IDENTIFICATION OF CONTINUOUS-TIME SYSTEMS : A TUTORIAL

H. Unbehauen* and G. P. Rao**

* *Control Engineering Laboratory, Faculty of Electrical Engineering, Ruhr-University 44780 Bochum, Germany.*
** *Engineering Systems Division, WED, Abu Dhabi, UAE (On deputation from Indian Institute of Technology, Kharagpur)*

Abstract: This paper aims at taking the reader on a guided tour of the field of identification of continuous-time systems. It presents a birds eye view of the continuous-time related aspects of the greater field of system identification. Continuous-time based contributions to system identification began in the nineteen-fifties but were overshadowed by a 'go completely digital' spirit which was spurred by parallel developments in digital computers during the following two decades. The nineteen seventies have witnessed a resurgence of continuous-time spirit and the field of continuous-time system identification has now matured to merit a review as is intended here. This paper is divided into three parts. An overview of the basic techniques of identification of continuous-time systems in a unified framework is presented in Part A. Parts B and C outline some recent developments in the identification of linear systems and nonlinear systems, respectively.

Keywords: System identification; continuous-time systems; Markov parameters; time moments; irreducible models; orthogonal functions; nonlinear systems; linear estimation; Hartley modulation functions; bilinear systems; Hammerstein model.

Part A: Overview of the general methodology in a unified framework

1. INTRODUCTION

It is essential to 'know and understand' a system before it is handled, i.e., manipulated or controlled. A system is known through modelling and identification and understood by analysis. Modelling and identification happen to be a conjugate pair of activities in the process of developing knowledge about a system. They are prerequisites to the practice of automatic control. Modelling by itself is a very vast area rich in a host of well established methods which are based on a variety of principles. Among the many variants, modelling on the basis of physical principles can hardly be overemphasized, particularly for physical systems. Application of the physical principles in modelling a physical system gives us a mathematical description with key parameters in generic form. The resulting model with generic parameters actually represents a class of models out of which the search for a particular member is conducted through the process of identification and system parameter estimation. Many surveys of the vast field of system identification may be found in the literature. For the purpose of the present discussion, two surveys in the past, one by Aström and Eykhoff (1971) in general and the other by Unbehauen and Rao (1990) with respect to continuous-time systems in particular are useful. At the outset,we briefly discuss the general premises of the two main settings in system identification, namely , the 'all digital or discrete time (DT) setting' and the 'setting which respects the inherent continuous-time spirit' and attempts to place the continuous-time (CT) aspects in the system identification perspective in order to preserve the overall picture of the field of system identification by integrating certain characteristic features needed to support continuous-time interests. The various continuous-time approaches are revisited and reviewed, with a relatively sharp focus on the developments which followed an earlier survey by the authors themselves (Unbehauen and Rao 1990).

2. THE SYSTEM IDENTIFICATION PROBLEM

The system identification problem, as defined by Zadeh (1962) in general, is depicted in Figure 1. It is characterized by three ingredients: a class of models, a class of input signals and a criterion. An attempt at the solution to the problem of system identification will be successful and the results useful if the problem is well posed in terms of the above entities. The class of models should be appropriate and the set of signals should have the property of persistent excitation relative to the model class. Usually a recursive algorithm enhances the utility of the

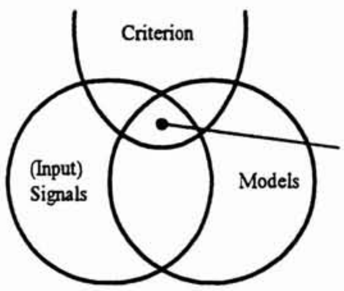


Fig. 1. The system identification problem.

method. System identification is usually not an independent exercise for its own sake; it serves a greater purpose depending on what we intend to do further with the identification results. Aström and Eykhoff (1971) and much of the literature following it, discuss the role of the three entities in the system identification problem. For instance, a classification of the main methods is based on the criterion giving rise to the output error (OE), equation error (EE), prediction error (PE) methods, etc. Different forms of model which arise due to the nature of the system require different methods of treatment. In particular, CT models call for special (and additional) signal processing considerations as will be seen later. The general setting of the CT based identification methods is shown in Fig. 2. Our discussion is thus centered only on parametric models.

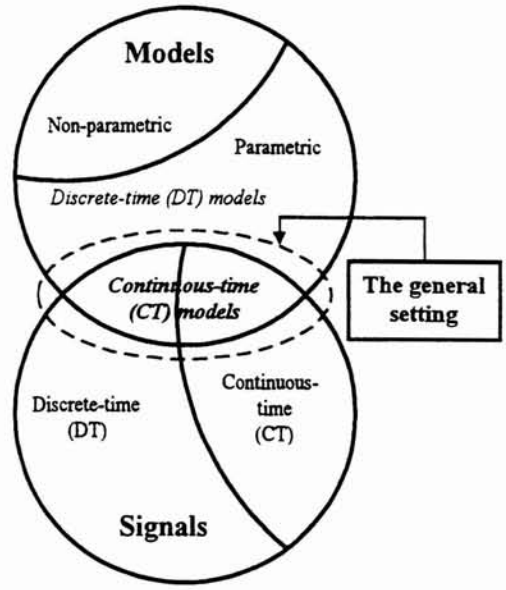


Fig. 2. The general setting for identification of continuous time systems.

The significance of continuous-time models

Much has been elaborated on the significance of CT models (Unbehauen and Rao 1987, 1990; Sinha and Rao 1991). We review below the main arguments in favour of CT models.

- Models of physical systems derived from physical principles are inherently continuous in time, because physical laws on which such modelling is based are in CT.
- CT models support a better understanding of the physical behaviour of the system under consideration. The model parameters are strongly correlated with the physical properties of the system.
- Undue sensitivity issues with respect to model parameters do not arise, which do in the event of discretization.
- Partial knowledge if present, is preserved in CT models. If a CT model containing some known parameters is discretized, these are lost in the process of discretization.
- Discretization of CT models may give rise to unnatural nonminimum phase character.
- Conventional DT methods are not in harmony with the CT spirit; in the limit of reduced sampling period, they do not converge to the results corresponding to the original CT model. The return from the conventional DT model to the original CT model is not easy.

3. REQUIREMENTS FOR CONTINUOUS-TIME MODEL IDENTIFICATION

The main difficulty in handling CT models is due to the presence of the derivative operator(s) associated with the input and output signals. While these signals are available by measurement with the attendant corruption by noise, direct generation of the required derivatives is practically undesirable. This difficulty is to be removed by preprocessing the signals in such a way that the undesirable derivative operations are favourably realized. Alternatively, the discretization of the CT model is to be made in terms of an unconventional discrete time (UDT) operator which is in harmony with its CT counterpart in the sense that the DT model converges to the original CT version as the sampling interval approaches zero. The various approaches reported in the literature may be classified with reference to the use of the ingredients in the general setting as shown in Figure 3 in which the model class is denoted as CT/UDT for reasons given above.

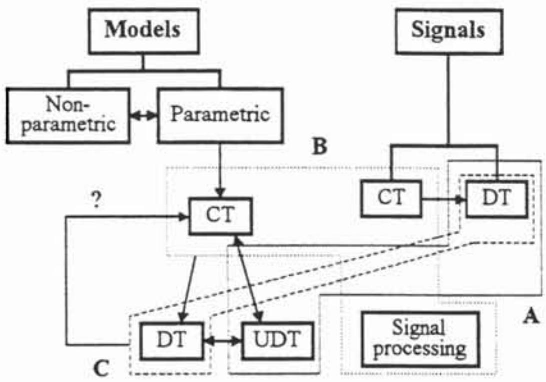


Fig. 3. Approaches to identification of continuous-time systems in the general setting.

4. CLASSIFICATION OF APPROACHES TO IDENTIFICATION OF CT SYSTEMS

The various approaches to the identification of continuous-time systems may be classified into three broad categories as indicated in Figure 3. These are based on the general framework shown in Fig.4

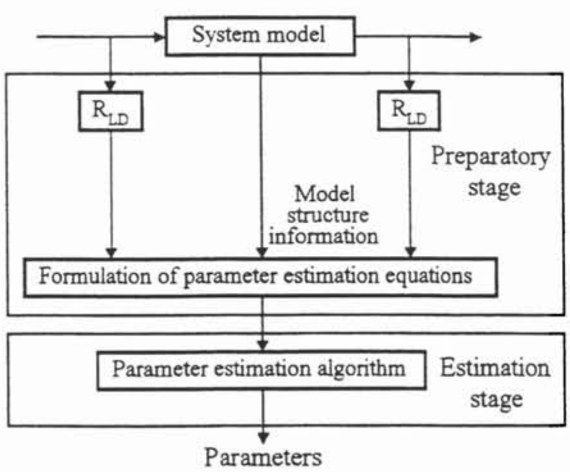


Fig.4. The general framework for identification of continuous-time systems.

According to Figure 3 these classes are:

A: Approaches using DT signals to identify a DT model which is then converted into native CT form.

B: Approaches using CT signals to directly identify a native CT model.

C: Approaches using DT signals giving rise to a UDT model which converges to its native CT.

The need to generate the time derivative terms in CT models is eliminated by a class of signal processing techniques denoted by the operation R_{LD}. The forerunner of this class of techniques is the method of modulating functions (Shinbrot 1957). Fig. 5 shows the family tree of the various signal processing methods belonging to this class.

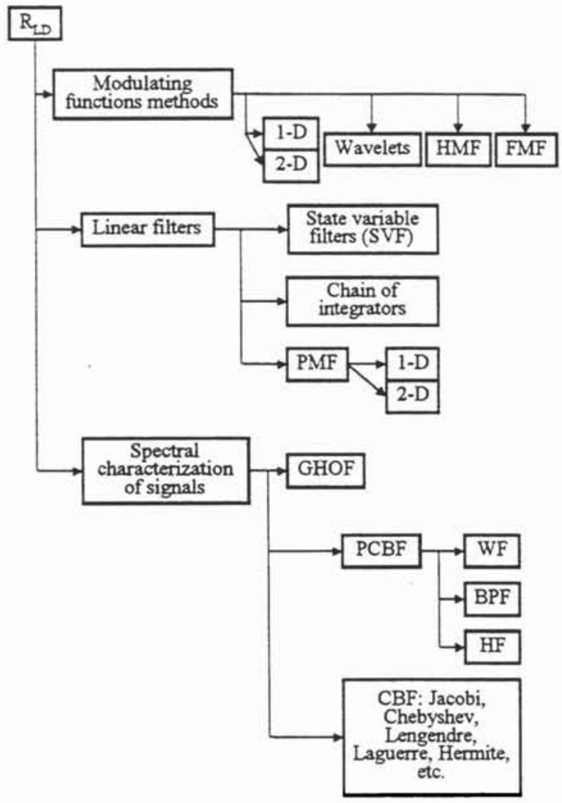


Fig. 5. Several variants of the signal preprocessing operation R_{LD} (PMF: Poisson Moment Functionals; GHOF: General Hybrid Orthogonal Functions; PCBF: Piecewise Constant Basis Functions; CBF: Continuous Basis Functions; WF: Walsh Functions; HF: Haar Functions; BPF: Block Pulse Functions; HMF: Hartley Modulating Functions; FMF: Fourier Modulating Functions).

5. CONTINUOUS-TIME MODELS OF DYNAMICAL SYSTEMS

We will consider first linear time-invariant asymptotically stable dynamical systems with input $u(t)$ and output $y(t)$. The input-output description of such a system in terms of its unknown transfer function $G(s)$, is

$$Y(s) = G(s)U(s) + N(s)\,. \tag{1a}$$

The term $G(s)U(s)$ represents the component of response of the system to the input $u(t)$, $n(t)$ represents the stochastic part of $y(t)$ and s is the complex variable $s = \sigma + j\omega$. In the case of a multi-input/multi-output (MIMO) system, $\boldsymbol{u}(t)$, $y(t)$ and $\boldsymbol{n}(t)$ represent vectors of appropriate dimensions and $\underline{G}(s)$ denotes transfer function matrix (TFM).

For models which are linear in their parameters, a generic equation of the form

[Transposed vector of measurements][Parameter vector]=[A single measurement of the output]

is first developed. Using this equation, and measurements at several instants of time, a set of equations is developed and cast in the form:

[Matrix of measurements][Parameter vector]=[Output measurement vector]

5.1 The preparatory stage

This stage involves certain operations which are required to convert the derivative terms in continuous-time models from their abstract and ideal forms $(d^k / dt^k)\{.\}$ into those which are realizable either by signal processing and/or computation. These two operations are linear; the former is dynamic and the latter is algebraic. As a simple illustration of this stage, let us consider a first order transfer function model

$$Y(s) / U(s) = G(s) = b / (as + 1)\,,$$

which corresponds to the differential equation

$$a\,dy(t) / dt + y(t) = bu(t)\,.$$

By making observations at $t_k = kT_s$, $k=1,2,3,\ldots$ and $T_s = \text{const.}$ a system of equations is generated as follows:

$$av(k) + y(k) = bu(k)\,,$$

in which $v(k)$ should be appropriately realized. Consider a set of known modulating functions:

$$\{\varphi_n(t)\}, n = 1,2,\ldots, t \in [0, t_0], \varphi_n(0) = \varphi_n(t_o) =$$
$$d\varphi_n / dt\big|_0 = d\varphi_n / dt\big|_{t_0} = 0, n = 1,2,\ldots$$

with derivatives known up to an adequate degree. In the case of the present first order model, the first derivative will suffice. Multiply the differential equation throughout by $\varphi_n(t)$, integrate over $[0, t_0]$ to get

$$a\int_0^{t_0} \varphi_n(t)\frac{dy}{dt}dt + \int_0^{t_0} \varphi_n(t)y(t)dt = b\int_0^{t_0} \varphi_n(t)u(t)dt.$$

Integrating the first term by parts and using the terminal conditions,

$$\int_0^{t_0} \varphi_n(t)y(t)dt - a\int_0^{t_0} \frac{d\varphi_n}{dt}y(t)dt = b\int_0^{t_0} \varphi_n(t)u(t)dt,$$
$$n = 1,2,\ldots.$$

The signal related terms in this equation are computable, albeit off-line. The generic *transposed vector of measurements* in this case is given as

$$[\int_0^{t_0} \frac{d\varphi_n}{dt}y(t)dt \quad \int_0^{t_0} \varphi_n(t)u(t)dt] \text{ for } n = 1,2,\ldots,$$

the *parameter vector* as $\left[a\, b\right]^T$ and the generic *measurement of the output* is

$$\int_0^{t_0} \varphi_n(t)y(t)dt\,,\ n = 1, 2, \ldots.$$

The computation in the modulating function method can be rendered on-line as a measurement by choosing $\varphi_n(t)$ as those arising out of the impulse response functions of the various stages of a filter chain having identical elements, each having a transfer function of the form $1/(s+\lambda)$. This leads to the so called Poisson moment functional (PMF) method. The PMF transformation of a signal $y(t)$ about $t=t_0$ gives the Poisson moment functionals

$$M_k\{dy/dt\} \underline{\underline{\Delta}} \int_0^{t_0} [(t_0-t)^k / k!]\exp[-\lambda(t_0-t)]\frac{dy}{dt}dt.$$

The PMF's of the derivatives of the process signals $y(t)$ and $u(t)$ can be expressed as linearly weighted sums of the PMF's of these signals themselves.

In this case, the *transposed* generic *vector of measurements* in this case is given as

$$[(M_{k-1}\{y\}t_0 - \lambda M_k\{y\}t_0 - p_k(t_0)y(0)) \quad M_k\{u\}t_0]$$

and the generic *output measurement* is $Mk\{y\}t_0$. In the above, $p_k(t)$ happens to be the inverse Laplace transform of the $1/(s+\lambda)^{k+1}$.

A set of equations may be developed either by taking PMF transformation at the minimal level and varying t_0 or by PMF transformation at different levels at a fixed t_0 or a combination of both. It is the former strategy that is usually preferred for its simplicity and possibility for on-line implementation. Higher order derivative terms of the process signals give rise to their initial values in the measurement vector. Since these are unknown, they should be separated and included in the parameter vector as additional unknowns to be estimated together with the usual system parameters. When coupled with a simulation stage, the combined algorithm becomes one of joint state and parameter estimation which is of considerable importance.

If λ is chosen to be very large relative to the time constants of the system under identification, and if the PMF transformation is taken about a large t_0, the effect of the initial conditions becomes insignificant. Consequently, the terms associated with them can be dropped from the measurement vector. The resulting algorithm estimates only the usual system parameters. However, this estimation would be at the cost of excessive passage of noise through the measurements into the estimates. The book by Saha and Rao (1983) is devoted to the PMF method and its many aspects.

PMF transformation with λ=0 is of particular significance. This leads to the so called integral equation approach. The integral operation may be automatically performed by a digital filter on sampled process signals. A typical linear integration filter extensively applied (Zhao et al 1991, 1992; Yang et al 1993) is shown in Fig. 6. This is a digital signal module in which q^{-1} denotes the shift operator and $\{p_k^j\}$ a set of weights specified by the chosen integration formula for the j-th integral.

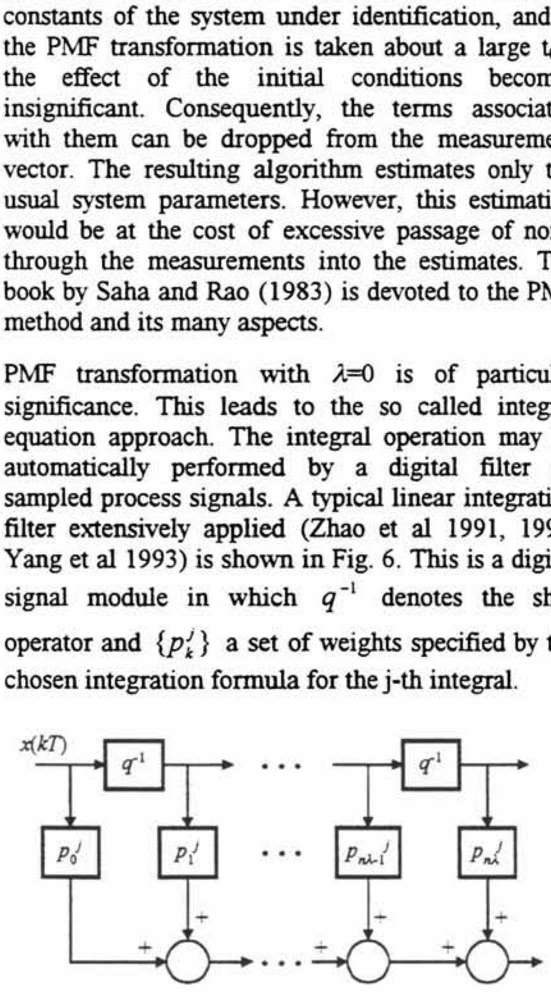


Fig. 6. The linear integrating filter.

An interesting way to realize the integrals in the 'integral equation approach' is by representing the process signals in series of orthogonal functions $\{\theta_k(t),\ k=1,\ 2, \ldots \infty\}$ over the interval $t \in [0, t_0]$ For the sake of simplicity of illustration, let us consider the first two components of the expansion in the case of all the signals involved in the example:

$$y(t) \approx y_1\theta_1(t) + y_2\theta_2(t)$$

$$u(t) \approx u_1\theta_1(t) + u_2\theta_2(t)$$

and insert them in the integral equation of the system

$$ay(t) - ay(0)s(t) + \int_0^t y(\tau)d\tau = b\int_0^t u(\tau)d\tau, t \in [0, t_0)$$

in which $s(t)$ denotes a unit step function at $t=0$, having s_1 and s_2 as its spectral components of $\theta_1(t)$ and $\theta_2(t)$. Further let

$$\int_0^t \theta_1(\tau)d\tau \approx e_{11}\theta_1(t) + e_{12}\theta_2(t)$$

$$\int_0^t \theta_2(\tau)d\tau \approx e_{21}\theta_1(t) + e_{22}\theta_2(t) \ .$$

The integral equation is transformed into algebraic form in which the *measurement matrix* becomes

$$\begin{bmatrix} y(0)s_1 - y_1 & u_1e_{11} + u_2e_{21} \\ y(0)s_2 - y_2 & u_1e_{12} + u_2e_{22} \end{bmatrix},$$

and the output *measurement vector* takes the form

$$\begin{bmatrix} y_1e_{11} + y_2e_{21} \\ y_1e_{12} + y_2e_{22} \end{bmatrix}.$$

The integral equation approach has been hosted by a wide range of systems of orthogonal functions. These include the systems of continuous functions such as Fourier, Chebyshev, Jacobi, Laguerre, Legendre, Hermite polynomial systems (Datta and Mohan 1995) and the systems of piecewise constant functions (Rao 1983) such as Walsh, Haar and block pulse functions (BPF). The class of general hybrid orthogonal functions (GHOF) proposed by Patra and Rao (1996) capture the features of continuity of the continuous systems and of discontinuities of the piecewise constant systems. The GHOF are capable of efficiently representing a wide range of signals encountered in practice, including those occurring in switched systems. In the book by Patra and Rao (1996) an extensive list of bibliography on the subject of orthogonal functions is given. The list is mapped on to different fields of applications in systems and control in separate tabular summaries.

The methods outlined above represent the process R_{LD} in the primary stage of identification of continuous systems. Following the signal processing operation, the task of obtaining the system of equations from the generic equation is accomplished by the application of certain operational matrices in an elegant algebraic framework that was developed over the seventies and eighties.

5.2 Unconventional discrete time methods

The differential equation representing the model of a continuous-time system can be discretized to get descriptions in the discrete δ - domain which has different variants depending on the choice of approximation. Table 1 shows some possibilities in which each form of δ is related to the conventional discrete time operator q^{-1} defined as $q^{-1}x(k) = x(k-1)$. Thus the δ-operator describes a DT approximation of the CT differentiation operator.

Table 1 Unconventional DT operator δ based on different approximations and its relation to the conventional backward shift operator (T_s sampling time).

	δ	**Description**
1	q^{-1}	backward shift operator
2	$\dfrac{1-q^{-1}}{q^{-1}T_s}$	backward differences based approximation
3	$\dfrac{2}{T_s}\dfrac{1-q^{-1}}{1+q^{-1}}$	trapezoidal approximation
4	$\dfrac{1-q^{-1}}{T_s}$	forward differences based approximation

The various possibilities for different methods of discretization of CT models and the interelationships among them together with their relationship with signal preprocessing R_{LD} are summarised in Fig. 7. A detailed discussion of these relationships is given in Mukhopadhyay et al (1992).From the resulting system of equations, the parameters can be estimated either by an en-bloc computation or by a recursive algorithm. The latter is made possible by virtue of the inherent nature of some signal processing methods. In particular, the BPF method and the PMF method lend themselves to recursive estimation. The δ-operator as it is known, refers only to the version based on the backward shift operator. In its more general form it is referred to as the γ-operator. The Gamma form refers to this case.

Recent developments with respect to the PMF technique include studies on the choice of the Poisson filter constant λ and its influence on the estimation. The bias distribution problem has also been studied in the pursuit of appropriate design parameters for the estimation algorithm. These are available in Bapat (1995). In the next stage of identification of continuous-time systems, standard procedures are applied as in the case of identification of discrete-time systems.

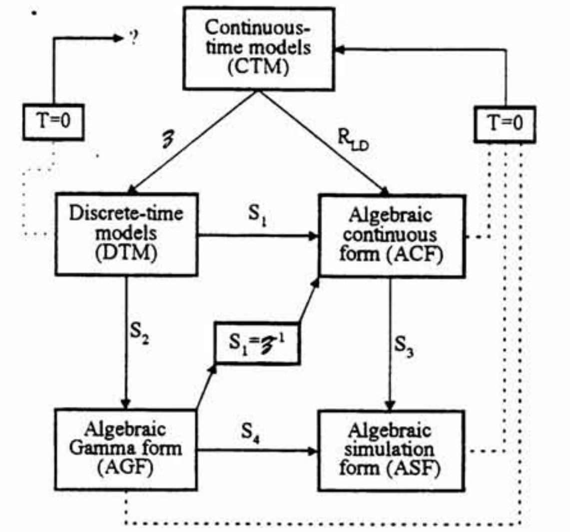


Fig. 7. Reduction of the calculus of continuous-time systems into algebra.

5.3 The estimation stage

After the preparatory stage, we now enter the estimation stage in the identification of continuous-time systems. Referring to the schemes of Fig. 8 and Fig. 9, we will discuss the various approaches in this stage.

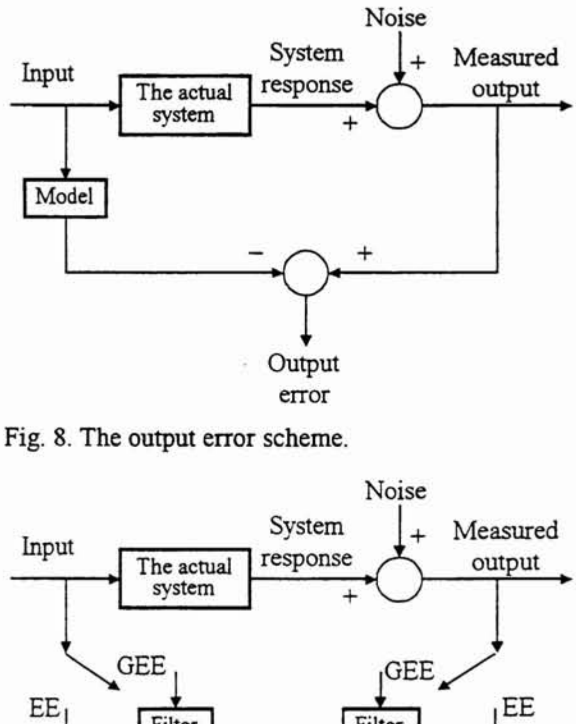


Fig. 8. The output error scheme.

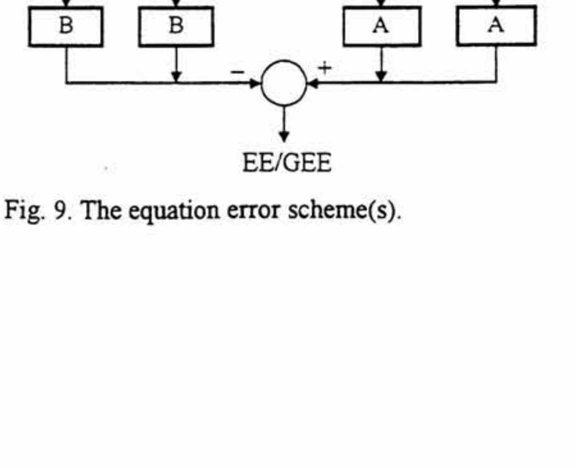


Fig. 9. The equation error scheme(s).

The model is considered in the form

$$Y(s) = G(s)U(s) + N(s)$$
$$Y(s) = G(s)U(s) + H(s)W(s) \qquad (1b)$$

in which the second term accounts for the combined effects of $n(t) = \mathcal{L}^{-1}\{N(s)\}$, unmodelled dynamics (due to model simplification) and possibly of unknown initial conditions. This term is generally called as the noise model and $w(t) = \mathcal{L}^{-1}\{W(s)\}$ denotes white noise. Table 2 summarizes the different approaches to estimation, where $H(s)$ is a filter transfer function and $C(s)$ and $D(s)$ are polynomials in s.

Table 2 Different approaches to parameter estimation

Model structure	Name	H
I	Least squares (LS), Instrumental variable (IV)	*1*
II	Extended matrix model I	*1/C*
III	Extended matrix model II	*D*
IV	General	*D/C*

Consider the set of measurements sampled at equal intervals of length T_s

$$\mathbf{Y}^N = \{u(k), y(k), k = 1,\ldots,N\} \ . \qquad (2)$$

Given $\mathbf{Y}^N$ and some prior knowledge of the dynamics of the system, the identification problem is to obtain $G(s)$ in terms of its parameters which best describes the dynamics of the system in some sense by minimizing a chosen norm of the modelling error.

Let $G(s,\theta)$ denote an estimate of the transfer function in which $\theta \in \mathbf{R}^n$ is the parameter vector. In terms of $G(s,\theta)$, the input-output description becomes

$$y(t) = \mathcal{L}^{-1}\{G(s,\theta)U(s) + H(s,\theta)W(s)\}, \qquad (3)$$

in which the second term accounts for the combined effects of $n(t)$, unmodelled dynamics (due to model simplification) and possibly of unknown initial conditions. This term is generally called as the noise model. The focus of our attention in the present treatment is on the first term in (3). The treatment is also applicable to the second term in (3).

A general discrete time parametric approximation for (3) is the polynomial black-box model

$$A(\delta,\theta)y(k) = \frac{B(\delta,\theta)}{F(\delta,\theta)}u(k) + \frac{C(\delta,\theta)}{D(\delta,\theta)}e(k), \qquad (4)$$

that is the DT approximation of the CT differential equation of the DT Box-Jenkins (1970) version (Unbehauen and Rao 1987). In this, δ denotes a DT approximation to the CT differentiation operator $p = d/dt$ and not the usual backward shift operator, $u(k)$ and $y(k)$ are the samples of input and output signals respectively, and $e(k)$ is a sequence of independent and uniformly distributed zero mean random variables. $A(\delta,\theta)$, $B(\delta,\theta)$, $C(\delta,\theta)$, $D(\delta,\theta)$, and $F(\delta,\theta)$ are polynomials in δ whose coefficients are arranged to form the parameter vector θ. Specific cases of these polynomials lead to particular models such as auto-regressive (AR), moving average (MA), auto-regressive and moving average (ARMA) and so on. In particular, to characterize stationary stochastic processes, the following ARMA model is considered.

$$A(\delta,\theta)y(k) = C(\delta,\theta)e(k), \qquad (5)$$

where

$$A(\delta,\theta) = \delta^{n_A} + a_1\delta^{n_A-1} + \ldots + a_{n_A},$$

and

$$C(\delta,\theta) = c_0\delta^{n_{CA}} + c_1\delta^{n_C-1} + \ldots + c_{n_C},$$

whose coefficients appear in the AR and MA portions respectively of the model. In the present context, the terms AR, MA and ARMA refer to the CT context, where δ denotes a DT approximation to the CT differential operator and not the usual backward shift operator.

The ARMA model of (5) is commonly used in spectral estimation and time-series analysis. In system identification, where the goal is to characterize the dynamic input-output relation of the underlying process, the following model is suitable:

$$A(\delta,\theta)y(k) = B(\delta,\theta)u(k), \qquad (6)$$

where

$$A(\delta,\theta) = \delta^n + a_1\delta^{n-1} + \ldots + a_n, \text{ and}$$
$$B(\delta,\theta) = b_1\delta^{n-1} + b_2\delta^{n-2} + \ldots + b_n.$$

Here the MA part is formed from the usually known process input signal. Without resorting to create a new name, we refer to

$$G(\delta,\theta) = \frac{B(\delta,\theta)}{A(\delta,\theta)}, \qquad (7)$$

as "deterministic ARMA". This is nonlinear in the parameters. With this model structure, the model output error (OE) in sampled form is

$$\varepsilon_{OE}(k) = y(k) - \frac{B(\delta,\boldsymbol{\theta})}{A(\delta,\boldsymbol{\theta})}u(k). \tag{8}$$

A parameter estimation criterion is to minimize

$$J_{OE}(k) = \sum_{k=1}^{N} \varepsilon_{OE}^2(k) \tag{9}$$

with respect to θ. Since the output error of (8) is nonlinear-in-parameters, this is a case of nonlinear optimization. In an attempt to simplify the situation, most of the identification approaches resort to the equation error (EE).

$$\varepsilon_{EE}(k) = \frac{A(\delta,\boldsymbol{\theta})}{E(\delta)}y(k) - \frac{B(\delta,\boldsymbol{\theta})}{E(\delta)}u(k) \tag{10}$$

and a criterion

$$J_{EE}(k) = \sum_{k=1}^{N} \varepsilon_{EE}^2(k). \tag{11}$$

Here $\frac{1}{E(\delta)}$ is a linear-dynamic operator, of adequate order for the removal of the need for direct differentiation of process data (Unbehauen and Rao 1987). These operators also serve the purpose of prefilters used for removing unimportant frequencies from the process data. Since (10) is linear-in-parameters, parameter estimation is simplified to linear recursive least-squares (RLS) estimation. However, EE minimization has its disadvantages.

- **Biased estimation:** The parameter estimates will be biased when the EE is not white(Eykhoff 1974). Variants of the ordinary least-squares (LS) algorithm such as "generalized least-squares" and "instrumental vairables" (Söderström and Stoica 1989) are applied to remove the bias. These and other "bias compensating least-squares" methods (Zhao et al. 1991, 1992; Yang et al. 1993) are computationally demanding (Mukhopadhyay et al. 1991). These approaches assume that the measurements are actually generated by an ARMA model, and that the measurement noise is Gaussian distributed. However, the performance of some of these may not be satisfactory when there is significant modelling error, as this component of error may not be Gaussian distributed.

- **Reducible models (for MIMO systems):** Consider a v_i-input v_o-output system described by the "transfer function" matrix (TFM)

$$\underline{G}(\delta,\boldsymbol{\theta}) = \begin{bmatrix} G_{11}(\delta,\boldsymbol{\theta}) & G_{12}(\delta,\boldsymbol{\theta}) & \cdots & G_{1v_i}(\delta,\boldsymbol{\theta}) \\ G_{21}(\delta,\boldsymbol{\theta}) & G_{22}(\delta,\boldsymbol{\theta}) & \cdots & G_{2v_i}(\delta,\boldsymbol{\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ G_{v_o1}(\delta,\boldsymbol{\theta}) & G_{v_o2}(\delta,\boldsymbol{\theta}) & \cdots & G_{v_ov}(\delta,\boldsymbol{\theta}) \end{bmatrix} \tag{12}$$

and

$$G_{ij}(\delta,\boldsymbol{\theta}) = \frac{B_{ij}(\delta,\boldsymbol{\theta})}{A_{ij}(\delta,\boldsymbol{\theta})}.$$

EE formulation necessitates a canonical form having a least common denominator (CD) of all the elements of the TFM. The CD considerably inflates the unknown parameter vector. To reduce this inflation partially, the TFM is decomposed into multiple-input/single-output (MISO) sub-models with several CDs limited only to the rows of the TFM. In this way, a two-stage algorithm was proposed by Diekmann and Unbehauen(1979) for DT model identification, and its CT version by Mukhopadhyay et al. (1991). A Gauss-Seidel type iterative algorithm that does not require a CD was later suggested by Rao et al. (1984). The approaches to estimation of irreducible CT models are surveyed in (Mukhopadhyay et al. 1991, Mukhopadhyay and Rao 1991).

- **Distribution of estimation errors**: Modelling of physical processes is usually associated with certain amount of undermodelling. This coupled with noise in the measurements, results in biased estimates. Though it is possible to eliminate bias due to measurement noise, the bias resulting from undermodelling can not be eliminated at all. It can be distributed over a range of frequencies by careful design of the identification experiment (Wahlberg and Ljung 1986) such that such undermodelling is not harmful in the context of the final application of the resulting model. With ARMA modelling, the problem of experiment design for a prescribed distribution of bias over a range of frequencies is not simple and straight forward.

Using Parseval's theorem, the frequency domain description of the EE criterion (11) in the limit as $T_s \to \infty$, is

$$J_{EE}(\omega) = \int_0^{\infty} \left| \frac{A(j\omega,\boldsymbol{\theta})}{E(j\omega)} U(j\omega) \right|^2 \cdot \left| G^0(j\omega) - \frac{B(j\omega,\boldsymbol{\theta})}{A(j\omega,\boldsymbol{\theta})} \right|^2 d\omega \tag{13}$$

where $U(j\omega)$ is the Fourier transform of the input signal and G^0 denotes the true model. The first term on the right-hand side of (13) is a weighting function that manipulates the second term (bias) over a range of frequencies. With the chosen ARMA model structure, it is clear that this weighting function is a function of the yet-unknown $A(\delta, \theta)$ which renders on-line experiment design as impossible. Off-line design, however, is shown to be possible by Bapat (1995).

Part B: Recent developments in linear continuous-time system identification

1. MODELS FOR LINEAR ESTIMATION (MOVING AVERAGE FORMS)

Recent developments are in the direction of linearizing the estimation problem, $G(\delta,\theta)$ is linear if the second and higher order derivatives of $G(\delta,\theta)$ with respect to θ vanish for all θ and the linearity of a parametrization is different from the linearity of the model in terms of its input-output behavior. Even nonlinear models can be linearly parametrized.

One situation in which the ARMA model of (6) is linearized (with respect to θ) is when its denominator $A(\delta,\theta)$ is fixed as some appropriate $A(\delta)$ which leads to the description

$$G(\delta,\theta) = \sum_{i=1}^{n-1} \frac{b_i \delta^{n-i}}{A(\delta)} = \theta^T \mathscr{B}(\delta) \tag{14}$$

in which

$$\theta = [b_1 b_2 \ldots b_n]^T \text{ and } \mathscr{B}(\delta) = \frac{1}{A(\delta)}\left[\delta^{n-}{}_1\delta^{n-2}\ldots\delta^0\right]$$

A linear-in-parameters model is therefore obtained as

$$Y(\delta) = \frac{B(\delta,\theta)}{A(\delta)} U(\delta), \tag{15}$$

whereby EE = OE and estimation (minimization) is linear. This leads to an advantageous situation with the following possibilities:

- **Robust estimation**: In the limit as $N \to \infty$, the LS estimate $\hat{\theta}$ in the presence of zero-mean disturbances tends to $\hat{\theta}^*$, where $\hat{\theta}^*$ is the limiting estimate in the absence of disturbances. In particular, if the disturbance term is Gaussian and there is no modeling error, the LS estimate $\hat{\theta}$ is asymptotically normal with mean $\hat{\theta}^*$ and a covariance proportional to the variance of the disturbance. This holds good even for coloured disturbances uncorrelated with the input. This implies that the LS estimation is robust to zero-mean disturbances. Note that the estimates will still be "biased" due to the inherent under-modelling.

- **Irreducible model estimation**: With MIMO TFM models, since the denominators do not include unknown parameters, the CD formulation does not inflate the parameter vector.

- **Simplified error distribution problem**: The weighting function in (13) now equals $\left|\frac{A(j\omega)}{E(j\omega)}U(j\omega)\right|^2$. The absence of the unknown θ in this weighting function permits on-line experiment design for a prescribed bias distribution.

- **Gray-box modeling**: The fixed denominator polynomial $A(\delta)$ in the linear-in-parameters model (14) serves as an additional design variable allowing for effective incorporation of prior knowledge of the process dynamics. By an intelligent choice of this polynomial, even complex systems can be estimated significantly accurately with a smaller number of parameters.

These are the advantages of linear-in-parameters models in system identification. In these models the output is expressed as a linear combination of certain MA components of the input. This leads to the "generalized moving average model" (GMAM) formulation as

$$Y(\delta) = \left[\sum_{i=1}^{n} \theta_i \mathscr{F}_i(\delta)\right] U(\delta). \tag{16}$$

In this model, the moving-average components of the inputs are formed as the responses of a set of known filters $\{\mathscr{F}_i(\delta)\}$ to $u(k)$. These filters form the basis

$$\mathscr{B}(\delta) = [\mathscr{F}_1(\delta)\mathscr{F}_2(\delta)\ldots\mathscr{F}_n(\delta),]^T \tag{17}$$

of the GMAM structure

$$G(\delta,\theta) = \sum_{i=1}^{n} \theta_i \mathscr{F}_i(\delta). \tag{18}$$

With such a parametrization, the model output error

$$\varepsilon_{OE}(\delta) = y(k) - \sum_{i=1}^{n} \theta_i \mathscr{F}_i(\delta) U(\delta)$$

is linear in $\{\theta_i\}$, and consequently the minimization problem of the output error criterion (9) is linear.

Such models evolve very naturally from truncated power series expansions of the rational transfer function. For example, in the DT case, the system transfer function in the complex variable z^{-1} may be written as

$$G^0(z^{-1}) = \sum_{i=1}^{\infty} h_i z^{-i} , \quad (19)$$

where $\{h_i\}$ is the impulse response sequence. This implies that

$$G(\delta,\theta) = \theta^T \mathcal{B}(z^{-1}) , \quad (20)$$

where
$\theta = [\theta_1\theta_2\ldots\theta_n]^T$, and $\mathcal{B}(z^{-1}) = [z^{-1}z^{-2}\ldots z^{-n}]^T$
The quality of this approximation depends on the rate of convergence of the impulse response sequence. The poles of $G^0(z^{-1})$ close to the unit circle in the z-domain slow down the rate of convergence. Consequently a high model order for n is required for a given tolerance. For these reasons in rapidly sampled CT systems the rate of convergence of the approximation will be very slow, and in the limit as the sampling time $T_s \to 0$, the DT poles approach unity and consequently the approximation fails to converge. Furthermore, even in the case of convergent approximations, high model order is required as the memory of the basis (shift operator) is very short (unity). Therefore, model representations having better convergence properties and less sensitivity to sampling rate will be preferable.

In the CT case, the transfer function $G(s)$ may be expanded about $s \to \infty$ as a complex power series in s^{-1} as

$$G^0(s) = \sum_{i=1}^{\infty} h_i (s^{-1})^i \quad (21)$$

leading to the form

$$G(s,\theta) = \theta^T \mathcal{B}(s^{-1}) , \quad (22)$$

where
$\theta = [h_1h_2\ldots h_n]^T$ and $\mathcal{B}(s^{-1}) = [s^{-1}s^{-2}\ldots s^{-n}]^T$. It is well known that h_i are the CT Markov parameters (MP) of $G^0(s)$ which are defined as

$$h_i = \left. \frac{d^{i-1}}{dt^{i-1}} g^0(t) \right|_{t=0} \quad (23)$$

where $g^0(t)$ is the impulse response of $G^0(s)$.

Considering a similar expansion of $G^0(s)$ about s=0, one has models parametrized in terms of normalized time moments of the impulse response $g^0(t)$ of $G^0(s)$, i.e.,

$$G(s,\theta) = \theta^T \mathcal{B}(s) , \quad (24)$$

where $\theta = [m_1m_2\ldots m_n]_T$ and $\mathcal{B}(s) = [s\,s_2\ldots s^n]^T$
and

$$m_i = \frac{(-1)^i}{i!} \int_0^{\infty} t^i g^0(t)dt \quad (25)$$

are the normalized time moments.

Other basis functions are also possible. Well known among these are Laguerre and Kautz filters. Laguerre filters imply a basis,

$$\mathcal{B}_{LAG}(s) = \left[\frac{1}{s+\lambda} \frac{1}{s+\lambda}\left(\frac{s-\lambda}{s+\lambda}\right) \ldots \frac{1}{s+\lambda}\left(\frac{s-\lambda}{s+\lambda}\right)^{n-1} \right]^T ,(26)$$

with λ>0, and Kautz filters imply

$$\mathcal{B}_{KAUTZ}(s) = [\psi_1(s)\psi_2(s)\ldots\psi_n(s)]^T \quad (27)$$

where

$$\psi_{2k-1}(s,b,c) = \frac{s}{s^2+bs+c}\left[\frac{s^2-sb+c}{s^2+sb+c}\right]^{k-1} ,$$

and

$$\psi_{2k}(s,b,c) = \frac{1}{s^2+bs+c}\left[\frac{s^2-bs+c}{s^2+bs+c}\right]^{k-1} ,$$

with $b > 0$, $c > 0$, and $k = 1,2, \ldots$. Wahlberg (1994) discusses these bases in greater detail. The role played by the basis in continuous and discrete system modelling is discussed by Goodwin et al (1991).

2. MARKOV PARAMETER MODELS

However, in CT situation, the reference to Markov parameters is rare. This is because of the natural but difficult-to-compute form (23) in which Markov parameters are defined for CT systems.

The work of Dhawan et al. (1991) is the first attempt with MP models for SISO CT model identification. The MP model (22) is transformed into an integral equation in which the integrals are realized using block-pulse functions (Rao 1983), thereby avoiding the derivative route to the realization of Markov parameters. However, truncation of the MP model as in (22) often leads to poor approximation, due to which the estimation may fail to converge. A simple generalization of the original MP model to ensure convergent approximations may be found in Dhawan et al. (1991) and Küper (1992). Further generalization of the basis leading to flexible and well-behaved approximations was suggested by Subrahmanyam and Rao (1993) and has been extended to MIMO systems by Subrahmanyam et al. (1996).

2.1 Estimation of moving-average models

Consider a v_i-input, v_o-output MIMO system having a TFM $\underline{G}^0(s)$, and an input-output relationship

$$Y(s) = \underline{G}^0(s)U(s) + V(s)\,, \tag{28}$$

where $y \in \mathbf{R}^{v_o}, u \in \mathbf{R}^{v_i}$ and $v \in \mathbf{R}^{v_o}$. CT Markov parameters of this system are defined as the coefficients of the power series

$$\underline{G}^0(s) = \sum_{l=1}^{\infty} \boldsymbol{H}_l s^{-1},$$

where $\{\boldsymbol{H}_l\}$ is the Markov parameter sequence (MPS). Denote

$$\boldsymbol{H}_l = \begin{bmatrix} h_{l,11} & h_{l,12} & \cdots & h_{l,1v_i} \\ h_{l,21} & h_{l,22} & \cdots & h_{l,2v_i} \\ \vdots & \vdots & \ddots & \vdots \\ h_{l,v_o1} & h_{l,v_o2} & \cdots & h_{l,v_ov_i} \end{bmatrix}.$$

In terms of the MPS, and time-domain notation

$$y(t) = \sum_{l=1}^{\infty} \boldsymbol{H}_l \boldsymbol{u}^l(t) + v(t),$$

where $\boldsymbol{u}^l(t)$ is the l-th integral of $\boldsymbol{u}(t)$. Assuming absolute convergence of the MPS and thus uniform convergence of partial sums, a truncated MP model is obtained as,

$$y(t) = \sum_{l=1}^{n} \boldsymbol{H}_l u^l(t) + \boldsymbol{e}(t) \tag{29}$$

where $\boldsymbol{e}(t)$ includes the truncation (of the MPS) error and the contribution of unknown initial conditions in addition to the usual noise term $\boldsymbol{v}(t)$. This model is valid only when the power series expansion of $\underline{G}^0(s)$ is absolutely convergent. Note that, when the system is represented in the sampled domain as $\underline{G}^0(z^{-1})$, the resulting DT MPS is the impulse response sequence of the system. For asymptotically stable systems, the DT MPS is absolutely convergent. But, when represented in the CT domain, even stable systems may have diverging MP sequences. To ensure absolute convergence and to increase the rate of convergence of the approximation, a more general version of Markov parameters, called as Markov-Poisson parameters was suggested by Subrahman-yam and Rao(1993). In terms of these, $\underline{G}^0(s)$ is expanded as

$$\underline{G}^0(s) = \sum_{l=1}^{\infty} \overline{\boldsymbol{H}}_l \left(\frac{\beta}{s+\lambda}\right)^l U(s). \tag{30}$$

The Markov-Poisson parameters $\{\overline{\boldsymbol{H}}_l\}$ are related to the Markov parameters $\{\boldsymbol{H}_l\}$ as

$$\overline{\boldsymbol{H}}_l^* = \frac{1}{\beta^l} \sum_{i=1}^{l} {}^{l-1}C_{i-1} \lambda^{l-i} \boldsymbol{H}_i, \quad l = 1,\ldots; \beta > 1. \tag{31}$$

Thus, the model is

$$y(t) = \mathcal{L}^{-1}\left\{\sum_{l=1}^{n} \overline{\boldsymbol{H}}_l \left(\frac{\beta}{s+\lambda}\right)^l U(s)\right\} + e(t),$$

implying the basis

$$\mathcal{B}_{PF}(s) = \left[\frac{\beta}{s+\lambda}\left(\frac{\beta}{s+\lambda}\right)^2 \cdots \left(\frac{\beta}{s+\lambda}\right)^n\right]^T.$$

The elements of $\mathcal{B}_{PF}(p)$ are the well-known Poisson filters (Saha and Rao 1983) of increasing order in which λ and β are tunable parameters. This generalization improves the low-frequency predictive ability of the model. The choice of the filter parameter λ has to be made according to the *a priori* knowledge of the poles of the system. In general, a $\lambda > 0$ is well suited for over-damped systems with poles not very close to the imaginary axis of the s-plane. On the other hand, a $\lambda<0$ with a large β is appropriate when the (complex) poles of the system are arbitrarily close to the imaginary axis.

Parameter estimation may now be carried out by decomposing the problem into v_o MISO sub-problems and considering one sub-problem at a time or in parallel. In the sequel, one such MISO problem is considered and the subscript i, that denotes the row index, is dropped mainly for notational simplicity. Further, only n_j parameters are considered for the j-th element of the MISO problem. Approximating the derivative operator by δ in the parameter estimation equation we get

$$\tilde{y}(k) = \boldsymbol{\phi}^T(k)\boldsymbol{\theta}\,,$$

where

$$\phi^T(k) = [\varphi_1^T(k)\varphi_2^T(k)\ldots\varphi_{v_i}^T(k)],$$

$$\varphi_j^T(k) = [\mathcal{F}_{1_j}(\delta)u_j(k)\mathcal{F}_{2_j}(\delta)u_j(k)\ldots\mathcal{F}_{n_j}(\delta)u_j(k)]$$
$$j = 1,\ldots,v_i,$$

$$\theta = \left[\bar{h}_{1,1}\ldots\bar{h}_{n_1,1}\middle|\ldots\middle|\bar{h}_{1,v_i}\ldots\bar{h}_{n_{v_i},v_i}\right]^T,$$

and

$$\mathcal{F}_l(\delta) = \left(\frac{\beta}{s+\lambda}\right)^l.$$

Next, we define the cost function as

$$J(\theta) = [\theta - \hat{\theta}(0)]^T P^{-1}(0)[\theta - \hat{\theta}(0)] + \sum_{k=1}^{N}[y(k) - \phi^T(k)\theta].$$

The LS estimate that minimizes $J(\theta)$ is

$$\hat{\theta}(N) = \left[P^{-1}(0) + \sum_{k=1}^{N}\phi(k)\phi^T(k)\right]^{-1} \bullet \left[P^{-1}(0)\hat{\theta}(0) + \sum_{k=1}^{N}\phi(k)y(k)\right],$$

provided the inverse exists. This estimate may be calculated using the conventional recursive least-squares algorithm.

2.2 Irreducible ARMA model realisation

Given the estimates of $\bar{H}_l, l = 1,\ldots,n,$ the first step towards realization of an irreducible ARMA TFM model is to examine the columns or rows of the (pxq)-dimensional Hankel matrix $\mathcal{H}$ formed from the estimates as

$$\mathcal{H} = \begin{bmatrix} \bar{H}_1 & \bar{H}_2 & \ldots & \bar{H}_q \\ \bar{H}_2 & \bar{H}_3 & \ldots & \bar{H}_{q+1} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{H}_p & \bar{H}_{p+1} & \ldots & \bar{H}_{q+p-1} \end{bmatrix} \quad (32)$$

for predecessor independence (Gantmacher 1959, Wolovich 1974). In view of the MISO decomposition,

$$\bar{H}_l = \left[\bar{h}_{l,1}\bar{h}_{l,2}\ldots\bar{h}_{l,v_i}\right].$$

Interchanging columns, equation (32) may be written as

$$\mathcal{H} = [\mathcal{H}_1\mathcal{H}_2\ldots\mathcal{H}_{v_i}],$$

where $\mathcal{H}_j$, $j = 1,2,\ldots,v_i$ are the $p\times q$-dimensional Hankel matrices of the SISO elements of the MISO sub-model.

Thus, the problem of structural identification of the MISO model is also decomposed into equivalent problems of finding ranks of Hankel matrices of individual elements over a row. Singular value decomposition may be used for this purpose.

According to the "partial realization theory" (Tether 1970, Kalman 1970), given a finite sequence of Markov parameters, it is possible to find a finite dimensional realization whose first few Markov parameters are correspondingly equal to the given finite sequence of Markov parameters. Accordingly, given a finite estimate Markov-Poisson parameter sequence, irreducible TFM models can be derived solving the following equations together with (31):

$$h_{l,ij} = b_{l,ij} - \sum_{r=0}^{l-1} h_{r,ij}a_{l-r,ij}, \quad l = 1,2,\ldots,n_{ij},$$

$$h_{l+n_{ij},ij} = -\sum_{r=1}^{n_{ij}} h_{n_{ij}+l-r,ij}a_{r,ij}, \quad l = 1,2,\ldots$$

where the ijth element of the TFM is considered to be of the form

$$G_{ij}(s) = \frac{b_{1,ij}s^{n_{ij}-1}+\ldots+b_{n_{ij},ij}}{s^{n_{ij}} + a_{1,ij}s^{n_{ij}-1}+\ldots+a_{n_{ij},ij}}.$$

Supposing the system is of this ARMA form, some insight may be given regarding the nature of the MPS:

- The MPS is convergent when all the poles of all the elements of the TFM are inside the unit circle centered at the origin of the s-plane. Equivalently, the Markov-Poisson parameter sequence (MPPS) is convergent when all the poles of all the elements of the TFM are inside the circle of radius β centered at $(-\lambda,0)$ of the pole-zero plot. This circle may be termed as the zone of convergence of the sequence.
- The MPS (or MPPS) is finite if and only if all the poles of all the elements of the TFM lie at the origin of the zone of convergence.

2.3 Finization of MPS

The usual infinite-length MPS is finite (with length $\max_{i,j}\{n_{ij}\}$ when no CD is assumed, or $\max_j\{n_j\}$ when column-wise CD is assumed) only when the poles of each subsystem of the TFM lie at the origin of the convergence zone. For a known system, all poles can be placed at the center by state feedback. Then such a modified system will have a finite MPS. In the identification problem, since such state feedback can not be introduced as the system itself is unknown, it is possible to introduce the effect of pole-placement on the input-output measurement data, by some iterative pole-placement algorithm. For the sake of simplicity, the SISO case is considered in the following.

Consider the state equation of $G^0(s)$ in its controllable canonical form

$$\dot{x}(t) = Ax(t) + bu(t)$$

$$\text{where } A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}$$

and $b = [0\,0 \ldots 0\,1]^T$.

The matrix A can be written as

$$A = A_0 - bk^T$$

where

$$A_0 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \text{and } k^T = [a_n a_{n-1} \ldots a_1].$$

Therefore we have the state equation as

$$\dot{x}(t) = A_0 x(t) + b\bar{u}(t),$$

where $\bar{u}(t) = u(t) - k^T x(t)$ is the filtered input signal. The fictitious system described as above by the signal pair $\bar{u}(t)$ and $x(t)$ has a finite MPS, as the eigenvalues of A_o are all at zeros. Therefore, by transforming the original system into that described by the above, the approximation error due to truncation of the MPS can be made to vanish. This is equivalent to placing the poles of the system at the origin of the convergence zone. Based on this, a time-recursive and iterative algorithm was initially proposed by Subrahmanyam and Rao (1993) and extended to MIMO systems (Subrahmanyam et al. 1996).

3. TIME MOMENT MODELS

Like Markov parameters, time moments also play an important role in the field of reduced-order modelling. Despite the wealth of other mathematically sound methods available for reduced order modelling, the moment matching method is still considered as the simplest and is widely used. In the field of system identification, an approach for multivariable system identification has been recently proposed by Subrahmanyam et al. (1995).

3.1 Estimation of moving-average models

The TFM may be written in terms of the time moments which are related to the impulse response as

$$\underline{G}^0(s) = \int_0^\infty g^0(t)e^{-st}dt = \sum_{l=0}^\infty M_l s^l ,$$

where

$$M_l = \frac{(-1)^l}{l!}\int_0^\infty t^l g^0(t)dt, \quad l = 1,2,\ldots$$

happen to be the normalized time moments of the impulse response. Define

$$H_l = \begin{bmatrix} m_{l,11} & m_{l,12} & \ldots & m_{l,1v_l} \\ m_{l,21} & m_{l,22} & \ldots & m_{l,2v_l} \\ \vdots & \vdots & \ddots & \vdots \\ m_{l,v_o 1} & m_{l,v_o 2} & \ldots & m_{l,v_o v_l} \end{bmatrix}.$$

In terms of the time moment sequence (TMS) $\{M_l\}$, the system input-output relation becomes

$$y(t) = \sum_{l=0}^\infty M_l u^{(l)}(t) + v(t),$$

where $u^{(l)}(t)$ is the l-th derivative of $u(t)$. Assuming absolute convergence of TMS and thus uniform convergence of partial sums, similar to the case of MP modelling, the truncated TM model is,

$$y(t) = \sum_{l=0}^n M_l u^{(l)}(t) + e(t). \tag{33}$$

To validate the use of the above model even for systems with diverging TMS, additional exponential scaling of the series will be necessary to ensure convergence.

To avoid the direct use of derivatives, (33) is operated on both sides by a $(n+1)$th order Poisson filter $\beta^{n+1}/(s+\lambda)^{n+1}$ (Saha and Rao 1983). Denoting

$$\mathcal{F}_{l,n+1}(s) = \beta^{n+}{}_1 \frac{s^l}{(s+\lambda)^{n+1}}, \quad l = 0,1,\ldots,n,$$

the time moment (TM) model is

$$\mathcal{F}_{0,n+1}(s)Y(s) = \sum_{l=0}^{n} M_l \mathcal{F}_{l,n+1}(s)U(s) + E(s). \qquad (34)$$

For the i-th row of (34) (dropping the subscript i in all relevant symbols), taking into account n_j time moments of the jth MISO subsystem, and letting $n = \max_j \{n_j\}$, the parameter estimation equation in discrete-time is obtained as

$$\widetilde{y}^*(k) = \phi^T(k)\theta,$$

where

$$\phi(k) = \left[\varphi_1(k)\ \varphi_2(k) \ldots \varphi_{v_i}(k)\right]^T,$$

$$\varphi_j(k) = [F_{0,n+1}(\delta)u_j(k)\ F_{1,n+1}(\delta)u_j(k)\ldots$$
$$F_{n_j,n+1}(\delta)u_j(k)]^T, \quad j = 1,\ldots,v_i,$$

and

$$\theta = \left[\ m_{0,1},\ldots,m_{n_1,1} \middle| \ldots \middle| m_{0,v_i},\ldots,m_{n_{v_i},v_i}\ \right]^T.$$

Parameter estimation may now be carried out with the usual least-squares algorithm.

3.2 Irreducible ARMA model realization

Given the estimates of M_l, $l = 1,\ldots,n$, an irreducible ARMA TFM model can be realized in a manner similar to the case of Markov parameter models. Let

$$A_{ij}(s) = 1 + a_{1,ij}s + \ldots + a_{n_{ij},ij}s^{n_{ij}},$$

$$B_{ij}(s) = b_{0,ij} + b_{1,ij}s + \ldots + a_{n_{ij}-1,ij}s^{n_{ij}-1},$$

and

$$M_l = \left\{\ m_{l,ij}; i = 1,\ldots,v_o, j = 1,\ldots,v_i\ \right\}.$$

Given the estimates of M_l, $l = 1,\ldots,n$, the TFM elements can be obtained by solving the following equations.

$$m_{l,ij} = b_{l,ij} - \sum_{r=0}^{l-1} m_{r,ij}a_{l-r,ij}, \quad l = 0,\ldots,n_{ij}-1,$$

and

$$m_{l+n_{ij},ij} = -\sum_{r=1}^{n_{ij}} m_{n_{ij}+1-r,ij}a_{l,ij}, \quad l = 1,2,\ldots.$$

Supposing the system is of this ARMA form, the following remarks are in order:

1. The TMS is convergent if all the poles of all the elements of the TFM are outside the unit circle centered at the origin of the s-plane. This circle is the zone of the convergence of the sequence.

2. The TMS is finite if and only if all the elements of the TFM are denominator free (i.e. have denominator as 1).

For this special case, Subrahmanyam et al. (1995) proposed an iterative algorithm that finitizes the sequence so as to eliminate the truncation error.

3.3 Finitization of TMS

The TMS is finite when all the subsystems of the TFM are denominator free, in which case the length of the TMS is $\max_j\{n_j\}$ and modelling will not involve unmodelled dynamics. This situation can be met by adding fictitious zeros to each subsystem, to cancel their respective denominators. In an identification experiment, this is achievable for ARMA systems as illustrated below for the SISO case

$$Y(s) = \frac{B(s)}{A(s)}U(s).$$

If the denominator $A(s)$ is known, we can write

$$Y(s) = \frac{B(s)}{A(s)}U(s) = \sum_{i=1}^{n-1} b_i s^i \overline{U}(s),$$

where $\overline{U}(s) = \frac{1}{A(s)}U(s)$. Thus the model between $\overline{u}(t)$ and $y(t)$ has a finite TMS. Therefore, by estimating the denominators and then canceling them, in an iterative way, it is possible to finitize the TMS, so that the truncation error is removed iteratively. Such an iterative algorithm with detailed analysis was presented by (Subrahmanyam et al. 1995).

4. CHOICE OF PARAMETRIC FORM

In reality, modelling error is inevitable and the performance (viz. predictive ability) of the estimated models depends on the choice of model structure and the prior knowledge embedded into the chosen model structures, for a given model order. In the

GMAM structure, the following variants are considered for CT system modeling:

◊ Motivated by Markov-Poisson parameter models, with a Poisson filter chain:

$$\mathcal{B}_{PF}(\delta) = \left[\frac{\beta_c}{\delta + l}\left(\frac{\beta_c}{\delta + l}\right)^2 \cdots \left(\frac{\beta_c}{\delta + l}\right)^n\right]^T \quad (35)$$

and,

◊ Motivated by TM models, with a state-variable filter (SVF)

$$\mathcal{B}_{SVF}(\delta) = \left[\frac{1}{E(\delta)} \frac{\delta}{E(\delta)} \cdots \frac{\delta^{n-1}}{E(\delta)}\right]^T \quad (36)$$

where $1/E(\delta)$ is a nth order stable filter. A typical choice is a nth order Poisson filter.

The following issues are now studied via numerical examples:

- **Predictive ability**: The above two choices $\mathcal{B}_{PF}(\delta)$ and $\mathcal{B}_{SVF}(\delta)$ are related through a linear nonsingular transformation (for $\lambda \neq 0$), e.g., for n=4 and β=1.

$$\mathcal{B}_{SVF}(\delta) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -\lambda \\ 0 & 2 & -2\lambda & \lambda^2 \\ 1 & -3\lambda & 3\lambda^2 & \lambda^3 \end{bmatrix} \mathcal{B}_{PF}(\delta)$$

Hence for a given model order, models based on these two sets will have the same predictive ability.

- **The numerical behavior** of the estimation algorithm is dictated by the condition number of the matrix

$$\boldsymbol{R} = \sum_{k=1}^{N} \boldsymbol{\phi}(k)\boldsymbol{\phi}^T(k).$$

It has been pointed out by (Subrahmanyam and Rao 1993) that use of $\mathcal{B}_{PF}(\delta)$ results in high condition numbers of the above matrix, as these functions are overlapping and non-orthogonal. On the other hand, the second set $\mathcal{B}_{SVF}(\delta)$ is near-orthogonal (Goodwin et al 1991), which improves the condition number.

- **Numerical conditioning** may be improved if an intelligently chosen linear transformation of these sets of basis functions is made before parameter estimation commences. When such transformation results in an orthogonal set, the numerical properties of the algorithm will be significantly improved. A popular orthogonal basis is in terms of Laguerre filters:

$$\mathcal{B}_{LAG}(\delta) = \left[\frac{1}{\delta + \lambda} \frac{1}{\delta + \lambda}\left(\frac{\delta - \lambda}{\delta + \lambda}\right) \cdots \frac{1}{\delta + \lambda}\left(\frac{\delta - \lambda}{\delta + \lambda}\right)^{n-1}\right]^T.$$

The required linear transformations are for (n = 4 and β = 1),

$$\mathcal{B}_{LAG}(\delta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & -2\lambda & 0 & 0 \\ 1 & -4\lambda & 4\lambda^2 & 0 \\ 1 & -6\lambda & 12\lambda^2 & -8\lambda^3 \end{bmatrix} \mathcal{B}_{PF}(\delta)$$

and

$$\mathcal{B}_{LAG}(\delta) = \begin{bmatrix} 1 & 3\lambda^2 & 3\lambda & 1 \\ -\lambda^3 & -\lambda^2 & \lambda & 1 \\ \lambda^3 & -\lambda^2 & \lambda & 0 \\ 1 & 3\lambda^2 & 3\lambda & 1 \end{bmatrix} \mathcal{B}_{PF}(\delta).$$

Part C: Recent developments in nonlinear continuous-time system identification

1. STATE OF THE ART

The methods dealt with in Part B of this paper for the identification of linear continuous-time systems involve essentially the computation of suitable measures of the derivatives of the input-output data, the use of which converts the identification problem to one of solving a set of algebraic equations. These measures are usually integrated or filtered versions of the data, or the coefficients in an orthogonal series expansion. Unfortunately these methods cannot be applied for the identification of nonlinear continuous-time systems. The main difficulty is that, in general, the nonlinear differential equations (NDE) describing these systems, are not integrable, which makes the application of the integral or filtering methods impossible. Only if the NDE is exactly integrable, i.e. if its terms can be written as pure derivatives of some computable function of the measured signals, can these methods be applied. This class of systems will be subsequently referred to as *integrable nonlinear systems*.

The second difficulty is that of computational burden. Further computations are required, to obtain the approximation for the derivatives and the various nonlinear product terms. The only exception to the above is the Fourier series expansion, where the coefficients can be computed using fast algorithms for a discrete Fourier transform (DFT). Another advantage of this method is the simple relation that exists between the Fourier spectra of a signal and its derivatives. This approach was introduced by Pearson (Pearson and Lee 1985, Pearson 1988, 1992, Pearson and Pan 1991, Pearson et al. 1993) and will be discussed in more detail later.

Because of these reasons, use of the standard linear techniques has been mostly restricted to only very special classes of nonlinear systems, such as bilinear systems, Hammerstein systems etc. As already mentioned in (Patra and Unbehauen, 1995) piecewise constant orthogonal expansions were applied by several authors for the identification of the above categories of specific nonlinear systems. The delayed state variable filter approach proposed in (Tsang and Billings, 1994) looks promising for the identification of nonlinear continuous-time systems from sampled data records. Delayed filtered inputs and outputs and associated higher-order derivatives collected from the state variable filters are used for the identification of the unknown system parameters using an orthogonal least-squares estimate. Another recently proposed approach (Liu et al. 1996) is based on a sequential identification scheme for nonlinear continuous-time systems with unknown nonlinearities using a resource allocating neural network. However, besides good simulation results, no details on estimated parameters are provided.

Apart from the above approaches, there also exist techniques for identifying nonlinear continuous-time state-spaces models, which employ the quasi-linearisation or invariant embedding principles to convert the task of parameter estimation to one of solving a nonlinear two-point boundary value problem. However, these methods are also computationally quite demanding.

Because of these difficulties mentioned above most of the efforts for identification of nonlinear systems in recent years have been towards the use of discrete-time models, such as Kolmogorov-Gabor polynomials (Kortmann and Unbehauen 1987, Kortmann et al. 1988, Flunkert 1992, Pottmann et al. 1993), artificial neural networks (Narendra and Parthosarathy 1990, Chen et al. 1990, Billings et al. 1993, Chen et al. 1992, Junge and Unbehauen 1996), fuzzy models (Takagi and Sugeno 1985, Pedrycz 1984, Sugeno and Tanaka 1991, Kortmann and Unbehauen 1996).

Under the consideration of the actual state of art in the field of nonlinear continuous-time system identification, it seems that modulating functions methods are the most promising ones. These will be dealt with in the next section.

2. MODULATING FUNCTIONS METHODS

In 1957 the concept of modulating functions was introduced by Shinbrot which is commonly known as Shinbrot's method of moment functionals. Basically, the use of modulating functions is to convert a differential expression involving input-output signals on a specified time interval into a sequence of algebraic equations. Shinbrot (1957) has presented a general theory of the so-called equations-of-motion methods for the analysis of dynamical systems and chosen a modulating functions of the form

$$\phi_m(t) = \sin^2(\omega_m t),\ 0 \le t \le T, \tag{37}$$

where $\omega_m = m\pi / T$ and $m = 1,2,\ldots,M$. The function $\phi_m(t)$ was quite arbitrary but must satisfy the two-point boundary conditions given by $\phi_m(t) = \dot{\phi}_m(t) = 0$ at $t = 0$ and $t = T$. This approach was also applied to the field of system identification depending on several choices of modulating functions: such as trigonometric- (Pearson and Lee, 1983), Fourier- (Pearson and Lee, 1985; Pearson, 1988, 1992; Pearson and Pan, 1991; Pearson, et al., 1993), Hermite- (Jalili et al., 1992), spline-type- (Maletinsky, 1979), Hartley-modulating functions (Patra and Unbehauen, 1995; Unbehauen, 1995; Daniel-Berhe and Unbehauen, 1996), etc.

Pearson et al. (Pearson and Lee, 1985; Pearson, 1988, 1992; Pearson and Pan, 1991; Pearson, et al., 1993) have used Fourier modulating functions and developed a class of nonlinear input-output differential operator models which can be implemented for parameter identification. Based on Shinbrot's method of moment functionals, Pearson used the following Fourier modulating functions (FMF) for the formulation of a least-squares (LS) parameter identification technique, i.e. an nth order class of FMF on $[0,T]$ is defined by

$$\begin{aligned}\phi_m(t) &= \frac{1}{T} e^{-jm\omega_0 t}\left(e^{-j\omega_0 t} - 1\right)^n \\ &= \frac{1}{T}\sum_{i=0}^{n}(-1)^{n-1}\binom{n}{i} e^{-j(m+i)\omega_0 t},\end{aligned} \tag{38}$$

where $m = 0,\pm 1,\pm 2,\ldots$ referred to as the modulating frequency index, $j = \sqrt{-1}, \omega_0 = 2\pi / T$ plays the role of a resolving frequency, and $\phi_m(t)$ is a sequence of order n modulating functions if it is sufficiently smooth and the $2n$ end point conditions $\phi_m^{(\nu)}(t) = 0$ at $t = 0$ and $t = T, \nu = 0,1,\ldots,n-1$ are satisfied for each m. The general form of a nonlinear input-output differential operator deterministic model essentially developed by Pearson (1992) is given by

$$\sum_{i=0}^{n_1}\sum_{k=1}^{n_2} g_i(\boldsymbol{\theta})F_{ik}(\boldsymbol{u},y)P_{ik}(p)E_k(\boldsymbol{u},y)=0, \quad (39)$$

where $g_i(\boldsymbol{\theta})$ are given functions of the parameter vector $\theta, F_{ik}(\boldsymbol{u},\boldsymbol{y})$ and $E_k(\boldsymbol{u},\boldsymbol{y})$ are specified functions of the I/O-pair $(\boldsymbol{u},y), P_{ik}(p)$ are fixed polynomials of degree n in the differential operator $p=d/dt$ and $g_0=1$. Thus, the initial step in this method is to rearrange a model in the form of (39) which can describe a large number of physical systems. Then it is possible to specify a cost function $J(\boldsymbol{\theta})\geq 0$ using the FMF-method for a given I/O-data set over the observation time interval, and minimizing $J(\boldsymbol{\theta})$ will lead to a one-shot least squares estimate of the system parameters. The Fourier modulating functions method have shown promising results for the parameter identification of linear and nonlinear input-output differential operator models.

The spline-type modulating functions method (Maletinsky, 1979) is based on an equation error, and is linear in the parameters. By this method the parameters of a differential equation can be estimated via least-squares or instrumental variable techniques. The approach provides a special model of the process, which is linear in the parameters and based on weighted integration of the measured input-output signals of the process and further more suited for digital implementation. The method has been applied to a linear system and shows interesting results. However, the method requires further studies for nonlinear continuous-time systems identification.

The other encouraging physically-based nonlinear continuous-time systems identification approach is by means of Hartley modulating functions (HMF) (Patra and Unbehauen, 1995; Unbehauen, 1995; Daniel-Berhe and Unbehauen, 1996). The HMF-method has been introduced by Patra and Unbehauen (1995) as a new member of the modulating functions family given by

$$\phi_m(t)=\sum_{i=0}^{n}(-1)^i\binom{n}{i}cas(n+m-i)\omega_0 t, 0<t\leq T, \quad (40)$$

where $\operatorname{cas}\omega t=\cos\omega t+\sin\omega t, m=0,\pm1,\pm2,\ldots$ is referred to as the modulating frequency index, $\omega_0=2\pi/T$ plays the role of a resolving frequency, and $\phi_m(t)$ is a member of the family of an nth order HMF if it is sufficiently smooth and the two-point boundary conditions

$$\phi_m^{(\nu)}(t)=0 \text{ for } t=0 \text{ and } t=T, \nu=0,1,\ldots,n-1 \quad (41)$$

are satisfied for each m, where $\phi_m^{(\nu)}(t)$ is the νth derivative of a member of a family of HMF $\{\phi_m(t)\}$. This modulating function is closely related to the Fourier modulating functions. Compared to the very efficient FMF-method the HMF-method has, however, as an important additional advantage that the HMFs are real-valued, and the Hartley-spectra can be computed efficiently with the help of fast algorithms for discrete Hartley transformation (DHT).

This new methodology is applicable to a large class of nonlinear continuous-time systems by defining a set of HMF for characterizing the continuous process signals. The approach has adopted a class of non-linear differential operator models essentially developed by Pearson et al., and has been applied to parameter identification by formulating an efficient least-squares algorithm. The method has been implemented to Hammerstein, integrable and convolvable non-linear models and shows promising results. Furthermore, attempts have been made to apply the method for identification of physically-based continuous-time dynamics and encouraging results have been obtained.

3. ILLUSTRATIVE SIMULATION STUDIES OF HMF-METHOD

In this section some basics of the recently developed parameter identification technique of HMF-method using three typical demonstrative studies of Hammerstein, integrable and convolvable nonlinear continuous-time systems will be discussed on the basis of simulated examples.

Basically, the method uses modulating functions of the form given in (40), Hartley transform and its spectrum. The Hartley transform (HT), first introduced by Hartley (1942), is an alternative formulation of a harmonic functional transform similar to the Fourier identity. It can be obtained from the Fourier transform by replacing the exponential function

$e^{-j\omega t}=\cos\omega t-j\sin\omega t$ by $\cos\omega t+\sin\omega t=\operatorname{cas}\omega t$.

The Hartley transform of a signal $x(t)$ for the continuous- and discrete-time case are given by

$$\text{Continuous case}\begin{cases} H_x(\omega)=\mathcal{H}_c\{x(t)\} \\ \qquad =\int_{-\infty}^{\infty}x(t)\operatorname{cas}\omega t\,dt \\ x(t)=\mathcal{H}_c^{-1}\{H_x(\omega)\} \\ \qquad =\dfrac{1}{2\pi}\int_{-\infty}^{\infty}H_x(\omega)\operatorname{cas}\omega t\,d\omega \end{cases} \quad (42)$$

and

Discrete case
$$\begin{cases} H_x(l) = \mathcal{H}_d\{x(\frac{kT}{N})\}, l = 0,1,\ldots,N-1 \\ \quad = \frac{1}{N}\sum_{k=0}^{N-1} x(\frac{kT}{N})\mathrm{cas}(\frac{2\pi lk}{N}) \\ x(\frac{kT}{N}) = \mathcal{H}_d^{-1}\{H_x(l)\} \\ \quad = \sum_{l=0}^{n-1} H_x(l)\mathrm{cas}(\frac{2\pi lk}{N}) \end{cases} \quad (43)$$

respectively. Let $\overline{H}_x(m\omega_0)$ be the mth HMF spectral component of a continuous signal $x(t)$ given by

$$\begin{aligned}\overline{H}_x(m\omega_0) &= \int_0^T x(t)\phi_m(t)dt \\ &= \sum_{i=0}^{n}(-1)^i\binom{n}{i}\int_0^T x(t)\mathrm{cas}(n+m-i)\omega_0 t dt \\ &= \sum_{i=0}^{n}(-1)^i\binom{n}{i}H_x((n+m-i)\omega_0). \quad (44)\end{aligned}$$

If $x^{(\nu)}(t), \nu = 1,2,\ldots,n$ is the νth derivative of the signal $x(t)$, then its corresponding *HMF spectrum* is given by

$$\begin{aligned}\overline{H}_x^{(\nu)}(m\omega_0) = \sum_{i=0}^{n}(-1)^i\binom{n}{i}\mathrm{cas}'(\nu\pi/2)(n+m-i)^\nu \\ \cdot\omega_0^\nu H_x((-1)^\nu(n+m-i)\omega_0), \quad (45)\end{aligned}$$

where $\mathrm{cas}'(\cdot) = \cos(\cdot) - \sin(\cdot)$. In addition, if the signals $x_1(t)$ and $x_2(t)$ are in the form of $x_1(t)x_2^{(\nu)}(t)$, then the HMF spectrum for such a product is given by

$$\begin{aligned}\overline{H}_{x_1,x_2}^{0,\nu}(m\omega_0) &= E_{x_1}(m\omega_0)\cdot\overline{H}_{x_2}^{(\nu)}(m\omega_0) \\ &\quad + O_{x_1}(m\omega_0)\cdot\overline{H}_{x_2}^{(\nu)}(-m\omega_0) \quad (46) \\ &= H_{x_1}(m\omega_0)\otimes\overline{H}_{x_2}^{(\nu)}(m\omega_0),\end{aligned}$$

where

$E_{x_1}(m\omega_0) = \frac{1}{2}\left[H_{x_1}(m\omega_0) + H_{x_1}(-m\omega_0)\right]$ and

$O_{x_1}(m\omega_0) = \frac{1}{2}\left[H_{x_1}(m\omega_0) - H_{x_1}(-m\omega_0)\right]$

are the even and odd parts of $H_{x_1}(m\omega_0)$ respectively, and the operator $\otimes$ symbol represents the two convolutions in short form for simplicity. For demonstration of the method, let us first consider a Hammerstein continuous-time nonlinear dynamic system.

3.1 HAMMERSTEIN NONLINEAR CONTINUOUS-TIME SYSTEM

Let us first consider a Hammerstein continuous-time nonlinear dynamic system having single input $u(t)$ and output $y(t)$. The system could be modelled via a state space description. Its nonlinear differential equation model is given by

$$\ddot{y}(t) + a_1\dot{y}(t) + a_2 y(t) = b_0 u(t) + b_1 u_3(t), \quad (47)$$

where $[a_1\ a_2\ b_0\ b_1] = [3\ 2\ 2\ 1]$. Multiplication of (47) by $\phi_m(t)$ of $n=2$ and integrating over $[0, T]$ implies

$$\begin{aligned}\int_0^T \ddot{y}(t)\phi_m(t)dt + a_1\int_0^T \dot{y}(t)\phi_m(t)dt \\ + a_2\int_0^T y(t)\phi_m(t)dt = b_0\int_0^T u(t)\phi_m(t)dt \\ + b_1\int_0^T u^3(t)\phi_m(t)dt. \quad (48a)\end{aligned}$$

Substituting the definition of $\phi_m(t)$ according to (40) into (48a), integrating until all the derivatives shift to $\phi_m(t)$, simplifying and applying the Hartley spectrum notations discussed above leads to

$$\begin{aligned}\overline{H}_y^{(2)}(m\omega_0) + a_1\overline{H}_y^{(1)}(m\omega_0) + a_2\overline{H}_y(m\omega_0) \\ = b_0\overline{H}_u(m\omega_0) + b_1\overline{H}_{u^3}(m\omega_0). \quad (48b)\end{aligned}$$

Let $z(m\omega_0) = \overline{H}_y^{(2)}(m\omega_0)$, then (48b) can be rewritten as a linear regression,

$$z(m\omega_0) = \varphi^T(m\omega_0)\,\theta(m\omega_0) + \varepsilon(m\omega_0) \quad (49)$$

where $\varphi^T(m\omega_0) = \left[-\overline{H}_y^{(1)}(m\omega_0)\ \ -\overline{H}_y(m\omega_0)\ \ \overline{H}_u(m\omega_0)\ \ \overline{H}_{u^3}(m\omega_0)\right]$

$\theta(m\omega_0) = \left[a_1\ a_2\ b_0\ b_1\right]^T$.

Let a sequence of observation be made for $m = 0, \pm1, \pm2, \ldots, \pm M$. Then, $(2M+1)$ regression equations of the form (49) can be represented as a vector equation

$$z(m\omega_0) = \Psi(m\omega_0)\ \theta(m\omega_0) + \varepsilon(m\omega_0), \quad (50)$$

where $\Psi^T(m\omega_0) = \left[\varphi(-M\omega_0)\ldots\varphi(-\omega_0)\ \varphi(0)\ \varphi(\omega_0)\ldots\varphi(M\omega_0)\right]$

and $z^T(m\omega_0) = \left[z(-M\omega_0)\ldots z(-\omega_0)\ z(0)\ z(\omega_0)\ldots z(M\omega_0)\right]^T$,

and the similarly $\varepsilon(m\omega_0)$. The objective function is to minimize a cost function of the form given by

$$\begin{aligned}J(\theta) &= \frac{1}{2}\sum_{m=-M}^{M}\varepsilon^2(m\omega_0,\theta) \\ &= \frac{1}{2}\sum_{m=-M}^{M}\left[z(m\omega_0) - \varphi^T(m\omega_0)\theta(m\omega_0)\right]^2. \quad (51)\end{aligned}$$

Minimizing $J(\boldsymbol{\theta})$ with respect to the unknown parameter vector $\boldsymbol{\theta}(m\omega_0)$, we obtain the least squares estimate of $\boldsymbol{\theta}(m\omega_0)$ as,

$$\hat{\boldsymbol{\theta}}(m\omega_0) = \left[\boldsymbol{\Psi}^T(m\omega_0)\boldsymbol{\Psi}(m\omega_0)\right]^{-1} \cdot \boldsymbol{\Psi}^T(m\omega_0)\,\boldsymbol{z}(m\omega_0). \quad (52a)$$

Here, several modifications of the least squares (LS) scheme are possible, such as the introduction of a positive definite symmetric weighting matrix $\boldsymbol{W}$ in the definition of $J(\boldsymbol{\theta}) = \frac{1}{2}\boldsymbol{\varepsilon}^T(m\omega_0,\boldsymbol{\theta})\boldsymbol{W}\boldsymbol{\varepsilon}(m\omega_0,\boldsymbol{\theta})$, which implies weighted LS estimate of $\boldsymbol{\theta}(m\omega_0)$ given by

$$\hat{\boldsymbol{\theta}}_w(m\omega_0) = \left[\boldsymbol{\Psi}^T(m\omega_0)\boldsymbol{W}\boldsymbol{\Psi}(m\omega_0)\right]^{-1} \cdot \boldsymbol{\Psi}^T(m\omega_0)\boldsymbol{W}\,\boldsymbol{z}(m\omega_0). \quad (52b)$$

In a similar way, a parameter estimation procedure for integrable and convolvable nonlinear systems can also be formulated, for instance, for *integrable* continuous-time nonlinear models of the type

$$\sum_{\nu=0}^{n_1}\sum_{i=0}^{n_2} \beta_{\nu i}\, f_i^{(\nu)}(u(t),y(t)) = 0, \qquad (53)$$

where $f_i^{(\nu)}(u(t),y(t))$ is the νth derivative of a known differentiable function $f_i(u(t),y(t))$. Multiplication of (53) by $\phi_m(t)$ and integrating over $[0, T]$ results in

$$\sum_{\nu=0}^{n_1}\sum_{i=0}^{n_2} \beta_{\nu i}\, \overline{H}_{f,i}^{(\nu)}(m\omega_0) = 0. \qquad (54)$$

In addition, for the more general *convolvable* category of continuous-time nonlinear models of the form

$$\sum_{\nu=0}^{n_1}\sum_{i=0}^{n_2}\sum_{k=0}^{n_3} \beta_{\nu ik}\, g_k(u(t),y(t))\, f_i^{(\nu)}(u(t),y(t)) = 0, \quad (55)$$

where $g_k(u(t),y(t))$ is another specified function of the input/output pair $(u(t),y(t))$, modulating (55) with $\phi_m(t)$ and applying (46) leads to the following form

$$\sum_{\nu=0}^{n_1}\sum_{i=0}^{n_2}\sum_{k=0}^{n_3} \beta_{\nu ik}\left[H_{g,k}(m\omega_0)\otimes \overline{H}_{f,i}^{(\nu)}(m\omega_0)\right] = 0. \quad (56)$$

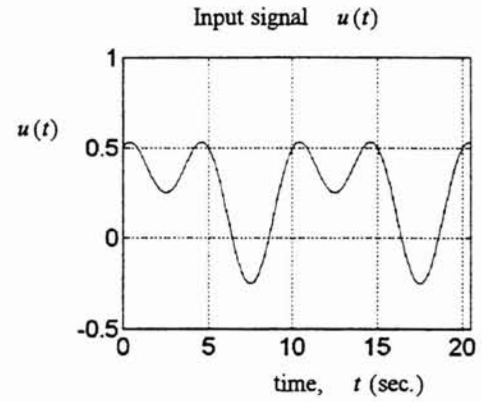
With a suitable normalization, (54) and (56) can be written in the regression form of (49) and the parameters can be estimated by the standard least squares technique as shown above. Some of the main advantages of this new class of modulating functions are that a set of algebraic equations with real coefficients results, the formulations are free from boundary conditions, and the computations can be made using fast algorithms for discrete Hartley transformation.

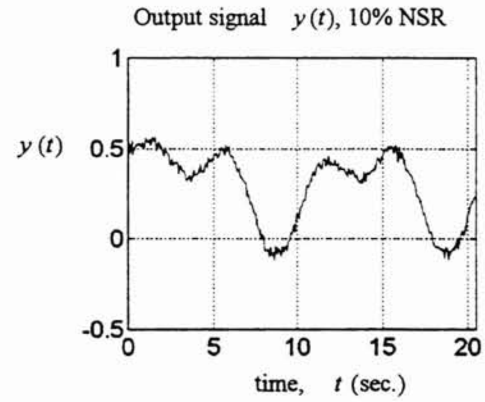
In general, if a system is arranged into the nonlinear input-output differential operator model of the form (39) while the I/O-data are available over an arbitrary period of time and with specified functions of $F_{ik}(\boldsymbol{u},\boldsymbol{y}), E_k(\boldsymbol{u},\boldsymbol{y})$ and fixed polynomials $P_{ik}(p)$ as well as given functions $g_i(\boldsymbol{\theta})$ of the unknown parameter vector $\boldsymbol{\theta}$, then without loss of generality we can rearrange the model in the regression form of (49) and (50) for a class of nonlinear continuous-time systems with the help of (54) and (56) both for the integrable and convolvable category of nonlinear model formulations, respectively.

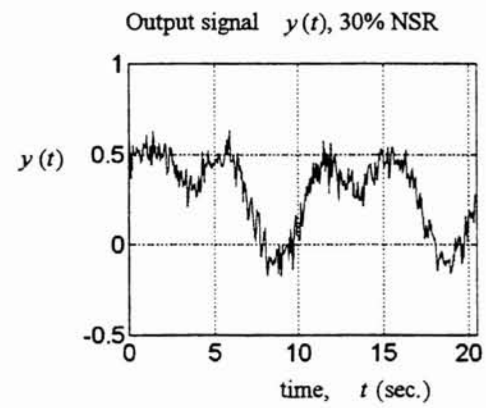
Now let us observe the basic computational requirements for all simulation studies:

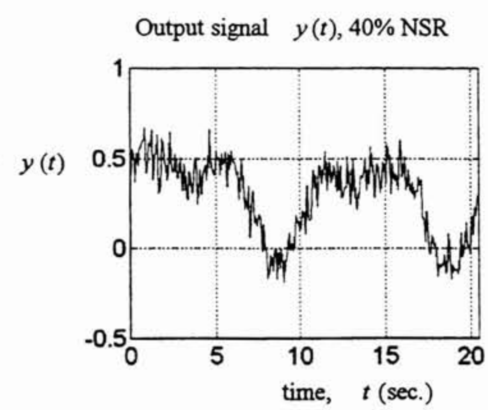
1) Solve the system differential equation using Runge-Kutta formulas over the given interval in order to obtain the input-output data of the system to be identified.
2) Mix noise with data for Monte Carlo simulation, i.e., corrupt the output signal $y(t)$ by additive white Gaussian noise with a chosen value of Noise-to-Signal ratio (NSR).
3) Compute the Hartley transform of the I/O-data using Simpson's rule for better approximation.
4) Compute the Hartley spectra and some of the derivatives, i.e., $\overline{H}_y, \overline{H}_y^{(1)}, \overline{H}_y^{(2)}, \overline{H}_u, \overline{H}_{u^3}$ etc.
5) Perform the operation $\otimes$ of (46) if any (this applies for a convolvable continuous-time nonlinear system simulation study).
6) Estimate the parameters using weighted least-squares.
7) Do steps (2) up to (6) for a required number of Monte Carlo runs. Calculate the average value of the parameters and the variance/standard deviation.
8) Repeat step (2) up to (7) for each of Noise-to-Signal ratio (NSR).

For the above Hammerstein nonlinear differential equation simulation study, the input signal was $u(t) = 0.25 + 0.25\sin(0.2\pi t + 0.25\cos(0.4\,\pi t)$ over a record length of T=20.48 sec. Furthermore, the record length was discretized into N=256 sub-intervals, but only $(2M+1) = 9$ significant spectral data points have been used in estimation. The response plots of I/O-signals and data are shown in Figures 10, 11 and Table 3, respectively. Twenty five Monte Carlo runs were made for each of several NSR and the weighted least squares estimate was used with a weighting factor inversely proportional to frequency. The simulation plots and data show that the Hartley modulating function method was able to estimate the parameters of Hammerstein nonlinear continuous-time systems in the presence of noticeable additive measurement noise precisely.









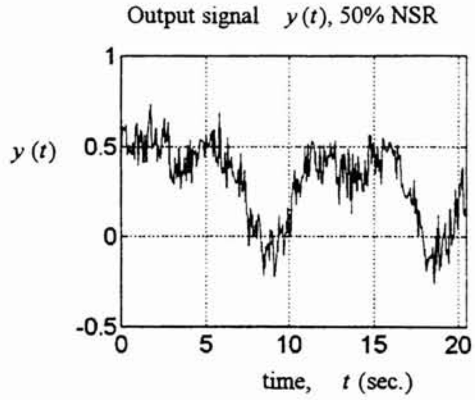


Fig. 10. Input-output response for the Hammerstein nonlinear continuous-time system example.

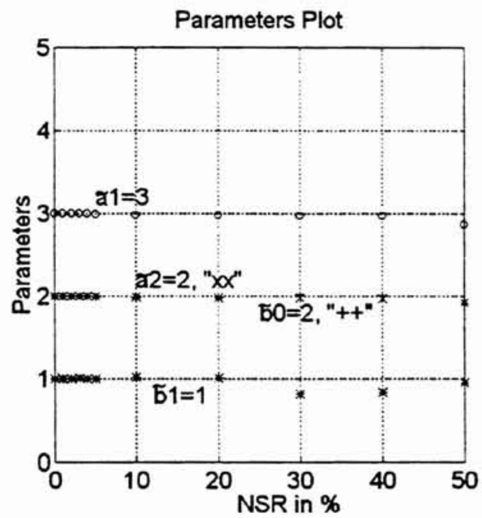


Fig. 11. Parameters estimation of the Hammerstein non-linear continuous-time system example.

Table 3 Estimated parameters, ①, and their standard deviation, ②, for the Hammerstein non-linear continuous-time system example

NSR	① - Parameters mean, $\tilde{a}_i, \tilde{b}_j$			
in %	$\tilde{a}_1\{\cong 3\}$	$\tilde{a}_2\{\cong 2\}$	$\tilde{b}_0\{\cong 2\}$	$\tilde{b}_1\{\cong 1\}$
0%	3.0000	2.0000	2.0000	1.0000
1%	3.0019	2.0012	2.0015	1.0002
2%	3.0036	2.0024	2.0017	1.0040
3%	3.0014	2.0003	1.9990	1.0065
4%	2.9999	2.0002	1.9998	1.0031
5%	2.9911	1.9959	1.9951	0.9995
10%	2.9781	1.9882	1.9829	1.0143
20%	2.9785	1.9836	1.9801	1.0072
30%	2.9754	1.9831	2.0186	0.8210
40%	2.9700	1.9760	2.0054	0.8398
50%	2.8655	1.9212	1.9151	0.9524

(Table 3 continued)

NSR in %	② - Parameters standard deviation $\sigma_{\tilde{a}_1}$	$\sigma_{\tilde{a}_2}$	$\sigma_{\tilde{b}_0}$	$\sigma_{\tilde{b}_1}$
0%	0.0000	0.0000	0.0000	0.0000
1%	0.0075	0.0043	0.0067	0.0123
2%	0.0127	0.0071	0.0098	0.0291
3%	0.0188	0.0112	0.0152	0.0477
4%	0.0193	0.0112	0.0166	0.0643
5%	0.0279	0.0162	0.0240	0.0744
10%	0.0577	0.0334	0.0511	0.1512
20%	0.0954	0.0533	0.0810	0.2443
30%	0.2025	0.1082	0.1699	0.4542
40%	0.2742	0.1566	0.2776	0.6784
50%	0.2203	0.1224	0.2355	0.8186

3.2. INTEGRABLE NONLINEAR CONTINUOUS-TIME SYSTEM

Now consider a simulation of an integrable nonlinear differential system given by

$$\ddot{y}(t) + a_1\, \dot{y}(t) + a_2\, y_2(t)\dot{y}(t) + a_3\, y(t) = a_4\, u(t), \quad (59)$$

where $[a_1\, a_2\, a_3\, a_4] = [-2\ \ 2\ \ 1\ \ 2]$. Modulating (59) with a HMF of $n = 2$ and using $y^2\dot{y} = d(y^3)/3dt$ implies

$$\overline{H}_y^{(2)} + a_1\, \overline{H}_{y1}^{(\,)} + \tfrac{a_2}{3}\, \overline{H}_{y^3}^{(1)} + a_3\, \overline{H}_y = a_4\, \overline{H}_u. \quad (60)$$

Implementing the aforementioned computational procedure for the simulation study leads to the response plots of the I/O-signals and data shown in Figures 12, 13 and Table 4, respectively. The input was $u(t) = 1/4\ \{1 + \sin(0.628t) + \cos(1.257t)\}$ over a record length of T=16 sec. Here, one hundred Monte Carlo runs were made for each of several NSR even if with a smaller number of runs also satisfactory results are obtained. The weighted LS estimation was used with a weighting factor inversely proportional to the frequency. Moreover, the record length was discretized into $2^{10} = 1024$ subintervals. These give rise to the same number of HMF spectra coefficients, but only a small number of $(2M + 1) = 19$ as a significant part of the spectrum was used for parameter estimation. The simulation plots and data show that the Hartley modulating function method was able to estimate the parameters of continuous-time systems in the presence of noticeable additive measurement noise again very precisely.

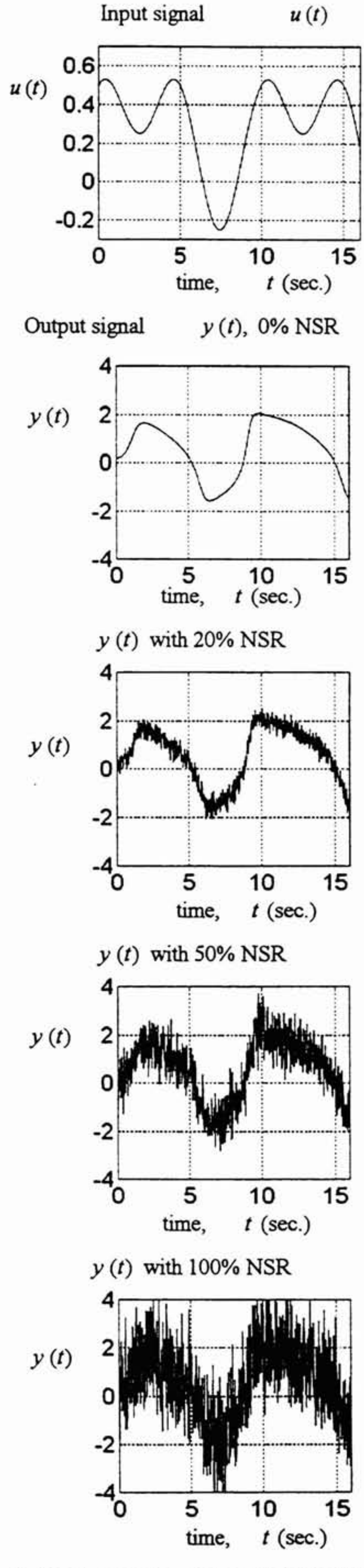


Fig. 12. Input-output response plots for the integrable nonlinear differential system example.

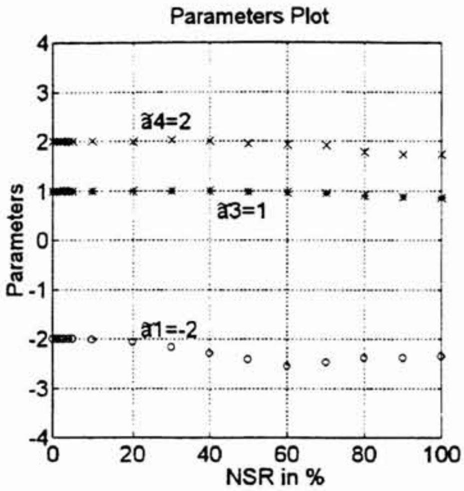


Fig. 13. Parameters estimation of the integrable non-linear differential system example.

Table 4 Estimated parameters,①, and their standard deviation,②, for the integrable nonlinear differential system example

NSR in %	① - Parameters mean, $\tilde{a}_i$ $\tilde{a}_1\{\cong\text{-}2\}$	$\tilde{a}_2\{\cong 2\}$	$\tilde{a}_3\{\cong 1\}$	$\tilde{a}_4\{\cong 2\}$
0%	-2.0000	2.0000	1.0000	2.0000
1%	-1.9997	1.9993	0.9999	1.9997
2%	-1.9997	1.9982	0.9999	1.9988
3%	-2.0038	2.0016	1.0004	2.0014
4%	-2.0042	2.0013	0.9982	1.9974
5%	-2.0027	1.9946	0.9995	1.9964
10%	-2.0182	1.9941	0.9973	1.9981
20%	-2.0590	1.9548	0.9923	1.9894
30%	-2.1737	1.9404	0.9981	2.0079
40%	-2.2933	1.8940	0.9911	2.0012
50%	-2.4062	1.8132	0.9750	1.9408
60%	-2.5450	1.7245	0.9578	1.9169
70%	-2.4695	1.4839	0.9390	1.9076
80%	-2.3861	1.2758	0.8963	1.7644
90%	-2.3799	1.1186	0.8686	1.7297
100%	-2.3470	0.9880	0.8591	1.7195

NSR in %	② - Parameters standard deviation, $\sigma_{\tilde{a}_i}$ $\sigma_{\tilde{a}_1}$	$\sigma_{\tilde{a}_2}$	$\sigma_{\tilde{a}_3}$	$\sigma_{\tilde{a}_4}$
0%	0.0000	0.0000	0.0000	0.0000
1%	0.0000	0.0000	0.0000	0.0000
2%	0.0100	0.0141	0.0000	0.0100
3%	0.0173	0.0200	0.0000	0.0141
4%	0.0245	0.0300	0.0000	0.0173
5%	0.0300	0.0374	0.0100	0.0245
10%	0.0574	0.0678	0.0173	0.0458
20%	0.1005	0.1200	0.0332	0.0812
30%	0.1934	0.2076	0.0520	0.1261
40%	0.2955	0.2890	0.0693	0.1789
50%	0.3908	0.3461	0.0883	0.2381
60%	0.5143	0.4167	0.1086	0.2557
70%	0.5884	0.4045	0.1072	0.2704
80%	0.6851	0.4225	0.1166	0.2777
90%	0.7610	0.3974	0.1140	0.3305
100%	0.9502	0.4580	0.1319	0.3558

3.3. CONVOLVABLE NONLINEAR CONTINUOUS-TIME SYSTEM

The third simulation example is a convolvable nonlinear differential system given by

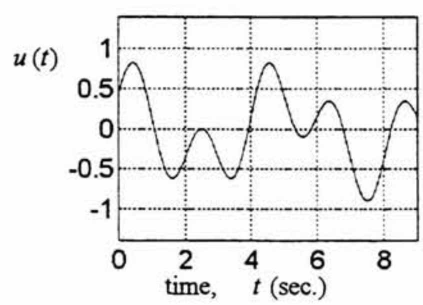
$$\ddot{y}(t) + a_1\, y(t) + a_2\, \dot{y}(t) \\ - a_3 y(t)u(t) - a_4 \dot{y}(t)u(t) = a_5 u(t), \quad (61)$$

where $[a_1\, a_2\, a_3\, a_4\, a_5] = [2\ \ 1\ \ 1\ \ 2\ \ 1]$. Similarly, modulating (61) with a HMF of n=2 implies

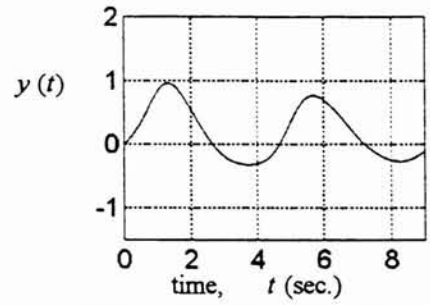
$$\overline{H}_y^{(2)} + a_1 \overline{H}_y + a_2 \overline{H}_{y1}^{(\,)} \\ - a_3 H_y \otimes \overline{H}_u - a_4 H_{y1}^{(\,)} \otimes \overline{H}_u = a_5 \overline{H}_u. \quad (62)$$

Here, the input signal was $u(t) = 0.45\,\{\cos(1.257t) + \sin(3.142t)\}$ over a record length of T=9 sec. After implementing the aforementioned computational procedures for the simulation study, the response plots of the I/O-signals and data shown in Figures 14, 15 and Table 5 are obtained, respectively. Note that one hundred Monte Carlo runs were made for each of several NSR even if with a smaller number of runs also satisfactory results are obtained. The weighted LS estimation was used with a weighting factor inversely proportional to the frequency. Moreover, the record length was discretized into $2^{10} = 1024$ sub-intervals. These give rise to the same number of HMF spectra coefficients, but only a small number of $(2M+1) = 33$ as spectral data points have been used, and the results show the high performance of the proposed algorithm to estimate the parameters of nonlinear continuous-time systems in the presence of noticeable additive measurement noise.

Input signal $u(t)$



Output signal $y(t)$, 0% NSR



(Figure 14 to be continued)

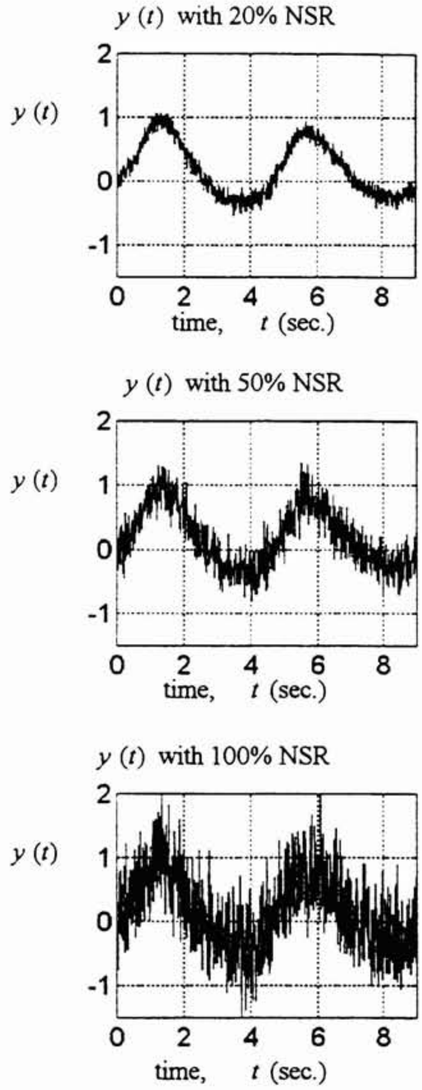


Fig. 14. Input-output response plots for the convolvable nonlinear differential system example.

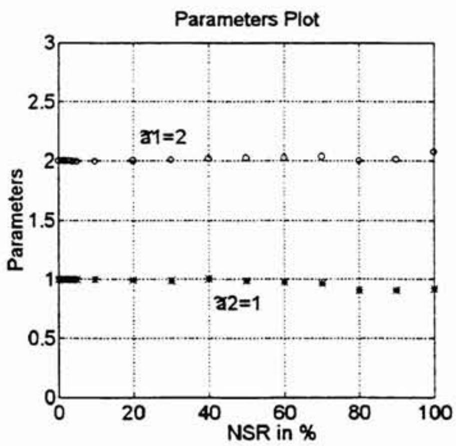


Fig. 15. Parameters estimate of the convolvable non-linear differential system example.

Table 5 Estimated parameters, ①, and their standard deviation, ②, for the convolvable nonlinear differential system example.

NSR in %	① - Parameters mean, $\tilde{a}_i$: $\tilde{a}_1\{\cong 2\}$	$\tilde{a}_2\{\cong 1\}$	$\tilde{a}_3\{\cong 1\}$	$\tilde{a}_4\{\cong 2\}$	$\tilde{a}_5\{\cong 1\}$
0%	2.0000	1.0000	1.0000	2.0000	1.0000
1%	1.9999	1.0001	0.9989	1.9995	1.0001
2%	2.0001	0.9997	1.0005	2.0003	1.0000
3%	1.9999	0.9996	0.9979	2.0005	0.9998
4%	1.9981	0.9986	0.9977	1.9960	0.9985
5%	1.9989	0.9956	1.0024	1.9967	0.9946
10%	1.9979	0.9990	0.9984	1.9974	0.9977
20%	2.0016	0.9942	1.0148	1.9963	0.9913
30%	2.0082	0.9842	0.9856	1.9964	0.9803
40%	2.0176	1.0047	1.0467	2.0204	1.0089
50%	2.0248	0.9845	1.0128	1.9943	1.0019
60%	2.0247	0.9760	0.9873	1.9917	0.9950
70%	2.0339	0.9619	1.0888	2.0277	0.9487
80%	2.0029	0.9068	0.9632	1.9205	0.9044
90%	2.0137	0.9035	1.0687	1.9236	0.9172
100%	2.0745	0.9180	1.1739	2.0080	0.9146

NSR in %	② - Parameters standard deviation, $\sigma_{\tilde{a}_i}$: $\sigma_{\tilde{a}_1}$	$\sigma_{\tilde{a}_2}$	$\sigma_{\tilde{a}_3}$	$\sigma_{\tilde{a}_4}$	$\sigma_{\tilde{a}_5}$
0%	0.0000	0.0000	0.0000	0.0000	0.0000
1%	0.0000	0.0000	0.0000	0.0000	0.0000
2%	0.0000	0.0000	0.0141	0.0100	0.0100
3%	0.0000	0.0100	0.0173	0.0141	0.0100
4%	0.0100	0.0100	0.0224	0.0141	0.0173
5%	0.0100	0.0141	0.0316	0.0173	0.0173
10%	0.0173	0.0265	0.0583	0.0316	0.0400
20%	0.0361	0.0548	0.1269	0.0707	0.0806
30%	0.0616	0.0781	0.1942	0.1095	0.1158
40%	0.0794	0.1020	0.2366	0.1393	0.1513
50%	0.0964	0.1449	0.3367	0.1811	0.2045
60%	0.1058	0.1513	0.3956	0.2205	0.2300
70%	0.1463	0.1855	0.4510	0.2820	0.2865
80%	0.1520	0.2283	0.4793	0.2532	0.3295
90%	0.1597	0.2598	0.5420	0.3161	0.3809
100%	0.2030	0.2846	0.6377	0.3610	0.4189

In general, the Hartley modulating functions method is a very effective technique for parameter estimation of nonlinear continuous-time systems which are linear in their parameters. Here, simulation examples of a Hammerstein, an integrable nonlinear and a convolvable nonlinear continuous-time systems were used to examine the performance of the proposed algorithm. Thus, from the simulation plots and data, one can observe that the algorithm was able to estimate the parameters of continuous-time systems in the presence of noticeable additive measurement noise very precisely.

CONCLUSIONS

A survey of the various approaches to the problem of identification of continuous-time systems is attempted in this paper. The focus is particularly on lumped linear and non-linear models and on those developments that followed an earlier survey by the present authors themselves (Unbehauen and Rao 1990). The various approaches have been outlined in a unified framework and the significance of continuous-time parametric models of physical systems has been discussed from the view point of control engineering. Recent developments on the application of models that result in linear estimation, i.e., models that are in a general moving average (MA) form, have been discussed.

In the case of identification of linear systems, Markov parameter and time moment models have been assessed. They are then generalized and related to the other form of transfer function expansion in terms of signal preprocessing features and filters. In the case of identification of nonlinear systems, some recent developments based on Hartley modulating functions are discussed and their promise and potential have been demonstrated with the help of simulated examples. Particularly in the identification of nonlinear systems, further research is needed to provide solutions to the many problems that still exist and to render the identification methods online.

REFERENCES

Aström, K.J., and P. Eykhoff (1971). System identification - A survey. *Automatica*, 7, 123-162.

Bapat, V.N. (1993). *Some extensions to Poisson moment functional based estimation of continuous-time models of dynamical systems.* Ph.D. Thesis, Department of Electrical Engineering, Indian Institute of Technology, Kharagpur, India.

Billings, S.A., H.B. Jamaluddin and S. Chen (1992). Properties of neural networks with applications to modelling nonlinear dynamical systems. *Intern. J. of Control*, **55**, 193-224.

Box, G.E.P. and G.W. Jenkins (1970). *Time series analysis: Forecasting and control* (2nd Edition), Holden Day, San Francisco, USA.

Chen, S., S.A. Billings and PM. Grant (1990). Nonlinear system identification using neural netwoks. *Intern. J. of Control*, **51**, 1191-1214.

Chen, S., S.A. Billings and P. M. Grant (1992). Recursive hybrid algorithm for nonlinear system identification using radial basis function networks *Intern. J. of Control*, **55**, 1051-1070.

Daniel-Berhe, S. and H. Unbehauen (1996). Parameter estimation of nonlinear continuous-time systems using Hartley modulating functions. In *Proc. IEE Int. Conference on Control*, UKACC, Exeter, 228-233.

Daniel-Berhe, S. and H. Unbehauen (1996). Application of the Hartley modulating functions method for the identification of the bilinear dynamics of a dc motor. In *Proc. of 35th IEEE Conference on Decision and Control*, Kobe, Japan, 1533-1538.

Datta, K.B. and B.M. Mohan (1995). *Orthogonal Functions in Systems and Control*. World Scientific, Singapore.

Dhawan, R.K., A. Sahai, D.V. Nishar and G.P. Rao (1991). Recursive estimation of Markov parameters in linear continuous-time SISO systems via block-pulse functions. In *Proc. of IFAC Symposium on Identification and System Parameter Estimation,* Budapest, Hungary, 1495-1500.

Diekmann, K. and H. Unbehauen (1979). Recursive identification of multiple-input multiple-output systems. In *Proc. of IFAC Symposium on Identification and System Parameter Estimation,* Darmstadt, Germany, 423-429.

Eykhoff, P. (1974). *System identification*. Wiley, New York, USA.

Fine, N.J. (1949). On the Walsh functions (1949). *Trans. Amer. Math. Soc.*, **65**, 372-414.

Flunkert, H.U. (1992). *Regelstrategien auf der Basis reduzierter nichtlinearer Modelle*. VDI-Verlag, Düsseldorf, Germany.

Gantmacher, F.R. (1959). *Applications of the theory of matrices.* Interscience Publishers, Inc., New York, USA.

Garnier, H., P. Sibille, H.L. Nguyen and T. Spott (1994). A bias-compensating least-squares method for continuous-time system identification via Poisson moment functionals. In *Proc. of IFAC Symposium on System Identification and Parameter Estimation, SYSID'94*, Copenhagen, Denmark, **3**, 675-680.

Garnier, H., P. Sibille and T. Spott (1994). Influence of the initial covariance matrix on recursive LS estimation of continuous models via generalized Poisson moment functionals. In *Proc. of IFAC Symposium on System Identification and Parameter Estimation, SYSID'94*, Copenhagen, Denmark, **3**, 669-674.

Gawthrop, P.J. (1987). *Continuous-time self-tuning control.* IEE Publication series, London, UK.

Goodwin G.C., B. Ninness and V. Poor (1991). Choice of basis functions for continuous and discrete system modeling. In *Proc. of IFAC Symposium on Identification and System Parameter Estimation*, Budapest, Hungary, 1179-1184.

Goodwin G.C., M. Gevers and D.Q. Mayne (1991) Bias and variance distribution in transfer function estimation. In *Proc. of IFAC Symposium on Identification and System Parameter Estimation*, Budapest, Hungary, 952-957.

Hartley, R.V.L. (1942). A more symmetrical Fourier analysis applied to transmission problems. *Proc. Inst. Radio Engrs.*, 144-150.

Jalili, S.A., J.A. Jordan and R.D.L. Mackie (1992). Measurement of the parameters of all-pole transfer functions using shifted Hermite modulating functions. *Automatica*, **38**, 613-617.

Junge, Th. F. and H. Unbehauen (1996). Off-line identification of nonlinear systems using structurally adaptive radial basis functions. In *Proc. 35th IEEE Conference on Decision and Control*, Kobe, Japan, 943-948.

Kalman R.E. (1970). *Aspects of network and systems theory*. Holt, Reinhart and Winston, New York, USA.

Karanam, V.R., P.A. Frick and R. R. Mohler (1978) Bilinear system identification by Walsh functions. *IEEE Trans. Autom. Cont.*, *AC*-**23**, 709-713.

Kortmann, M. and H. Unbehauen (1988) Two algorithms for model structure determination of nonlinear dynamic systems with application to industrial processes. In *Proc. of the 8th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, Beijing, People's Republic of China (Pergammon).

Kortmann, M., K. Janiszowski and H. Unbehauen (1988). Application and comparison of different identification schemes under industrial conditions. Intern. J. of Control, **48**, 2275-2296.

Kortmann, P. and H. Unbehauen (1996). Identification of the structure of Fuzzy Models. In *Proc. of Fuzzy 96*, Zittau, Germany, 36-46.

Kung, F.C. and D.H. Shih (1986). Analysis and identification of Hammerstein model nonlinear delay systems using block-pulse function expansion. *Intern. J. of Control*, **43**, 139-147.

Küper P. (1992). *Identifikation kontinuierlicher dynamischer Systeme mit Hilfe von Markov-Parameters*. Studienarbeit ESR-9138, Lehrstuhl für Elektrische Steuerung and Regelung, Ruhr Universität, Bochum, Germany.

Liu, G.P., V. Kadirkamanathan and S.A. Billings (1996), Stable sequential identification of continuous nonlinear dynamical systems by growing radial basis function networks. *Intern. J. of Control*, **65**, 53-69.

Ljung L. (1987). *System iIdentification: Theory for the user*. Prentice Hall, Inc., Englewood Cliffs, N.J., USA.

Maletinsky, V. (1979). Identification of continuous dynamical systems with spline-type modulating functions method. In *Proc. 5th IFAC Symp. on Identification and System Parameter Estimation*, Darmstadt, Germany, 275-281.

Middleton R.H. and G.C. Goodwin (1990). *Digital control and estimation: A unified approach*. Prentice Hall, Inc., Englewood Cliffs, N.J., USA.

Mukhopadhyay S. and G.P. Rao (1991). Integral-equation approach to joint state and parameter estimation in continuous-time MIMO systems., *IEE-Proc., Part D*, **138**, 93-102.

Mukhopadhyay S., A. Patra and G.P. Rao (1991). Irreducible model estimation for MIMO systems, *Intern. J. of Control*, **53**, 223-253.

Mukhopadhyay S., A. Patra and G.P. Rao (1992). New class of discrete-time models for continuous-time systems. *Intern. J. of Control*, **55**, 1161-1187.

Narendra, K.S. and K. Parthasarathy (1990). Identification and control of dynamical systems using neural networks. *IEEE Trans. on Neural Networks*, **1**, 4-27.

Norton J.P. (1986). *An introduction to identification*. Academic Press, New York, USA.

Patra, A. (1989). *General hybrid orthogonal functions and some applications in systems and control*. Ph.D. Thesis, Department of Electrical Engineering, Indian Institute of Technology, Kharagpur-721302, India.

Patra, A. and G.P. Rao (1993). *General hybrid orthogonal functions and their applications in systems and control*. Internal Report, Department of Electrical Engineering, Indian Institute of Technology, Kharagpur-721302, India.

Patra, A. and H. Unbehauen (1995). Identification of a class of nonlinear continuous-time systems using Hartley modulating functions. *Intern. J. of Control*, **62**, 1431-1451.

Patra A. and G.P. Rao (1996).: *General hybrid orthogonal functions and their application in* *systems and control.* LNCIS Series Vol. 213, Springer-Verlag, Berlin, Germany.

Pearson, A.E. and F.C. Lee (1983). Time limited identification of continuous systems using trigonometric modulating functions. In *Proc. of 3rd Yale Workshop on Applic. of Adap. Syst.*, New Haven, CT, USA, 168-173.

Pearson, A.E. and FC. Lee (1985). On the identific-ation of polynomial input/output differential systems. *IEEE Trans. Autom. Cont.*, *AC*-**30**, 778-782.

Pearson, A.E. and F.C. Lee (1985). Parameter ident-ification of linear differential systems via Four-ier based modulating functions. *Control-Theory and Advanced Technology*, **1**, 239-266.

Pearson, A.E. (1988). Least squares parameter ident-ification of nonlinear differential I/O models. In *Proc. of 27th IEEE Conference on Decision and Control'88*, Austin, USA, 1831-1835.

Pearson, A.E. and J.Q. Pan (1991). Frequency anal-ysis via the method of moment function. In *Proc. of 30th IEEE Conference on Decision and Control'91*, Brighton, UK, 2024-2025.

Pearson, A.E. (1992). Explicit parameter identific-ation for a class of nonlinear I/O differential operator models. In *Proc. of 31st IEEE Conference on Decision and Control' 92*, Tucson, Arizona, 3656-3660.

Pearson, A.E., Y. Shen and J.Q. Pan (1993). Discr-ete frequency formats for linear differential sys-tem identification. In *Proc. of 12th IFAC World Congress*, Sidney, Australia, **VII**, 143-148.

Pedrycz, W. (1984). Identification in fuzzy systems. In *Proc. IEEE Trans. on Systems, Man, and Cybernetics*, **14**, 361-366.

Pottmann, M., H. Unbehauen and D.E. Seborg (1993). Application of a general multi-model approach fir identificationof highly nonlinear processes - a case study. *Intern. J. of Control*, **57**, 97-120.

Rao, K.V., P.A. Frick and R.R. Mohler (1976). On bilinear system identification by Walsh func-tions. In *Proc. of 4th IFAC Symp. on Identi-fication and System. Parameter Estimation*, Tiblisi, USSR, **3**, 350-359.

Rao G.P. (1983). *Piecewise constant orthogonal functions and their applications to systems and control.* LNCIS-Series, Vol. 55, Springer-Verlag, Berlin, Germany.

Rao G.P., K. Diekmann and H. Unbehauen (1984). Parameter estimation in large-scale inter-con-nected systems. In *Proc. of IFAC Symposium on Identification and System Parameter Estimation*, Budapest, Hungary, 729-733.

Rao G.P. and A.V.B. Subrahmanyam (1996). Models in generalized MA form for identification of continuous-time systems. In *"Statistical methods in control and signal processing"* Eds. Katayama and Sugimoto, Marcel Dekker, New York, USA.

Saha, D.C. and G.P. Rao (1980). Identification of lumped linear systems in the presence of unknown initial conditions via Poisson moment functionals. *Intern. J. of Control*, **31**, 637-644.

Saha D.C. and G.P. Rao (1983). *Identification of* continuous-time systems - A Poisson moment *functional approach.* LNCIS Series Vol. 56, Springer-Verlag, Berlin, Germany.

Saha, D.C. and S.K. Mandal (1990). Recursive least squares parameter estimation in SISO systems via Poisson moment functionals. Part 1: Open loop systems. *Int. J. Systems Sciences*, **21**, 1205-1216.

Saha, D. C., B. K. Roy and V. N. Bapat (1991). Some generalizations and further uses of the Poisson moment functionals (PMF) approach to continuous-time model estimation. In *Proc. of the 9th IFAC Symp. on Identification and System Parameter Estimation*, Budapest, Hungary, 1324-1333.

Shinbrot M., (1957). On the analysis of lLinear and nonlinear systems. *Trans. ASME*, **79**, 547-452.

Sinha N.K., G.P. Rao (Eds.) (1991). *Identification of continuous-time systems: Methodology and computer implementation.* Kluwer-Academic Publishers, Dordrecht, The Netherlands.

Söderström T. and P. Stoica (1989). *System identification.* Prentice Hall, Hemel Hempstead, UK.

Subrahmanyam A.V.B. and G.P. Rao (1993). Identification of continuous-time SISO systems via Markov parameter estimation. *IEE-Proc., Part D*, 140, 1-10.

Subrahmanyam A.V.B., D.C. Saha and G.P. Rao (1995). Identification of continuous-time MIMO systems via time moments. *Control Theory and Advanced Technology*, **10**, 1359-1378.

Subrahmanyam, A.V.B., D.C. Saha and G.P. Rao (1995). Identification of continuous-time MIMO systems via time moments. *Control-Theory and Advanced Technology*, **10**, 1359-1378.

Subrahmanyam, A.V.B., D.C. Saha and G. P.Rao (1996). Irreducible continuous model identif-ication via Markov parameter estimation. *Automatica*, **32**, 249-253.

Sugeno, M. and K. Tanaka (1991). Successive identification of a fuzzy model and its applications to prediction of a complex system. *Fuzzy sets and Systems*, **42**, 315-334.

Takagi, T. and M. Sugeno (1985). Fuzzy Identification of Systems and its Application to Modeling and Control. *Proc. IEEE Trans. on Systems, Man, and Cybernetics*, **15**, 116-132.

Tether A. (1970). Construction of minimum state-variable models from finite input-output data. *IEEE Trans. AC*, **17**, 427-436.

Tsang, K.M. and S.A. Billings (1994). Identification of continuous-time nonlinear systems using delayed state variable filters. *Intern. J. of Control*, **60**, 159-180.

Unbehauen H. and G.P. Rao (1987). *Identification of continuous-time systems.* North Holland, Amsterdam, The Netherlands.

Unbehauen H. and G.P. Rao (1990). Continuous-time approaches to system identification - A survey. *Automatica*, **26**, 23-35.

Unbehauen, H. (1995). Identification of nonlinear continuous-time systems by Hartley-transformation. In *Proc. EURACO Workshop*, Florence, Italy, 323-333.

Wahlberg B. and L. Ljung (1986). Design variables for bias distribution in transfer function estimation. *IEEE Trans. AC*, **31**, 134-144.

Wahlberg B. (1991). Identification of resonant systems using Kautz filters. *In Proc. of IEEE Conference on Decision and Control*, Brighton, UK, 2005-2010.

Wahlberg B. (1994). Laguerre and Kautz models. In *Proc. of SYSID'94 10th IFAC Symposium on Identification and System Parameter Estimation*, Copenhagen, Denmark, 1-12.

Wolovich W.A. (1974). *Linear multivariable systems.* Springer-Verlag, New York, USA.

Yang Z.J., S. Sagara and K. Wada (1993). Identification of continuous-time systems from sampled input-output data using bias eliminating techniques. *Control Theory and Advanced Technology*, **9**, 53-75.

Young, P C. (1981). Parameter estimation of contin-uous-time models – a survey. *Automatica*, **17**, 23-39.

Zadeh L. A. (1962). From circuit theory to system theory. *Proc. IRE*, **50**, 856-865.

Zhao Z.Y., S. Sagara and K. Wada (1991). Bias-compensating least-squares method identification of continuous-time systems from sampled data. *Intern. J. of Control*, **53**, 445-461.

Zhao Z.Y., S. Sagara and M. Tomizuka (1992). A new bias-compensating least-squares method for continuous system identification in the presence of coloured noise. *Intern. J. of Control*, **56**, 1441-1452.