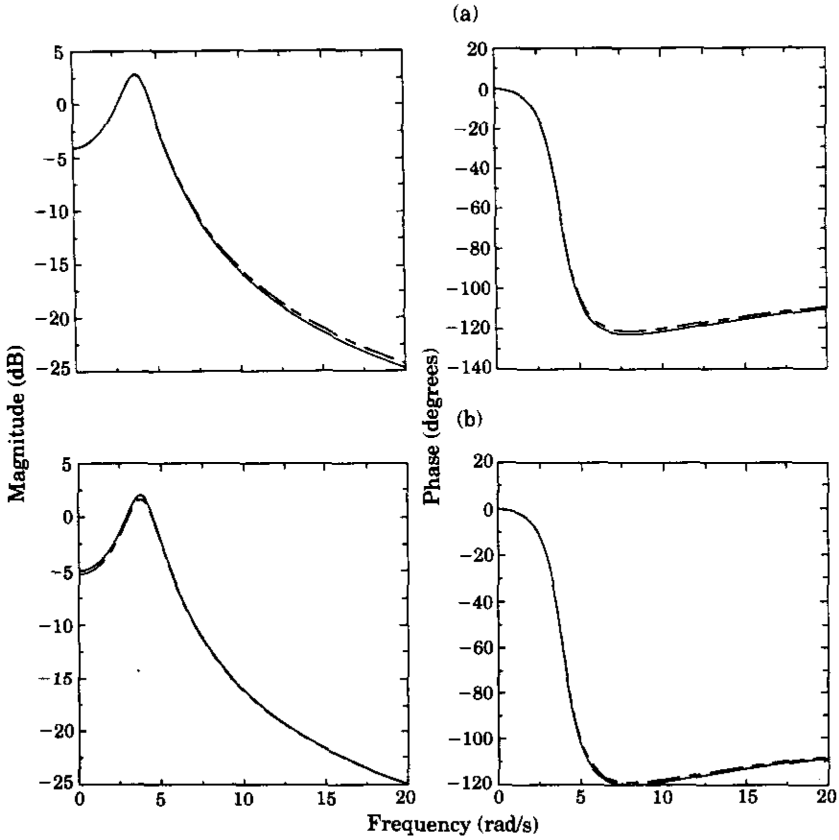

Figure 2. Frequency response curves of the estimated continuous-time system A: (a) transfer function $b(D)/a(D)$; (b) transfer function $c(D)/a(D)$. ——, Theoretical; – – –, estimated for $\sigma_w^2 = 0{\cdot}005$; — — —, estimated for $\sigma_w^2 = 0{\cdot}01$.

TABLE 3

Monte Carlo results for system A

	System parameters	Estimated parameters: Mean value	Estimated parameters: 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

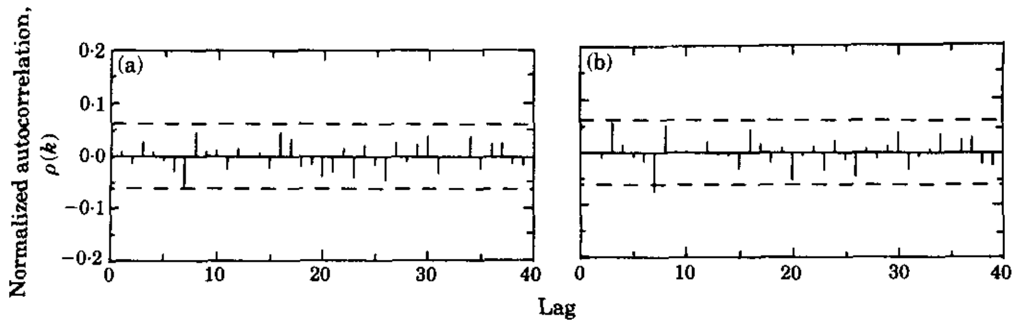

Figure 3. The normalized sample autocorrelation of the discrete residual sequence for the estimated system B: (a) $\sigma_w^2 = 0{\cdot}00005$; (b) $\sigma_w^2 = 0{\cdot}0001$.

based on 20 data records generated by different seed numbers and $(\sigma_w^o)^2 = 0{\cdot}005$ are presented.

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

$$(D^2 + 3D + 2)y(t) = u(t) + (D + 10)w(t),$$

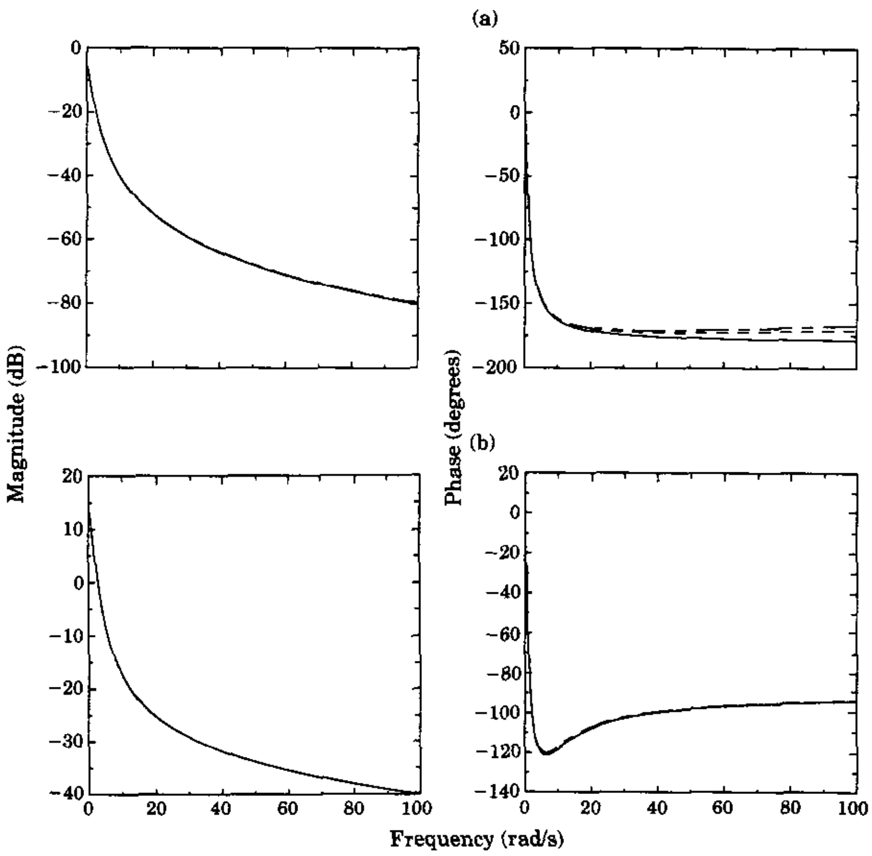

Figure 4. Frequency response curves of the estimated continuous-time system B: (a) transfer function $b(D)/a(D)$; (b) transfer function $c(D)/a(D)$. ——, Theoretical; – – –, estimated for $\sigma_w^2 = 0{\cdot}00005$; — · —, estimated for $\sigma_w^2 = 0{\cdot}0001$.

TABLE 4

Estimation results for system B at two different noise powers

	System parameters	Estimated parameters†	
		$\sigma_w^2 = 0{\cdot}00005$	$\sigma_w^2 = 0{\cdot}0001$
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	—	$1{\cdot}1084 \times 10^{-4}$	$2{\cdot}2154 \times 10^{-4}$

† For the simulation $\tau/\Delta t = 50$.

is considered based on data records that are 100 s long and generated with white noise sequences having spectral heights $(\sigma_w^\circ)^2 = 0{\cdot}00005$ and $(\sigma_w^\circ)^2 = 0{\cdot}0001$ (two cases). The integration step and BPF width were selected as in the previous example, that is $\Delta t = 0{\cdot}01$ s and $T/m = 10\Delta t$. In both of the considered cases, discrete ARMAX(2,2,1) models were estimated and successfully validated (Figure 3), from which the final

TABLE 5

Monte Carlo results for system B

	System 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}	$1{\cdot}14245 \times 10^{-4}$	$6{\cdot}92103 \times 10^{-6}$

TABLE 6

Order determination results for system C ($\sigma_w^2 = 0{\cdot}005$)

System order, n_a	*BIC*
2	$-8{\cdot}16820$
3	$-8{\cdot}77149$
4	$-8{\cdot}77039$

estimation results presented in Table 4 were obtained. The achievable accuracy is very good, and the estimated frequency response characteristics nicely match those of the corresponding theoretical curves (Figure 4). A Monte Carlo analysis based on 20 data records and $(\sigma_w^\circ)^2 = 0{\cdot}00005$ (Table 5) further confirms the method's performance characteristics.

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

$$(D^3 + 11D^2 + 424D + 1200)y(t) = (2D^2 + 60D + 800)u(t) + (D^2 + 12D + 225)w(t)$$

is considered based on data records that are 37·5 s long and generated with white noise sequences having spectral heights $(\sigma_w^\circ)^2 = 0{\cdot}005$ and $(\sigma_w^\circ)^2 = 0{\cdot}01$ (two cases). The integration step and BPF width were selected as $\Delta t = 3{\cdot}125 \times 10^{-3}$ s and $T/m = 8\Delta t$, 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 $[(\sigma_w^\circ)^2 = 0{\cdot}005]$, are presented in Table 7, Figure 6 and Table 8, respectively.

TABLE 7

Estimation results for system C at two different noise powers

	System parameters	Estimated parameters†	
		$\sigma_w^2 = 0{\cdot}005$	$\sigma_w^2 = 0{\cdot}01$
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_2	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

† For the simulation $\tau/\Delta t \simeq 100$.

TABLE 8

Monte Carlo results for system C

	System 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

9. CONCLUSIONS

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. The 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 endogenous and exogenous 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 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.

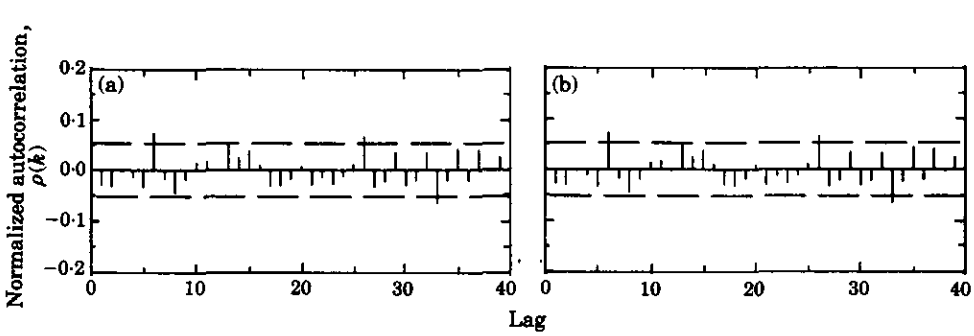

Figure 5. The normalized sample autocorrelation of the discrete residual sequence for the estimated system C: (a) $\sigma_w^2 = 0·005$; (b) $\sigma_w^2 = 0·01$.

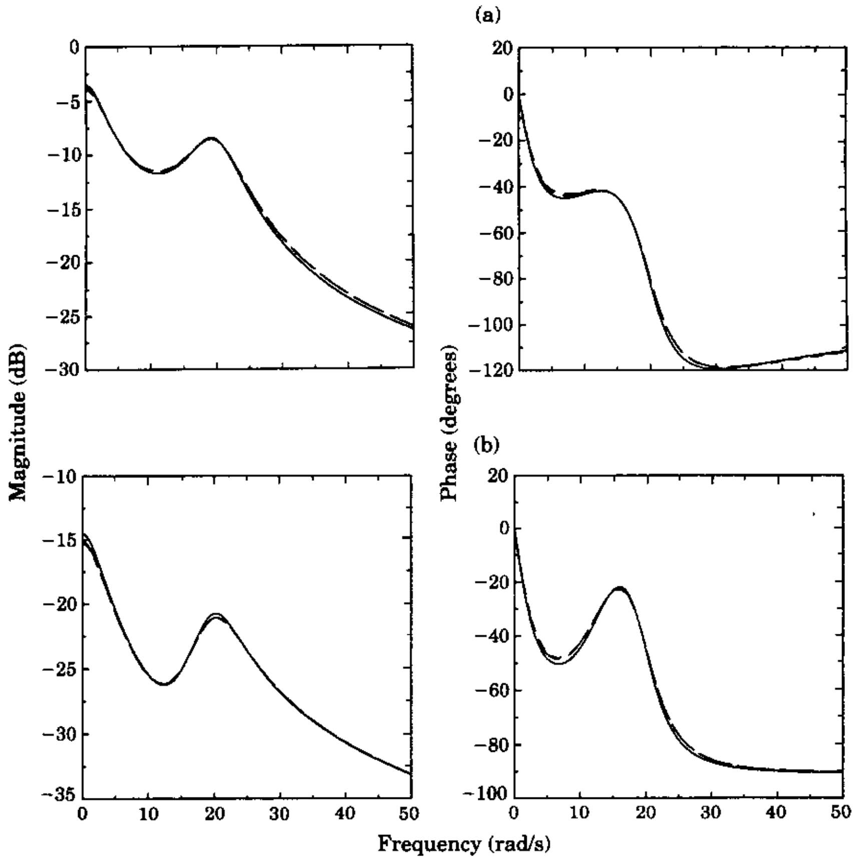

Figure 6. Frequency response curves of the estimated continuous-time system C: (a) transfer function $b(D)/a(D)$; (b) transfer function $c(D)/a(D)$. ——, Theoretical; ---, estimated for $\sigma_w^2 = 0{\cdot}005$; — — —, estimated for $\sigma_w^2 = 0{\cdot}01$.

Among the unique features and advantages of the proposed estimation method that make it especially attractive for applications are the following. (a) The fact that neither estimates of signal derivatives, nor direct discretizations associated with instantaneous sampling 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 not restricted to be uniformly (and reasonably slowly) sampled, but may in fact by in *analog* form or in very frequently sampled digital form. The choice is with the user and the available type of signal processing hardware. Apart from its obvious significance, this fact also implies that non-uniformly sampled and/or missing data can be accommodated if necessary. (d) The mapping relationship between the discrete and the original continuous-time system parameters is *linear*, so that sensitivity problems associated with highly non-linear transformations are eliminated, the computational complexity is reduced, and frequently available *a priori* system information can be readily incorporated into the estimation procedure.

The effectiveness and good performance characteristics of the proposed method were finally demonstrated via a number of numerical simulations using a digital implementation.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the financial support of the Indonesian Second University Development Project (P.N.) and the Whirlpool Corporation (S.D.F.) that enabled the undertaking of this research work.

REFERENCES

1. P. YOUNG 1981 *Automatica* **17**, 23–39. Parameter estimation for continuous-time models—a survey.
2. P. YOUNG and A. JAKEMAN 1980 *International Journal of Control* **31**, 741–764. Refined instrumental variable methods of recursive time-series analysis—part III: extensions.
3. H. UNBEHAUEN and G. P. RAO 1987 *Identification of Continuous Systems*. Amsterdam: North Holland.
4. B. WAHLBERG 1989 in *Proceedings of the Eighth IFAC/IFORS Symposium on Identification and System Parameter Estimation* (H. F. Chen, editor), 435–440. Oxford: Pergamon Press. On the identification of continuous time dynamical systems.
5. D. NEUMANN, R. ISERMANN and S. NOLD 1989 in *Proceedings of the Eighth IFAC/IFORS Symposium on Identification and System Parameter Estimation* (H. F. Chen, editor), 851–856. Oxford: Pergamon Press. Comparisons of some parameter estimation methods for continuous-time models.
6. S. SAGARA and Z. Y. ZHAO 1990 *Automatica* **26**, 63–74. Numerical integration approach to on-line identification of continuous-time systems.
7. H. UNBEHAUEN and G. P. RAO 1990 *Automatica* **26**, 23–35. Continuous-time approaches to system identification—a survey.
8. A. V. BALAKRISHNAN 1973 *Stochastic Differential Systems*. Lecture Notes in Economics and Mathematical Systems 84. Berlin: Springer-Verlag.
9. A. BAGCHI 1975 *Automatica* **11**, 533–536. Continuous time system identification with unknown noise covariance.
10. PHAM-DINH-TUAN 1977 *Biometrika* **64**, 385–399. Estimation of parameters of a continuous time Gaussian stationary process with rational spectral density.
11. M. B. PRIESTLEY 1981 *Spectral Analysis and Time Series*. New York: Academic Press.
12. J. H. VAN SCHUPPEN 1983 *Journal of Mathematical Analysis and Applications* **96**, 209–225. Convergence results for continuous-time adaptive stochastic filtering algorithms.
13. J. B. MOORE 1988 *Stochastic Processes and Their Applications* **27**, 195–215. Convergence of continuous time stochastic ELS parameter estimation.
14. J. E. LEE and S. D. FASSOIS 1992 *ASME Journal of Vibration and Acoustics* **114**, 260–271. Suboptimum maximum likelihood estimation of structural parameters from multiple-excitation vibration data.
15. K. J. ÅSTRÖM and C. G. KÄLLSTRÖM 1973 in *Proceedings of the Third IFAC/IFORS Symposium on Identification and System Parameter Estimation* (P. Eykhoff, editor), 415–424. Amsterdam: North Holland. Application of system identification techniques to the determination of ship dynamics.
16. M. S. PHADKE and S. M. WU 1974 *Journal of the American Statistical Association* **69**, 325–329. Modeling of continuous stochastic processes from discrete observations with applications to sunspot data.
17. S. M. PANDIT and S. M. WU 1975 *Biometrika* **62**, 479–501. Unique estimates of the parameters of a continous stationary stochastic process.
18. S. D. FASSOIS, K. F. EMAN and S. M. WU 1990 *ASME Journal of Dynamic Systems, Measurement, and Control* **112**, 1–9. Sensitivity analysis of the discrete-to-continuous dynamic system transformation.
19. S. M. PANDIT 1973 *Ph.D. Dissertation, University of Wisconsin at Madison*. Data-dependent systems: modeling, analysis, and optimal control via time-series.
20. S. M. WU 1977 *ASME Journal of Engineering for Industry* **99**, 708–714. Dynamic data system: a new modeling technique.
21. S. M. PANDIT and S. M. WU 1983 *Time Series and System Analysis with Applications*. New York: John Wiley.
22. P. M. ROBINSON 1980 in *Directions in Time Series* (D. Brillinger and G. C. Tiao, editors), 263–278. Institute of Mathematical Statistics. Continuous model fitting from discrete data.

23. R. H. JONES 1981 in *Applied Time Series Analysis II* (D. F. Findley, editor), 651–682. New York: Academic Press. Fitting a continuous time autoregression to discrete data.
24. V. SOLO 1984 in *Time Series Analysis of Irregularly Observed Data* (E. Parzen, editor), 325–345. Berlin: Springer-Verlag. Some aspects of continuous-discrete time series modelling.
25. J. D. SARGAN 1974 *Journal of the Royal Statistical Society, Series B* **36**, 74–90. Some discrete approximation to continuous time stochastic models.
26. A. R. BERGSTROM (editor) 1976 *Statistical Inference in Continuous Time Economic Models*. Amsterdam: North Holland.
27. A. R. BERGSTROM 1983 *Econometrica* **51**, 117–152. Gaussian estimation of structural parameters in higher order continuous time dynamic models.
28. L. LJUNG 1987 *System Identification—Theory for the User*. Englewood Cliffs, New Jersey: Prentice-Hall.
29. R. BEN MRAD and S. D. FASSOIS 1991 *ASME Journal of Vibration and Acoustics* **113**, 354–361. Recursive identification of vibrating structures from noise-corrupted observations—part I: identification methods.
30. C. R. WYMER 1972 *Econometrica* **40**, 565–577. Econometric estimation of stochastic differential equation systems.
31. C. H. WANG 1982 *International Journal of Control* **36**, 67–76. Generalized block-pulse operational matrices and their applications to operational calculus.
32. Z. H. JIANG and W. SCHAUFELBERGER 1985 *International Journal of Systems Science* **16**, 1559–1571. A new algorithm for single-input single-output system identification via block-pulse functions.
33. A. H. JAZWINSKI 1970 *Stochastic Processes and Filtering Theory*. New York: Academic Press.
34. C. P. KWONG and C. F. CHEN 1981 *International Journal of Systems Science* **12**, 745–751. The convergence properties of block-pulse series.
35. B. H. BHARUCHA and T. T. KADOTA 1970 *IEEE Transactions on Information Theory* **16**, 139–141. On the representation of continuous parameter processes by a sequence of random variables.
36. F. KRAUS and W. SCHAUFELBERGER 1990 *International Journal of Control* **51**, 931–942. Identification with block pulse functions, modulating functions, and differential operators.
37. C. T. CHEN 1984 *Linear System Theory and Design*. New York: Holt, Rinehart and Winston.
38. H. AKAIKE 1977 in *Applications of Statistics* (P. R. Krishnaiah, editor), 27–41. Amsterdam: North Holland. On entropy maximization principle.
39. S. BARNETT 1990 *Matrices: Methods and Applications*. New York: Clarendon Press.
40. H. CRAMER 1946 *Mathematical Methods of Statistics*. Princeton, New Jersey: Princeton University Press.
41. R. L. BURDEN and J. D. FAIRES 1985 *Numerical Analysis*. Boston: Prindle, Weber and Schmidt; third edition.

APPENDIX: THE OPERATIONAL MATRIX FOR INTEGRATION OF SIGNALS IN THE BPF REPRESENTATION

Consider the mth 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

$$\underbrace{\int_0^t \cdots \int_0^t}_{k\text{-fold}} y(t')\,\mathrm{d}t' \simeq \mathbf{X}^{\mathrm{T}}\mathbf{F}_k\boldsymbol{\Psi}(t), \tag{A1}$$

where

$$\mathbf{F}_k \triangleq \left(\frac{T}{m}\right)^k \frac{1}{(k+1)!}\begin{bmatrix} f_{k,1} & f_{k,2} & f_{k,3} & \cdots & f_{k,m} \\ 0 & f_{k,1} & f_{k,2} & \cdots & f_{k,m-1} \\ 0 & 0 & f_{k,1} & \cdots & f_{k,m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f_{k,1} \end{bmatrix}, \qquad [m \times m], \tag{A2}$$

and

$$f_{k,1} = 1, \qquad \forall k,$$
$$f_{k,i} = i^{k+1} - 2(i-1)^{k+1} + (i-2)^{k+1}, \qquad i = 2, 3, \ldots, m. \tag{A3}$$

The matrix $\mathbf{F}_k$ is called the kth order operational matrix for integration [31]. Finally, note that the approximation in equation (A1) is due to the truncation error associated with the mth order BPF signal representation, which, due to Theorem 1, converges to zero as $m \to \infty$.