

IDENTIFICATION OF LINEAR MULTIVARIABLE CONTINUOUS-TIME SYSTEMS

by

ZEINAB HANAFY ASHOUR ABU-EL-MAGD

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree of

Doctor of Philosophy

McMaster University

February, 1986



IDENTIFICATION OF LINEAR MULTIVARIABLE CONTINUOUS-TIME SYSTEMS

To the memory of my father

and

for MAHA, ALI and MOHAMED

DOCTOR OF PHILOSOPHY (1986)  
(Electrical Engineering)

McMaster University  
Hamilton, Ontario

**TITLE:** IDENTIFICATION OF LINEAR MULTIVARIABLE CONTINUOUS-  
TIME SYSTEMS

**AUTHOR:** Zeinab H. Ashour Abu-El-Magd  
B.Sc. (Elec. Eng., Cairo University)  
M.Eng. (Elect. Eng., McMaster University)

**SUPERVISOR:** Naresh K. Sinha, Chairman of the Electrical  
Engineering Department  
(McMaster University)  
B.Sc. Eng. (Banaras)  
Ph.D. (Manchester)

**NUMBER OF PAGES:** xiv, 171

## ABSTRACT

The problem of identification of linear multivariable continuous-time systems from input-output data is considered. A survey has been made to present the direct and the indirect approaches in identifying continuous-time systems from the samples of the observations. The direct approach with approximate integration seems to be more promising and hence it is adopted in this work. Three direct methods based on the use of block pulse functions, trapezoidal pulse functions and cubic splines have been compared and applied for multivariable systems. A comprehensive study has been conducted to analyze the effect of noise on the identification. The analysis was carried out first for the single-input single-output case and then extended to the multivariable case. A new approach is presented to overcome the combined effect of the errors in the approximation and additive white noise on the identification of continuous-time systems. The method consists of modelling the combined error term. Extensive simulations are conducted in order to illustrate the merits of the new procedure. The problem of order determination has been considered and three order determination tests have been studied and applied for continuous-time systems, two of them for the first time as far as the author is aware.

The problem of the selection of the structure that will give good conditioned parameterization is also considered. A new procedure to identify the structure in the input-output form is presented. This procedure is suitable for both stationary and non-stationary systems when a change in the structure occurs while the order remains constant. It uses the concept of overlapping parameterization to choose a better conditioned parameterization for the multivariable system whenever ill conditioning is detected. A switching criterion is presented based on the complexity principle which provides a good monitor of the conditioning of the parameterization as well as the suitability of the tested structure.

#### ACKNOWLEDGEMENTS

The author wishes to express her sincere appreciation to Dr. N.K. Sinha for his expert guidance and supervision throughout the course of this work. She also thanks Dr. D.P. Taylor, Dr. I. Yellowlee and Dr. J.F. MacGregor, members of her supervisory committee, for their continuing interest.

It is the author's pleasure to acknowledge useful discussions with other colleagues at McMaster. Without being exhaustive they are Dr. M. El-Sobki, Dr. A. Salama, Dr. I. El-Nahas, Dr. A. Karnik, J. Raol and S. Puthenpura.

Thanks are due to Mrs. Dianne Crabtree for her cheerful and excellent typing of the manuscript.

The author especially wishes to express her deep gratitude and appreciation for the encouragement, help and patience from her husband, Mohamed.

## TABLE OF CONTENTS

	PAGE
ABSTRACT	iii
ACKNOWLEDGEMENTS	v
LIST OF FIGURES	x
LIST OF TABLES	xiii
CHAPTER 1 INTRODUCTION	1
CHAPTER 2 IDENTIFICATION OF CONTINUOUS-TIME MODELS FROM INPUT-OUTPUT DATA: A SURVEY	7
2.1 Introduction	7
2.2 The Indirect Approach	8
2.2.1 Estimation of the Parameters of a Discrete- Time Model from the Input-Output Data	8
2.2.2 The Recovery of the Continuous-Time Model Corresponding to the Discrete-Time Model	10
2.3 The Direct Approach	12
2.3.1 Cubic Splines	14
2.3.2 Walsh Functions	16
2.3.3 Block Pulse Functions	18
2.3.4 Trapezoidal Pulse Functions	24



TABLE OF CONTENTS (continued)

	PAGE
2.4 Concluding Remarks	26
2.4.1 Promising Research Areas	29
CHAPTER 3 THE DIRECT APPROACH FOR SINGLE-INPUT SINGLE-OUTPUT CONTINUOUS-TIME SYSTEMS	31
3.1 Introduction	31
3.2 System Identification Using the Direct Approach	32
3.2.1 Comparison of Three Direct Methods	34
3.2.1.1 Identification in the Noise Free Case	35
3.2.1.2 Identification in the Presence of Noise	37
3.3 Effect of Approximate Integration on Identification	43
3.3.1 Problem Formulation	45
3.4 A Proposed Approach for Modelling the Error Term	49
3.4.1 The Noise-Free Case	51
3.4.2 The Noisy Case	56
3.5 Concluding Remarks	60
CHAPTER 4 PARAMETER ESTIMATION OF MULTIVARIABLE CONTINUOUS-TIME SYSTEMS	63
4.1 Introduction	63
4.1.1 Statement of the Problem	64
4.2 Identification in the Transfer Function Form	67
4.2.1 Identification in the Noise-Free Case	69
4.2.2 Identification in the Presence of Noise	72

**TABLE OF CONTENTS (continued)**

	PAGE
4.3 Effect of Approximate Integration on the Identification in the Multivariable Case	72
4.4 The Error Modelling Approach	76
4.4.1 Results of Simulation	77
4.5 Identification in the Input-Output Form	84
4.5.1 Results of Simulation	86
4.5.2 Parameter Estimation with the Error Modelling Approach	87
4.6 Concluding Remarks	94
<b>CHAPTER 5 ORDER DETERMINATION OF MULTIVARIABLE CONTINUOUS-TIME SYSTEMS</b>	<b>96</b>
5.1 Introduction	96
5.2 Reformulation of the Information Matrices	100
5.3 The Instrumental Variable Method of Model Order Identification	104
5.4 The Instrumental Determinant Ratio Test (IDR)	105
5.5 The Residual Error Technique (RET)	106
5.6 Results of Simulation	108
5.7 Concluding Remarks	119
<b>CHAPTER 6 STRUCTURE SELECTION FOR MULTIVARIABLE CONTINUOUS TIME SYSTEMS</b>	<b>125</b>
6.1 Introduction	125
6.2 A New Recursive Structure Selection Procedure for Multivariable Continuous-Time Systems in the Input-Output Form	127

TABLE OF CONTENTS (continued)

	PAGE
6.2.1 The Model Identification	130
6.2.2 The Switching Criterion	132
6.2.3 Selection and Transformation to a Better Structure	136
6.2.4 Summary	136
6.3 Results of Simulation	137
6.3.1 Noise-Free Case	138
6.3.2 The Noisy Case	139
6.4 Discussion	140
6.5 Concluding Remarks	141
CHAPTER 7 CONCLUSIONS	143
7.1 Suggestions for Further Research	147
APPENDIX A The Complexity Criterion	151
APPENDIX B Multistrucre Input-Output Models	159
REFERENCES	164

## LIST OF FIGURES

FIGURE		PAGE
2.1	The first four Walsh functions	17
2.2	Block pulse functions and integrations	20
2.3	$y(t)$ as approximated by the trapezoidal pulse functions	25
3.1	Autocorrelations and partial autocorrelations of the error series	48
3.2	Autocorrelations and partial autocorrelations of the residuals after modelling $e_t$	52
3.3	Autocorrelations and partial autocorrelations of the error term without modelling. Noise-free case	54
3.4	Autocorrelations and partial autocorrelations of the residuals with IMA (0,3,3) for the error term. Noise-free case	55
3.5	The autocorrelations and partial autocorrelations of $w_t$ without modelling at 20% noise level	57
3.6	The autocorrelations and partial autocorrelations of the residuals when modelling $w_t$ with ARMA(1,1) at 20% noise level	58
4.1	Autocorrelations and partial autocorrelations of the error series without modelling at 20% noise level. First subsystem	78
4.2	Autocorrelations and partial autocorrelations of the error series without modelling at 20% noise level. Second subsystem	79
4.3	Autocorrelations and partial autocorrelations of the error series after modelling (20% noise). First subsystem	81

LIST OF FIGURES (continued)

FIGURE		PAGE
4.4	Autocorrelations and partial autocorrelations of of the error series after modelling (20% noise). Secqnd subsystem	82
4.5	Autocorrelations and partial autocorrelations of the error series without modelling at 20% noise level. First subsystem	88
4.6	Autocorrelations and partial autocorrelations of the error series without modelling at 20% noise level. Second subsystem	89
4.7	Autocorrelations and partial autocorrelations of the error series after modelling at 20% noise level. First subsystem	92
4.8	Autocorrelations and partial autocorrelations of the error series after modelling at 20% noise level. Second subsystem	93
5.1	Iterative procedure of tests based on analysing the residuals	99
5.2	The estimation of $n_1$ with the ordinary IV test	111
5.3	The estimation of $n_3$ with the ordinary IV test	112
5.4	The estimation of the structural indices with the adaptive IV test	114
5.5	The determination of $n_1$ with the IDR test	115
5.6	The determination of $n_2$ with the IDR test	116
5.7	The determination of $n_3$ with the IDR test	117
5.8	The determination of $n_1$ with the IDR test at SNR = 10db	118
5.9	Residual difference plot of the first output	121
5.10	Residual difference plot of the second output	122

LIST OF FIGURES (continued)

FIGURE		PAGE
5.11	Residual difference plot of the third output	123
5.12	Residual difference plot of the second output at S = 150 and 600 samples	124
6.1	Flow chart of the structure selection procedure	129
6.2	Parameter and Residues Estimation	134

LIST OF TABLES

TABLE		PAGE
3.1	Comparison of the three approximate integration methods for the noise-free case	36
3.2	Comparison of execution time of the three direct methods to approximate the first, second and third integrals of 500 samples. (Cyber 730)	37
3.3	Second order system with 10% and 20% noise levels 500 input-output samples. $T = 0.01$ sec.	39
3.4	Comparison of three different sampling intervals	40
3.5	Third order system with 10% and 20% noise levels 1000 input-output samples. $T = 0.007$ sec.	42
3.6	Comparison of three different sampling intervals	44
3.7	Variance of the residuals with and without modelling $e_t$	53
3.8	Variance of the residuals and parameter error norm for different noise levels with and without modelling $w_t$	59
3.9	Third order system with $T = 0.02$ sec. 10% noise - 1000 input-output data	59
3.10	Third order system with $T = 0.01$ sec. 10% noise - 1000 input-output data	60
4.1	Comparison of three direct methods in the noise-free case. First subsystem	71
4.2	Comparison of three direct methods in the noise-free case. Second subsystem	71
4.3	Comparison of three direct methods in the noisy case	73

LIST OF TABLES (continued)

TABLE		PAGE
4.4	The values of $Q$ when $w_t$ is not modelled	80
4.5	Results of simulation of the given example after 400 iterations at $T = 0.05$ sec	83
4.6	Estimates of $Q$ when $w_t$ is modelled	84
4.7	Comparison of three direct methods in the noisy case	90
4.8	Estimated values of $Q$ when $w_t$ is modelled	91
4.9	Estimates of $Q$ when $w_t$ is modelled	91
4.10	Results of simulation of the given example after 200 iterations at $T = 0.1$ sec	95
5.1	Variance of the residuals White noise, SNR = 10 db, $S = 500$ samples adaptive auxiliary model	113
5.2	Time comparison based on $S = 600$ samples	113
6.1	Structure selection in noise-free case	138
6.2	Structure selection in noisy case	139



## CHAPTER 1

### INTRODUCTION

Identification and process parameter estimation is a rapidly developing field. The interest in system identification is essentially due to the needs of engineers working in process industries to obtain a better knowledge about their plants for improved control. Several survey papers and books [1-10] have been written on the subject and lots of control applications have benefitted from this abundant research work. Applications in bioengineering and in econometrics [1] are also developing in parallel. The advent of computers and their accessibility has revolutionized system identification techniques which have found their way into the new sophisticated fields of robotics, satellites control and artificial intelligence.

The research done in the sixties dealt only with single-input single output systems. Identification of multivariable systems was approached carefully in the seventies due to the considerable difficulties associated with their identification. Most of this work, however, was dedicated to discrete-time systems. On the other hand, the identified processes themselves are usually described in terms of continuous-time differential equations. Mathematical models encountered in the control of chemical processes, thermal systems,

rocket motor combustion, travelling wave systems, tracer kinetics in health care systems, high speed aerodynamic systems, etc..., involve differential equations.

In view of this wide field of applications specially in adaptive control, two basic approaches to identify continuous time-systems from input-output data have been developed. In the first, a discrete-time model from the samples of the observations is obtained and then a corresponding continuous-time model is derived. The other approach attempts to solve the problem directly and is based on obtaining approximate solutions of differential equations over a time interval. It is called the direct approach with approximate integration and should be particularly well suited to certain types of adaptive control where rapid identification is a prime requirement.

Several difficulties are, however, associated with the identification of continuous-time linear systems and need to be studied before real life applications could be realized.

Until very recently the literature lacked a reliable method to approximate the actual continuous-time signals from the available samples. Three methods have been proposed lately and have been applied mainly to single-input single-output systems. These are based on the use of block-pulse functions, trapezoidal pulse functions and cubic splines, respectively as approximating functions and are promising in fast recursive identification.

Most of the practical systems have considerable measurement noise, hence in certain cases the available information to the control

engineer is not accurate. In the discrete-time case the problem of noisy data is not as severe as in the continuous-time case. The approximation and other factors make the continuous-time system very sensitive to the noise level superimposed on it. Hence, a realistic identification technique should take into consideration the noise factor.

The problem of system identification consists of two major steps: structure determination and parameter estimation. Several structure and order determination procedures have been proposed and applied to discrete-time systems. Since continuous-time systems have been long neglected the application of such procedures with the new continuous-time identification techniques presents a new and important research area.

Multivariable systems, unlike single-input single-output systems can be represented within different structures. The problem is to choose one which gives a well conditioned parameterization. To solve this problem several approaches have been suggested in the literature, such as using physical a priori knowledge about the system and canonical parameterizations. But all these methods have disadvantages or are not suitable for practical situations. The overlapping parameterization approach presents a realistic approach to the problem. It calls for the transformation from a given structure to another equivalent one when ill conditioning is detected.

The major effort in this thesis is directed toward the problem of identification of linear time-invariant continuous-time

multivariable systems from samples of input-output data. The approach adopted in the identification is the direct approach by approximate integration. It is based on obtaining approximate expressions for the signals from their samples. The differential equations are then integrated using these approximations and the results are used for estimating the parameters of the continuous-time system. The three-aspects of the system identification problem: order determination, structure selection and parameter estimation, have been considered in this work.

In Chapter 2 a comprehensive survey, that discusses and exposes two relatively new approaches in continuous-time systems identification from input-output data, is conducted. The development of the indirect and the direct approaches as well as the various resulting techniques are given in detail. A constructive discussion of the advantages for each approach at the end of the survey clarifies and points out the needed and unaddressed research topics.

Chapters 3 and 4 discuss the problem of identification of continuous-time systems with the direct approach by approximate integrations. Three parameter estimation algorithms have been developed. In each algorithm a different method is used to approximate the input-output signals and their respective integrals. Block pulse functions, trapezoidal functions and cubic splines are used as approximating functions. A comparison based on simulated examples for single-input single-output and multivariable systems in both noise-free

and noisy cases is presented to compare computation time and the accuracy of the estimated parameters.

It has been noticed that when noise was added to any studied system, identification problems such as correlated residuals and inaccurate parameter estimates were created. So a comprehensive study is conducted to analyze the reasons behind the previously mentioned problems and to investigate whether the errors caused by the approximation constituted one of those reasons. The analysis is carried out first for the single-input single-output case and then extended to the multivariable case.

A new approach is presented to overcome the combined effect of the errors in the approximation and additive white noise on the identification of continuous-time systems. The method consists of modelling the combined error term. Extensive simulations are conducted in order to illustrate the merits of the new procedure.

Order determination is the first part of the identification problem and it is done, generally, off-line. In Chapter 5 three order determination tests have been studied and applied for continuous-time systems, two of them for the first time, as far as the author is aware. The sampled input, output signals have been integrated with the cubic spline technique and the information matrix has been reformulated to suit the new interpretation of the data. The three order determination tests have been compared according to the computation time, the number of input-output samples used to indicate the correct order and their robustness to added noise.

Chapter 6 presents a new procedure to identify the structure of continuous-time multivariable systems in the input-output form. This procedure is suitable for both stationary and non-stationary systems when a change in the structure occurs while the order remains constant. It uses the concept of overlapping (multistructural) parameterization to choose a better conditioned parameterization for the multivariable system whenever ill conditioning is detected.

A switching criterion is presented based on the complexity principle which provides a good monitor of the conditioning of the parameterization as well as the suitability of the tested structure. The inclusion of this criterion in the selection procedure answered the question of when to switch to another pseudostructure. This leads to the reduction of time required for computations when compared with the other structure selection procedures. All of the latter procedures lack a switching criterion.

Conclusions and suggestions for future investigation in the problem of identification of linear continuous-time multivariable systems are discussed in Chapter 7.

## CHAPTER 2

### IDENTIFICATION OF CONTINUOUS-TIME

#### MODELS FROM INPUT-OUTPUT DATA: A SURVEY

##### 2.1 Introduction

Finding the simplest system that will realize a prescribed input-output behaviour has been a fundamental problem in systems theory [11]. Only a few techniques were generally known and they were based on repeated differentiation and contained all the inherent disadvantages involved therein [11-13].

In the fifties the methods used to model the dynamic systems were pure analogue, such as the frequency response and transient response methods. During the sixties and with the introduction of the digital computer most of the research was directed toward obtaining discrete-time models for the dynamic systems. The interest in obtaining continuous-time models for continuous systems from the input and output samples has been renewed toward the end of the decade. Two new theories were developing from scattered research efforts. Those two theories, called the "direct" and the "indirect" approaches, will be discussed in the following sections. The evolution of both techniques and the major contributions are stated in sections 2.2 and 2.3. The advantages and the difficulties facing each approach are discussed in

section 2.4. Also, a comparison is drawn based on several published papers on the subject stressing the accuracy of the estimated parameters and the computation time. Several promising research topics are outlined in the area of continuous-time systems identification using the direct approach.

## 2.2 The Indirect Approach

The arrival of the digital computer generated a wider use of discrete-time mathematical models for continuous-time processes. Unfortunately such models do not always provide enough information for the analysis of the process. Many efforts have then been directed toward identifying continuous-time models from existing discrete-time models [14-16]. This is called the indirect approach. This approach calls for the derivation of an appropriate discrete model  $H(z)$  for a given analog system  $G(s)$  with undetermined parameters and then the fitting of the sampled input-output data by  $H(z)$  to estimate these parameters.

### 2.2.1 Estimation of the Parameters of a Discrete-Time Model from the Input-Output Data

Exact transformation of the continuous system into the discrete system does not exist, if there is no knowledge of the actual variation of the input between the sampling instants [17]. The unknown variations of the input signals between the sampling intervals could be interpreted as some kind of noise and this causes problems when



identifying the discrete-time model of the system [18-19]. We could use steps or rectangular pulses or white noise [16] as inputs to solve the identification problem but this may not be true for actual plant records.

The choice of the sampling interval is very important. A large value could mean losing important information about the system and a very small value may cause mathematical difficulties. It can be assumed that the sampling interval  $T$  is selected such that  $\lambda_f T \leq 0.5$  where  $\lambda_f$  is the eigenvalue of  $A$  farthest from the origin of the complex plane (Haykin [20]). The matrix  $A$  corresponds to the state space representation of the continuous system in equation (2.1).

$$\begin{aligned}\dot{\underline{x}}(t) &= A\underline{x}(t) + B\underline{u}(t) \\ \underline{y}(t) &= C\underline{x}(t) \\ \underline{z}(t) &= \underline{y}(t) + \underline{w}(t)\end{aligned}\tag{2.1}$$

where  $\underline{x}(t)$  is the  $n$ -dimensional state vector,  $\underline{u}(t)$  is the  $m$ -dimensional input vector and  $\underline{y}(t)$  is the  $p$ -dimensional noise-free output vector,  $\underline{z}(t)$  is the measured output vector and  $\underline{w}(t)$  is considered to be a zero-mean white noise sequences.

If a proper sampling interval is chosen [21] and the input signal did not vary between intervals then any discrete-time structure determination and parameter estimation method could be applied to the

samples of input and output data [22-23]. Putting the final identified model into a canonical discrete state space representation was also discussed by Guidorzi [24].

### 2.2.2 The Recovery of the Continuous-Time Model Corresponding to the Discrete-Time Model

Several authors have contributed to this part of the problem. The problem may be stated as follows [14]:

Given the pulse transfer function  $H(z^{-1})$ , determine the continuous system transfer function  $G(s)$ , so that the responses of these correspond closely at the sampling instants for all possible inputs. Many of the works done in the transformation area from continuous to discrete was dedicated only to univariate systems [14] [16-17]. The identification of the multivariable continuous-time systems through the indirect method has been discussed by Sinha and Zhou QiJie [25-26] and by Strmcnik and Bremsak [15]. Several transformations were suggested [14] based on the approximation of the input as a step or a ramp but they involved long calculations and were vulnerable in practical cases.

The bilinear  $z$  transformation [14] given by (2.2) or (2.3) may be more suitable and is more flexible in regard of input variation between sampling intervals.

$$s = \frac{2}{T} \left( \frac{1 - z^{-1}}{1 + z^{-1}} \right) \quad (2.2)$$

or

$$z^{-1} = \frac{2 - sT}{2 + sT} \quad (2.3)$$

A comparison has been made between all these methods both in the univariate and the multivariable cases. The reader may be referred to [14], [25].

The most well known indirect method is the state-transition method [26]. If the input is assumed to be held constant during each sampling interval hence equation (2.4) is the discrete-time equivalent of equation (2.1).

$$\underline{x}(k+1) = \underline{F}\underline{x}(k) + \underline{G}u(k) \quad (2.4)$$

$$\text{and } \underline{F} = e^{AT} \quad (2.5)$$

$$\underline{G} = \int_0^T e^{At} B dt \quad (2.6)$$

The problem of estimating the parameters of the discrete-time model described by equation (2.4) has been discussed by several authors [22-24]. The next problem, then, is to determine A and B from the estimates of F and G. This can be solved easily if F can be diagonalized. Some difficulty arises when the eigenvalues of F are either complex or negative [27].

### 2.3 The Direct Approach

It has been always desired to utilize the differential equation of the dynamic system in the determination of the system parameters. This way the parameter estimation problem will not be divided into two subproblems as in the "indirect" approach.

Techniques previously used involved direct differentiation [11-13] of the signals which had the disadvantage of amplifying the superimposed noise. Another approach was to derive a discrete-time equivalent of a continuous system by the use of finite-difference approximation. Lanczos [28] showed that the latter method yields inaccurate parameter estimates if the data is contaminated with measurement errors.

Based on the concepts of avoiding the noise accentuating operation of signal differentiation or the use of a modest equivalent discrete-time model, the "direct" approach has been gradually evolving.

Consider the dynamical single-input single-output system described by the following differential equation

$$A(D) y(t) = B(D) u(t) \quad (2.7)$$

or more explicitly in the canonical input-output form

$$D^n + a_{n-1} D^{n-1} y(t) + \dots + a_0 y(t) = b_m D^m u(t) + \dots + b_0 u(t) \quad (2.8)$$

$m < n$

The problem is to estimate the unknown parameters of the system without directly differentiating the input and output signals.

So, as pointed out by Rao and Sivakumar [29], if we perform a linear operation on each side of (2.7) this would enable the generation of transformed input and output signals easy to measure or to compute.

Shinbrot [30] introduced the technique of "method functions" based on multiplying the input and output signal families by suitable functions followed by integration over finite limits in time domain. Loeb and Cahan [31] used modulating functions in a similar manner to Shinbrot. Fairman et al. [32] employed filter chains with each unit having a transfer function of  $\frac{1}{s + \lambda}$ ,  $\lambda > 0$ .

It was Diamessis [33] who suggested utilizing successive integration in identification schemes for linear continuous systems. Hence the differential input-output model of equation (2.7) is converted into a linear algebraic model suitable for a least squares solution. The next natural step was to think of utilizing the digital computer as a tool in the parameter estimation process provided that a reliable approximate integration method is used. Since the only available information is the input and output samples, the input and output signals should be approximated in order to perform the operation of successive integration.

The direct approach makes use of Galerkin's approximation to the solution of the differential equation (2.9)

$$Dy = u(x) \quad (2.9)$$

thus

$$y(x) \approx \sum_{i=0}^N y(x_i)w_i(x) \quad (2.10)$$

The Galerkin approximating functions  $w_i(x)$  are free to be selected for better convergence rate [34]. In general  $w_i(x)$  are eigen functions, polynomials or cubic splines [35].

Several approximating functions were suggested in the literature and will be presented in the following sections.

### 2.3.1 Cubic Splines

The cubic spline function  $S(t)$  is a piecewise cubic function defined with the following properties.

- A -  $S(t_i) = y_i$  where  $y(t)$  is the function to be approximated
- B -  $S(t)$  has continuous first and second derivatives for  $0 \leq t \leq t_n$
- C -  $S(t)$  is a polynomial of order 3 in each interval  $t_{i-1} \leq t \leq t_i$ .

Bellman [36] proposed the use of splines for the identification of single-input single-output systems. Shridhar et al. [35] extended their use to the multivariable case. Their algorithm is based on using the cubic splines as approximating functions then integrating the differential equation with Simpson's rule. By simplifying the obtained

expression with the splines continuity relations, the parameters of a second order system can be easily computed. For higher order systems, they assumed that the various integrals for the different signals are available for measurements. Sinha and Zhou QiJie [37] developed a more complex recursive algorithm which enables the calculation of higher integrals using the cubic splines properties.  $S_1(t)$  is supposed to have the following form (2.11) in the interval  $[t_{i-1}, t_i]$

$$S_1(t) = \frac{t_i - t}{T} y_{i-1} + \frac{t - t_{i-1}}{T} y_i + (at+b)(t-t_{i-1})(t-t_i) \quad (2.11)$$

$$\text{where } T = t_i - t_{i-1} \quad (2.12)$$

The values of  $a$  and  $b$  are calculated by applying condition (B) and replaced in equation (2.11). The third order polynomial  $S_1(t)$  is then integrated by the trapezoidal rule from  $t = t_{i-1}$  to  $t = t_i$ . Recursive formulae were developed for the first, second and third order integrals [37] as a function of the available observations and the calculated derivatives of  $S_1(t)$ . Equation (2.13) gives the third order integral.

$$I_{3,i} = I_{3,i-1} + T I_{2,i-1} + \frac{T^2}{2} I_{1,i-1} + \frac{1}{30} T^2 y_i + \frac{2}{15} T^3 y_{i-1} \\ - \frac{1}{120} T^4 m_i + \frac{1}{60} T^4 m_{i-1} \quad (2.13)$$

$$\text{where } m_i = \left. \frac{d}{dt} S(t) \right|_{t = t_i}$$

### 2.3.2 Walsh Functions

The approximation of a function by a linear combination of a set of orthogonal basis functions was a classic tool in numerical analysis [28]. The Walsh functions [38] have long been known to constitute a complete orthonormal set of rectangular waveforms. Chen and Hsiao [39] expanded the systems variables in Walsh series and introduced an operational matrix to perform integration of Walsh functions. This is explained by the following.

Consider the first four ( $m = 4$ ) Walsh functions, in Figure 2.1,  $\phi_k(t)$ ,  $k = 0, 1, \dots, m-1$ .

$$\text{with } \phi_k(t) = R_s^{d_s}(t) R_{s-1}^{d_{s-1}}(t) \dots R_1^{d_1}(t) \quad (2.14)$$

$$k = 0, 1, \dots, 2^s - 1$$

$R_s(t)$ ,  $s = 1, 2, \dots, n$  are a set of orthonormal square waves for  $0 \leq t < 1$  with unit height and repetition rate equal to  $2^{s-1}$ .  $(d_s d_{s-1} \dots d_1)$  is the binary expression of the decimal number  $k$  [34], [40]. The initial conditions for (2.12) are  $\phi_0(t) = R_0(t) = 1$ . Integrating the four Walsh functions of Figure 2.1 from 0 to  $t$  we obtain

$$\int_0^t \begin{bmatrix} \phi_0(\tau) \\ \phi_1(\tau) \\ \phi_2(\tau) \\ \phi_3(\tau) \end{bmatrix} d\tau = \begin{bmatrix} 1/2 & -1/4 & -1/8 & 0 \\ 1/4 & 0 & 0 & -1/8 \\ 1/8 & 0 & 0 & 0 \\ 0 & 1/8 & 0 & 0 \end{bmatrix} \begin{bmatrix} \phi_0(t) \\ \phi_1(t) \\ \phi_2(t) \\ \phi_3(t) \end{bmatrix} \quad (2.15)$$



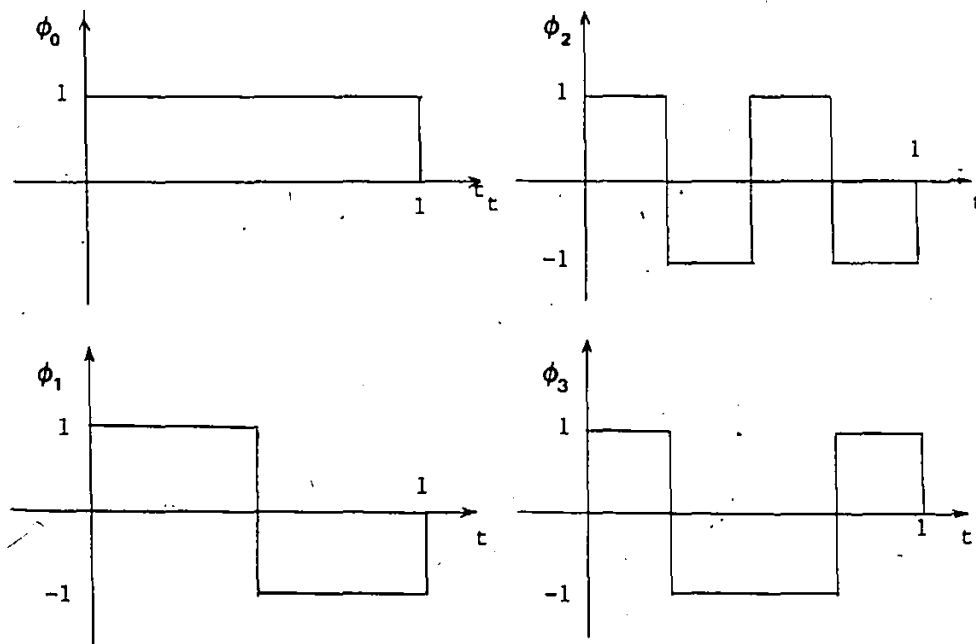


Figure 2.1 The first four Walsh functions

Equation (2.13) can be written in a more general form as

$$\int_0^t \underline{\phi}(\tau) d\tau = P\underline{\phi}(t) \quad (2.16)$$

with

$$P = \begin{bmatrix} P_m/2 & \frac{1}{2m} & I_m/2 \\ \frac{1}{2m} & I_m/2 & O_m/2 \end{bmatrix}, \quad P_1 = \frac{1}{2} \quad (2.17)$$

for  $m = 2^8$

Equation (2.17) is the key of the orthonormal series approach to system identification. The integration is reduced to the multiplication of  $\underline{\phi}(t)$  by  $P$ . Walsh functions have been used for identification by Chen and Hsiao [39], Rao and Sivakumar [41] and others. A review on this matter is presented by Tzafestas [34].

### 2.3.3 Block Pulse Functions

The set of the block pulse functions is more fundamental than the Walsh functions. In a unit interval they are defined by

$$\phi_i(t) = \begin{cases} 1 & \text{for } (i-1)/m < t < i/m, \quad i=1,2,\dots,m \\ 0 & \text{otherwise} \end{cases} \quad (2.18)$$



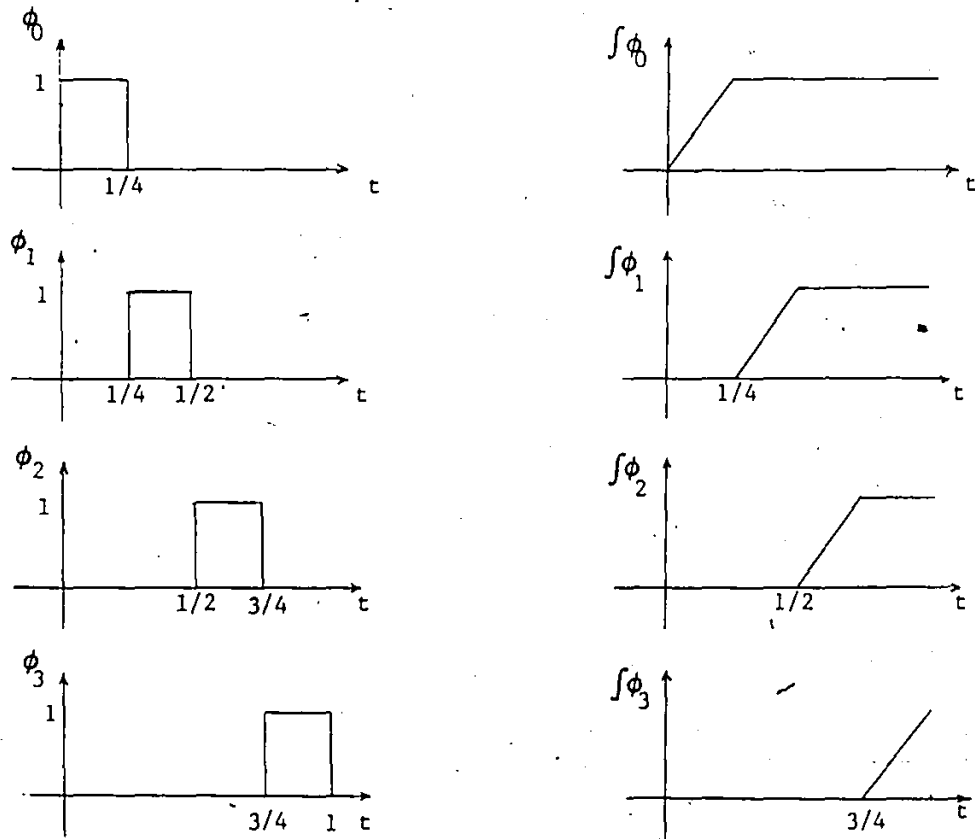


Figure 2.2 Block pulse functions and integrations

then been applied to the identification of bilinear systems [44] and to many other control problems [45-46]. They have been used for identification by Palanisamy and Bhattacharya [47] in the univariate case, and Sinha and Zhou QiJie in the multivariable case [25-26], [48]. The latter has developed a recursive algorithm which performs the numerical integrations without the need for matrix inversion. Their method will be discussed briefly.

Given any function  $y(t)$  integrable over  $(0, T_0)$ , it can be approximated as:

$$y(t) = \underline{\phi}'(t) \underline{y} \quad (2.20)$$

where  $\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_m \end{bmatrix}$  (2.21)

with the superscript ' representing transposition and  $y_i \triangleq$  the average value  $y(t)$  over the interval  $(i-1)T_s, iT_s$

$$T = \frac{T_0}{m}$$

The following integral may be approximated as

$$I_1(t) = \int_0^t y(\tau) d\tau = \int_0^t \underline{\phi}'(\tau) \underline{y} d\tau \quad (2.22)$$

$I_1(t) = T \underline{\phi}'^T P' \underline{y}$  and may be expressed in block pulse form

$$I_1(t) = \underline{\phi}'(t) \underline{I}_1 \quad \text{where } \underline{I}_1 = \begin{bmatrix} I_{1,1} \\ \vdots \\ I_{1,m} \end{bmatrix} \quad (2.23)$$

$I_1$  is the first integral of  $y(t)$  which can be expressed recursively by the relation:

$$I_{1,i} = I_{1,i-1} + \frac{T}{2} (y_i + y_{i-1}) \quad (2.24)$$

Similar derivations are carried out to find higher integrals. The general recursive relation to find  $I_{k,i}$  is expressed by,

$$I_{k,i} = I_{k,i-1} + T I_{k-1,i-1} + \frac{T^2}{2} I_{k-2,i-1} + \frac{T^3}{4!} I_{k-3,i-1} + \dots \quad (2.25)$$

$$+ \frac{T^{k-1}}{2^{k-2}} I_{1,k-1} + \frac{T^k}{2^k} (y_i + y_{i-1})$$

The use of recursive algorithms such as equation (2.25) makes the identification problem an easy task.

The block pulse functions have long been argued as an incomplete orthogonal set [49]. If it is really incomplete then it cannot be guaranteed for any given function, that an arbitrary small mean

error can be obtained by increasing the number of terms in the series. Kwong and Chen [45] have proved the completeness of the block pulse series by studying their convergence properties as  $m \rightarrow \infty$ .

Shieh et al. [43] incorporated a modification, equivalent to the inverse use of the trapezoidal rule, for the block pulse technique. Hung et al. [50] argued that for the case of linear time-invariant systems Shieh et al.'s modification is equivalent to the trapezoidal rule technique. They used the following method to identify the parameters of the single-input single-output system represented by the state equation,

$$\dot{\underline{x}} = \underline{A}\underline{x} + \underline{B}u \quad (2.26)$$

making use of the trapezoidal rule,

$$\underline{x}_k = \underline{A} \frac{\underline{x}_k + \underline{x}_{k-1}}{2} T + \underline{B}T \frac{u_k + u_{k-1}}{2} + \underline{x}_{k-1} \quad (2.27)$$

but they considered the state vector to be known. Sinha [26], [51], modified the method proposed by Hung et al. [50]. He applied the method in the multivariable case and showed that the discrete-time transfer function matrix obtained is identical to the one calculated by the bilinear z transformation [26].

### 2.3.4 Trapezoidal Pulse Functions

The basic idea behind the use of complete orthogonal functions such as the Walsh functions and the block pulse functions is to obtain piecewise constant solutions of linear differential equations over a time interval.

In the block pulse function approach the function is approximated by a sequence of rectangular pulses of height  $y_k$  equal to the mean value of  $y(t)$  over the interval  $kT \leq t < (k+1)T$ . Prasad and Sinha [52] found that a better approximation can be obtained by using piecewise linear approximation, where it is assumed that the function is varying linearly between sampling instants as seen in Figure 2.3 for the subinterval  $kT \leq t < (k+1)T$

$$y(t) = \frac{1}{T} [(k+1)T - t] y_k + (t - kT) y_{k+1} \quad (2.28)$$

$$\text{and } y_k \triangleq y(kT) \quad (2.29)$$

Approximate expressions for the successive integrals were derived in [52] using the trapezoidal rule of integration. The expression for the first integral of  $y(t)$  is given by

$$I_{1,k+1}(y) = I_{1,k}(y) + \frac{T}{2} (y_k + y_{k+1}) \quad (2.30)$$



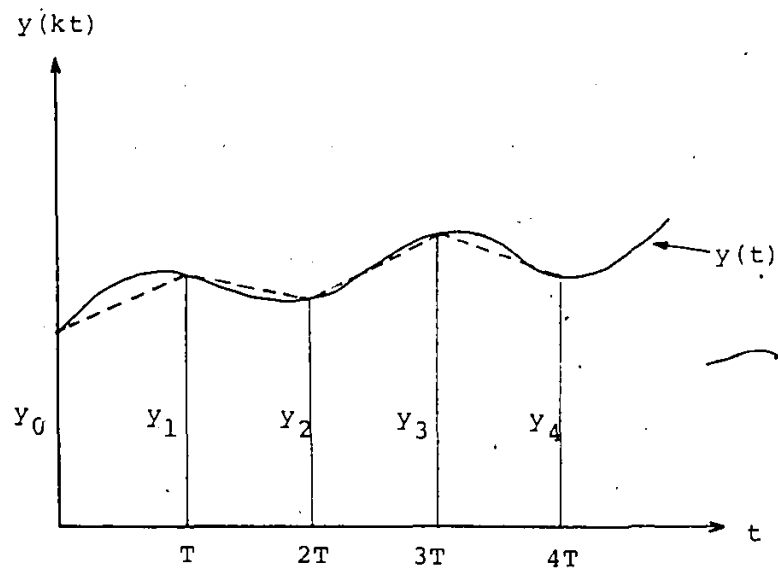


Figure 2.3  $y(t)$  as approximated by the trapezoidal pulse functions

where

$$I_{1,0}(y) = 0 \quad (2.31)$$

The  $n$ th integral of  $y(t)$  can be calculated using the following recursive formula

$$I_{n,k+1}(y) = I_{n,k}(y) + T I_{n-1,k}(y) + \frac{T^2}{2!} I_{n-2,k}(y) + \dots \quad (2.32)$$

$$+ \frac{1}{(n-1)!} T^{n-1} I_{1,k}(y) + \frac{n}{(n+1)!} T^n y_k + \frac{1}{(n+1)!} T^n y_{k+1}$$

with  $I_{n,0}(y) = 0$

Prasad and Sinha [52] used the trapezoidal-pulse functions approach to identify a second order single-input single-output system.

#### 2.4 Concluding Remarks

In this section the advantages and the problems of the indirect and the direct approaches are discussed. Promising research areas in the identification of continuous-time systems from sampled data are outlined.

The "indirect" method involves a long procedure and the accuracy of the solution depends on two factors. First, the satisfactory estimation of the discrete-time model from the available observations. The second difficulty is the derivation of the continuous model from the estimated discrete one so that the responses of both correspond closely at the sampling instants for all possible inputs. There is considerable literature available on the first part of the problem [53-58] which provides acceptable results. Relatively less has been published about the second part [14-17].

Problems are mainly caused by the empirical choice of the sampling interval and the inaccurate assumptions about the variations of the input between sampling instants. However, if the sampling intervals were carefully chosen such that  $\lambda_f T \leq 0.5$  as mentioned by Haykin [20], and the input matched the conditions set by the transformation methods, then the continuous-time system will be accurately identified from the input-output samples.

The determination of a suitable sampling interval from the input-output samples constitutes a problem in system identification. Several methods have been proposed [21], [59-60] focusing on obtaining a sampling interval which will realize  $\lambda_f T$  in the neighbourhood of 0.5 but are sometimes computationally expensive.

The direct method by approximate integration has several advantages over the indirect method. The direct approach does not divide the problem of the identification into two subproblems, as described in sections 2.2 and 2.3. It is in general more efficient

computationally due to the development of the recursive integration algorithms. These algorithms are, also, suitable for on-line identification. The direct method techniques smooth the effect of noise incorporated in the measurement because of the subsequent integrations of the input and the output signals. Sinha and Zhou Qijie [25-26] presented a valuable comparison between five discrete-time approximation methods in the multivariable case. Three of them belong to the indirect approach: the bilinear z-transformation, the state-transition method and the modified state transition method, and two direct methods: the trapezoidal rule and the block pulse function methods. The last three methods were proposed by Sinha and Qijie.

The comparison showed that the accuracy of the approximation depends, to a large extent, on the nature of the input applied to the continuous-time system. For example, if the input is not a piecewise-constant function of time, the state transition method does not give an accurate discrete-time model. In the noise-free case [26] the trapezoidal rule, the block pulse function and the bilinear z-transformation methods gave comparable results followed by the modified state transition method where the input is averaged over the sampling interval instead of being held constant. When the simulated systems were contaminated with a small amount of noise [25] the best results were obtained using the block pulse function method. This was attributed to the fact that the implied integration used by this method smooths out the effect of noise especially when the sampling interval is reduced.

Another reason is that all transformations require some differencing and this causes the noise to be accentuated.

Considering the relative computational effort required for each method [26] the bilinear z-transformation is found to be most convenient for digital simulation if the continuous-time system is described in the transfer function form. If the model is given in the form of state equations the trapezoidal rule is the most convenient for digital implementation.

#### 2.4.1 Promising Research Areas

The direct approach is attractive because of its simplicity and could possibly be used in a wide field of applications, such as digital simulation of continuous-time systems as well as in the identification of the system through samples of the input-output data. Further applications of these methods are possible in digital adaptive control.

Until now very little [25-26] and [35] of the work done has been directed to investigate the identification of multivariable continuous-time systems with the direct approach. Moreover, two of the most promising recursive approximate integration algorithms [37] [52] have only been applied to single-input single-output systems. Those algorithms utilize the trapezoidal pulse functions and the cubic splines as approximating functions.

The direct approach suffers some serious difficulties when the continuous-time system is contaminated with noise. Shridhar et al. [35], Sinha and Zhou [25] noticed that as the noise level increases the

parameter estimates become less accurate. And since in any practical case different noise levels are superimposed on the data this problem should be investigated in order to benefit from the advantages of the direct method in real situations.

In his survey on the parameter estimation methods for continuous-time models, Young [61] states that "it is clearly of little use having a sophisticated parameter estimation algorithm if the model structure is not firmly established". The selection of a suitable order for the dynamic system is an important research area. A lot of work [62-71] has been done for the discrete-time case. Young et al. [71] have applied an order determination procedure based on the instrumental variable method to continuous-time systems. As far as the author is aware, this is the only attempt to apply order selection methods to continuous-time systems beside the work reported in Chapter 5.

A multivariable system can be described within different structures so in addition to determining the order of the system, one should select a suitable structure. This is an interesting area for both discrete-time and continuous-time systems and a lot of work is required specially for on-line applications. Little has been published for discrete-time systems [72-76] and none in the continuous-time case. Those promising research areas are the subject of this thesis.

## CHAPTER 3

### THE DIRECT APPROACH FOR SINGLE-INPUT SINGLE-OUTPUT CONTINUOUS-TIME SYSTEMS

#### 3.1 Introduction

The identification of a suitable model for a process for control purposes is often done, using a digital computer from the samples of input and output observations. In general, the process itself is described in terms of continuous-time state equations. The direct approach attempts to solve this problem directly and is based on obtaining approximate solutions of differential equations over a time interval [34-37]. Several methods of approximation of the actual continuous-time input and output signals, from the available samples, have been proposed. These were presented in detail in Chapter 2. The use of this approximation overcomes the need of resorting to direct differentiation techniques. Actually, the original differential input-output model is converted to linear algebraic (or regression) model convenient for a direct (or a least squares) solution.

In this Chapter three direct methods based on the use of block pulse functions, trapezoidal pulse functions and cubic splines as approximating functions, are utilized in the identification of continuous-time single-input single-output systems. Their relative

performances are compared both in the absence and in the presence of measurement noise.

Since the direct methods are based on approximating the signals from their samples and on approximate integration techniques, a thorough study is performed in section 3.3 to investigate the effect of errors in the approximation on the identification of single-input single-output continuous-time systems. The study also examines the important case of noisy data and draws several new and very useful conclusions.

In section 3.4 a new approach is proposed to overcome the effect of approximate integration on the accuracy of the parameter estimates. A simulated example is presented to show the improvement when the new approach is applied.

### 3.2 System Identification Using the Direct Approach

Consider a single-input single-output system represented by the linear differential equation:

$$\sum_{i=0}^n a_i \frac{d^i}{dt^i} y(t) = \sum_{j=0}^m b_j \frac{d^j}{dt^j} u(t) \quad (3.1)$$

$m < n$

According to Diamessis [33] the successive integration of equation (3.1)  $n$  times over the interval  $(0, t)$  yields:



$$\sum_{i=0}^n a_i I_{n-i} y(t) = \sum_{j=0}^m b_j I_{n-j} u(t) \quad (3.2)$$

The input and output signals may be approximated with any of the previously discussed functions in Chapter 2.

In each case, the system differential equations can be integrated using these approximations. Expressing  $y(t)$  and  $u(t)$  of equation (3.1) in block pulse function series, for example, gives:

$$\left. \begin{aligned} u(t) &= U\phi(t) \\ y(t) &= Y\phi(t) \end{aligned} \right\} \quad (3.3)$$

Equation (3.2) may be written as

$$\sum_{i=0}^n a_i Y_{n-i} = \sum_{j=0}^m b_j U_{n-j} \quad (3.4)$$

where  $U_n = \int \dots \int U \phi(t) dt = UP \phi(t)$

$$Y_{n-i} = YP^{n-i} \quad (3.5)$$

$$U_{n-j} = UP^{n-j}$$

The use of recursive algorithms such as (2.13), (2.25) and (2.32) facilitates the computations of the different integrals. Since the input and output samples are known the respective integrals are easily

calculated. Hence, the only unknowns in equation (3.4) are the model parameters which can be estimated by the least squares algorithm.

### 3.2.1 Comparison of Three Direct Methods

The three direct methods are based on using block pulse functions, trapezoidal pulse functions and cubic splines as approximating functions. Each of these has basic equations for the calculations of various order integrals [37] [48] and [52]. The general recursive relations for the first two methods are given in equations (2.25) and (2.32). No general recursive formula has been developed for the cubic splines method. Sinha and Zhou [37] calculated up to the third successive integral and based on the same principle the author developed the recursive formulae of the fourth and fifth order integrals.

$$I_{5,i} = I_{5,i-1} + T I_{4,i-1} + \frac{T^2}{2!} I_{3,i-1} + \frac{T^3}{3!} I_{2,i-1} + \frac{T^4}{4!} I_{1,i-1} + \frac{T^5}{1120} y_1 - \frac{T^6}{4032} m_1 + T^5 x \quad (3.6)$$

$$7.44037 \times 10^{-3} y_{1-1} + T^5 \times 7.44095 \times 10^{-4} m_{1-1}$$

Computer programs have been developed in Fortran 77 for the three methods and then utilized with simulated systems to compare relative performance in noise-free and noisy environments.

### 3.2.1.1 Identification in the Noise-Free Case

Consider the following third order single-input single-output system given by the following transfer function:

$$G(s) = \frac{1}{(s+1)(s+2)(s+3)} = \frac{1}{s^3 + 6s^2 + 11s + 6}$$

with  $u(t) = \sin t$ , the corresponding output is calculated as

$$y(t) = 0.25e^{-t} - 0.2e^{-2t} + 0.05e^{-3t} - 0.1 \cos t$$

The results for three different sampling intervals are given in Table 3.1.

The parameters were estimated by the least squares algorithm [77]. As a measure of the accuracy in estimating the parameters of the continuous-time model the parameter error norm is calculated for each case and listed in the Table. The parameter error norm is defined as  $\|\underline{\theta} - \underline{\theta}_1\| + \|\underline{\theta}\|$  where  $\underline{\theta}$  and  $\underline{\theta}_1$  are the actual and estimated parameter vectors, respectively [21-22].

The use of cubic splines as approximating functions does clearly improve the parameter estimates of the system under consideration. This is due to the good approximation of the original signal the cubic splines provide. When the number of successive integrations exceeds two, the cubic splines in the noise-free case show superiority to the other two methods.

**Table 3.1 Comparison of the three approximate integration methods for the noise-free case**

Sampling Interval	Method	Number of Samples	$a_2$ (6)	$a_1$ (11)	$a_0$ (6)	$b_0$ (1)	Variance of Residuals	Parameter Error Norm
0.07	BPF	100	5.959	10.95	5.968	0.9953	$0.1824 \times 10^{-8}$	0.00519
	TPF		6.056	11.094	6.054	1.009	$0.201 \times 10^{-9}$	0.00878
	cub sp		6.00	11.00	6.00	1.00	$0.5043 \times 10^{-13}$	0.
0.007	BPF	1000	6.00	11.00	6.00	1.00	$0.1785 \times 10^{-12}$	0.
	TPF		6.00	11.00	6.00	1.00	$0.5605 \times 10^{-13}$	0.
	cub sp		6.00	11.00	6.00	1.00	$0.563 \times 10^{-18}$	0.
0.0035	BPF	2000	6.00	11.00	6.00	1.00	$0.1114 \times 10^{-13}$	0.
	TPF		6.00	11.00	6.00	1.00	$0.3498 \times 10^{-14}$	0.
	cub sp		6.00	11.00	6.00	1.00	$0.5618 \times 10^{-18}$	0.

The spline function, has the best approximation and minimum curvature properties [35], i.e., of all twice differentiable interpolating functions, the cubic spline provides the smoothest interpolation. So by its nature, the cubic splines give better approximation to a given function than the block pulse functions or the trapezoidal pulse functions and this property influenced the parameter estimation process. On the other hand the computation time is larger compared with the case of the block pulse functions or trapezoidal pulse functions as approximating functions, see Table 3.2. The latter methods have also the advantage of being easily incorporated in any on-line algorithm to obtain directly at each sampling instant the approximate value of the function and its successive integrals. To calculate the cubic spline function at a certain sampling instant the value of the first derivative at the same and previous sampling instants should be known in addition to the observed values.

**Table 3.2 Comparison of execution time of the three direct methods to approximate the first, second and third integrals of 500 samples. (Cyber 730)**

Method	BPF	TPF	Cubic Splines
Execution time in seconds for 500 iterations	0.0989999	0.0939999	0.238

### 3.2.1.2 Identification in the Presence of Noise

Two other simulated examples are presented to compare between the three direct methods when the single-input single-output system is contaminated with noise.

Example 1 Second Order System:

Consider the following transfer function:

$$G(s) = \frac{10}{(s+1)(s+10)} = \frac{10}{s^2 + 11s + 10}$$

The input to the system  $u(t)$  is taken as

$$u(t) = \cos 0.47 t + \cos 3.83 t - 2 \cos 7.29 t$$

Assuming zero initial conditions, the output of the system is given by:

$$\begin{aligned} y(t) = & -0.939951e^{-t} + 0.062662e^{-10t} + 0.9049026 \\ & \cos(0.47t - 0.486326) + 0.235916 \cos(3.83t - \\ & 1.681167) - 0.219636 \cos(7.29t - 2.064398) \end{aligned}$$

Two white noise sequences were generated by the computer and added to the output. The sequences had zero means and standard deviations equal to ten percent (10%) and twenty percent (20%) of the output signal respectively (corresponding to a signal to noise ratio (SNR) of 20 db and 14 db, respectively). In Table 3.3, the three approximate integrations methods are compared with respect to the accuracy of the estimated parameters for a sampling interval  $T$  of 0.01 sec.

The cubic splines method does not show any superiority in the noisy case. The three methods gave the same results, almost the same parameter estimates.

Table 3.3 Second order system with 10% and 20% noise levels 500 input-output samples.  $T = 0.01$  sec.

True Parameters	N.S.R. = 10%			N.S.R. = 20%		
	BPF	TPF	Cub. Sp.	BPF	TPF	Cub. Sp.
$a_1$ (11)	10.91	10.91	10.9	10.65	10.65	10.64
$a_0$ (10)	10.15	10.15	10.15	10.16	10.16	10.15
$b_0$ (10)	10.04	10.04	10.04	9.925	9.925	9.918
Parameter Error Norm	0.01	0.01	0.01	0.0219	0.0219	0.0222

While examining this example, it is worthy to illustrate the effect of the choice of the sampling interval on the identification. Three different sampling intervals are chosen, to show how a large sampling interval as well as a small one may affect the accuracy of the estimated parameters. The results are listed in Table 3.4. They were taken with the cubic splines as approximating functions at 10% and 20% noise levels, respectively.

It is clear that taking the sampling interval  $T$  equal to 0.01 sec. gives the best results compared to the other two sampling intervals. A large sampling interval causes loss of information, while a very small sampling interval causes numerical problems, makes the system more sensitive to the noise and increases the computation time as well [37].

Table 3.4 Comparison of three different sampling intervals

Noise to Signal Ratio N.S.R.	Number of Samples	Sampling Interval T(sec.)	Parameters (True Values)			Parameter Error Norm
			a <sub>1</sub> (11)	a <sub>0</sub> (10)	b <sub>0</sub> (10)	
10%	250	0.02	11.45	10.55	10.53	0.049
	500	0.01	10.9	10.15	10.04	0.01
	714	0.007	11.15	10.34	10.21	0.0238
20%	250	0.02	11.05	10.4	10.31	0.0284
	500	0.01	10.64	10.15	9.918	0.0222
	714	0.007	11.07	10.5	10.2	0.0303

A commonly used rule is that the sampling interval T should be selected in such a manner [20] that

$$\lambda_f T \leq 0.5 \quad (3.7)$$

where  $\lambda_f$  is the magnitude of the largest eigenvalue of the continuous-time model. In this example the equality sign is not suitable, it results in a very large value of T. As we proceed in this chapter we will find that a more appropriate rule-of-thumb to choose the sampling interval when using approximate integration methods is the following:

$$0.05 \leq \lambda_f T \leq 0.1 \quad (3.8)$$



Example 2 Third Order System:

In the noise free case, the cubic splines approach showed superiority over the other two techniques because the order of the system exceeded two. To investigate if this property still holds in the noisy environment a third order single-input single-output system is simulated with the following transfer function:

$$G(s) = \frac{s^2 + 2s + 5}{(s+0.1)(s+1)(s+10)} = \frac{s^2 + 2s + 5}{s^3 + 11.1s^2 + 11.1s + 1}$$

$$u(t) = \cos 0.5t + \cos 3t - 2 \cos 7t$$

The output was calculated to be:

$$y(t) = -0.21142 e^{-0.1t} + 0.42469 e^{-t} - 0.05463 e^{-10t} - 0.092269 \cos 0.5t + 0.84538 \sin 0.5t + 0.04413 \cos 3t + 0.0578562 \sin 3t - 0.1105 \cos 7t - 0.1055714 \sin 7t$$

In Table 3.5, the three direct methods are compared for the noise levels of 10% and 20%, respectively, for a sampling interval  $T = 0.007$  sec.

It is clear from Table 3.5 that even for a third order system, the cubic splines approach performed in the same level as the other two methods (the block pulse functions and the trapezoidal pulse functions approaches). Since the method of trapezoidal pulse functions requires less computations it is preferable to use it in practical cases.

**Table 3.5 Third order system with 10% and 20% noise levels 1000 input-output samples.  
T = 0.007 sec.**

Noise Level	N.S.R.	Method	Parameters (True Values)					Parameter Error	
			a <sub>2</sub> (11.1)	a <sub>1</sub> (11.1)	a <sub>0</sub> (1)	b <sub>2</sub> (1)	b <sub>1</sub> (2)	b <sub>0</sub> (5)	Norm
10%		BPF	11.64	10.16	1.021	1.032	2.151	4.673	0.068623
		TPF	11.64	10.16	1.021	1.032	2.151	4.674	0.068605
		cub.sp.	11.64	10.16	1.02	1.032	2.15	4.673	0.068614
20%		BPF	11.53	8.793	0.9961	1.029	2.163	4.149	0.15020
		TPF	11.53	8.794	0.9962	1.029	2.163	4.149	0.15015
		cub.sp.	11.52	8.792	0.9953	1.029	2.16	4.148	0.15016

The estimates of the parameters obtained for different sampling intervals with the trapezoidal pulse functions method are given in Table 3.6. It is clear that  $T = 0.007$  sec. gives the best results which corresponds to  $\lambda_f T = 0.07$ . This value lies within the region indicated in equation (3.8).

It can be noticed that as the noise level increases the parameter estimates deteriorate even after an appropriate sampling interval has been picked up carefully.

This problem is the topic of the following section.

### 3.3 Effect of Approximate Integration on Identification

The input and output signals and their integrals are approximated at the sampling instants either by the block pulse functions, the trapezoidal pulse functions, etc..., and substituted in the differential equation. There always exists an error in the approximation, but does this error affect the identification? Previous work [25] [35] [37] has been conducted to identify continuous-time systems with the direct approach in presence of noise with little success. In [25] the highest noise level added to the system to obtain reasonable parameter estimates was 1.5%, which is actually very low. Sinha and Zhou [37], Shridhar and Balatoni [35] showed also that the higher the noise level the less accurate are the estimates of the parameters. They also mentioned the presence of numerical problems.

Table 3.6 Comparison of three different sampling intervals

Noise Level	Number of Samples	Sampling Interval T	Parameters (True Values)						Parameter Error Norm
			11.1	11.1	1.	1.	2.	5.	
10%	700	0.01	9.82	10.72	0.8791	0.9423	1.772	4.695	0.08375
	1000	0.007	11.64	10.16	1.021	1.032	2.15	4.673	0.0686
	2000	0.0035	10.22	9.992	0.9833	0.9554	1.75	4.575	0.09
20%	700	0.01	8.391	10.17	0.7516	0.8744	1.508	4.324	0.1799
	1000	0.007	11.52	8.792	0.9953	1.028	2.16	4.148	0.15016
	2000	0.0035	8.968	8.814	0.9199	0.8888	1.444	4.078	0.19866

### 3.3.1 Problem Formulation

Let us study the effect of the approximation errors on the identification of continuous-time systems for a second order single-input single-output system and assume that the measurements are noise free to emphasize the effect of the errors in the approximation.

The system is described by the following differential equation

$$\frac{d^2}{dt^2} y(t) + a_1 \frac{d}{dt} y(t) + a_0 y(t) = b_1 \frac{d}{dt} u(t) + b_0 u(t) \quad (3.9)$$

Integrating twice and assuming zero initial conditions we obtain

$$y(t) + a_1 \int y(t) dt + a_0 \iint y(t) dt = b_1 \int u(t) dt + b_0 \iint u(t) dt \quad (3.10)$$

or

$$y(t) + a_1 I_1(y) + a_0 I_2(y) = b_1 I_1(u) + b_0 I_2(u) \quad (3.11)$$

Substituting the input and output signals and their integrals by their approximated values we get:

$$\hat{y}(t) + a_1 \hat{I}_1(y) + a_0 \hat{I}_2(y) = b_1 \hat{I}_1(u) + b_0 \hat{I}_2(u) \quad (3.12)$$

Subtracting equation (3.12) from (3.11):

$$\epsilon_1(y(t)) + a_1 \epsilon_2(I_1(y)) + a_0 \epsilon_3(I_2(y)) = b_1 \epsilon_4(I_1(u)) + b_0 \epsilon_5(I_2(u)) \quad (3.13)$$

$\epsilon$  constitute the error in the approximations.

It is well known that a function which is integrable over an interval can be approximated as:

$$f(t) \approx \sum_{i=0}^N f_i \phi_i(t) \quad (3.14)$$

$$\epsilon_1(f(t)) = f(t) - \sum_{i=0}^N f_i \phi_i(t) \quad (3.15)$$

where  $\epsilon_1$  is the error in approximating the function.  $\epsilon_2, \epsilon_3, \text{etc.}, \dots$  represents the errors in the integrals.

The errors in the approximation exist whether the system is contaminated with noise or noise-free. Equation (3.12) may be presented as follows:

$$\hat{y}(t) + a_1 \hat{I}_1(y) + a_0 \hat{I}_2(y) = b_1 \hat{I}_1(u) + b_0 \hat{I}_2(u) + e(t) \quad (3.16)$$

where  $e(t)$  is the resultant error from all approximation errors.

The errors in the integrations propagate and hence the residuals are correlated. This is better illustrated by the following example.

Consider the third order system of subsection 3.2.1.1 represented by the transfer function  $G(s)$ .

$$G(s) = \frac{1}{(s+1)(s+2)(s+3)}$$

We take 500 input, output data points, sampled at the rate of 0.007 sec. and use the cubic splines as approximating functions. The recursive least squares algorithm utilized in the identification converged to the true parameter values. But if we examine the autocorrelations and partial autocorrelations [see Figure 3.1] we find that the residuals are highly correlated. This is the result of the approximations in the integration.

Let us assume now that the approximation errors are negligible and consider a first order single-input single-output system, with an added measurement error term.

$$\frac{d}{dt} y(t) + a_0 y(t) = b_0 u(t) + n(t) \quad (3.17)$$

$n(t)$  is a white noise sequence.

Integrating (3.17) once we get:

$$y(t) + a_0 \int y(t) = b_0 \int u(t) + \int n(t) \quad (3.18)$$

So even if we assume that there was no error in the approximation while integrating, we do have a new term  $\int n(t)$  which is no longer white noise. It is a Wiener process.

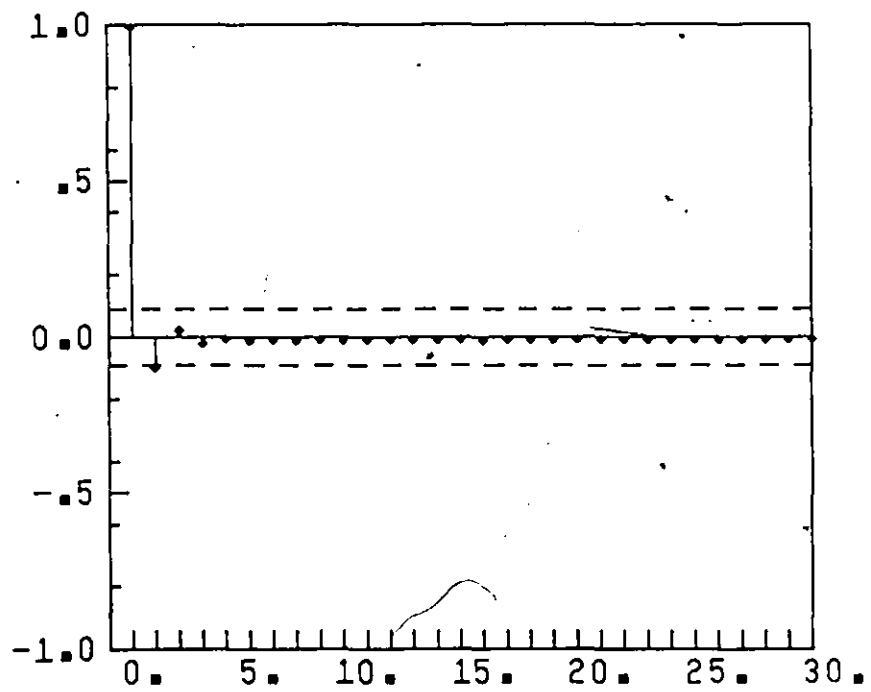
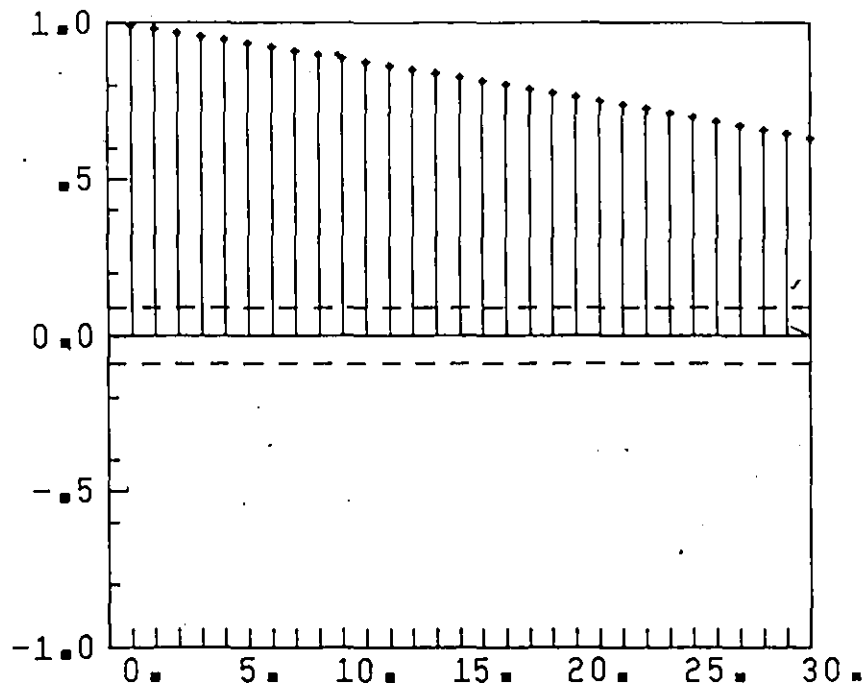


Figure 3.1 Autocorrelations and partial autocorrelations of the error series



The combination of both the errors in the approximation and the integrated noise term give birth to a new error series. When the order of the continuous-time system increases, the number of times the error term is integrated is higher. This results in less accurate estimated parameter values and correlated residuals.

### 3.4 A Proposed Approach for Modelling the Error Term

The presence of this unmodelled error term is responsible for all the problems encountered in identification in presence of noise [25], [35] and [37]. In previous work the error term was considered to represent a white noise sequence and was not accounted for in the identification.

A new approach is now presented to deal with the problems discussed previously. It consists of modelling the error term with the time series method of Box and Jenkins [70]. This is done by examining the autocorrelations and partial autocorrelations of the residuals (the noise series) and fitting a suitable model. The fitted model is checked for adequacy and is adjusted if it fails diagnostic tests.

The modelling of the resultant noise absorbs the errors and gives good estimates of the parameters. Although it is a continuous time system the noise is modelled with a discrete model. This hybrid approach is usually utilized [78] to overcome the problems associated with the identification of continuous noise models.

Because of the existence of the integrated noise term, non-stationary models such as IMA (0,d,q) and ARIMA (p,d,q) [equation (3.19)] are expected to be adequate models for the error term.

$$\phi(B) \nabla^d w_t = \theta(B) a_t \quad (3.19)$$

where  $\phi(B)$  is a stationary autoregressive operator of order  $p$   
 $d$  is the number of roots lying on the unit circle (degree of differencing)  
 $w_t$  is the combined error term  
 $\nabla = (1-B)$  where  $B$  is the backward shift operator  
 $\theta(B)$  is a stationary moving average operator of order  $q$   
 $a_t$  is a sequence of random shocks.

Now, if the model is adequate, it is possible to show [70] that,

$$\hat{a}_t = a_t + O\left(\frac{1}{\sqrt{N}}\right)$$

where  $N$  is the number of  $w$ 's. The autocorrelations  $r_k(\hat{a})$  of the residuals  $\hat{a}$ 's can yield valuable evidence concerning lack of fit and the possible nature of model inadequacy. If more than 5% of the autocorrelations exceed  $\pm \frac{1}{\sqrt{N}}$ , as it will be clearly indicated by a dashed line on each figure, then the model is not adequate.

### 3.4.1 The Noise-Free Case

#### Example 1:

Let us reconsider the same third order system of section 3.3.1. After examining the autocorrelations and partial autocorrelations of the residuals, given in Figure 3.1, several discrete models were tried to model the residuals. The best model found is described by:

$$\nabla^2 e_t = a_t \quad \text{or} \quad (1-B)^2 e_t = a_t \quad (3.20)$$

i.e., the second differencing of the errors in the approximation is a stationary series.

By looking at Figure 3.2 we can see that the residuals are no more correlated and that the model does fit the series.

#### Example 2:

Now let us present another example to demonstrate the effect of the error term on the identification of single-input single-output continuous-time systems. Consider the third order system of section 3.2.1.2 which is a difficult system because the poles are not close to each other. The system was represented by the following transfer function:

$$G(s) = \frac{s^2 + 2s + 5}{(s+0.1)(s+1)(s+10)} = \frac{s^2 + 2s + 5}{s^3 + 11.1s^2 + 11.1s + 1}$$

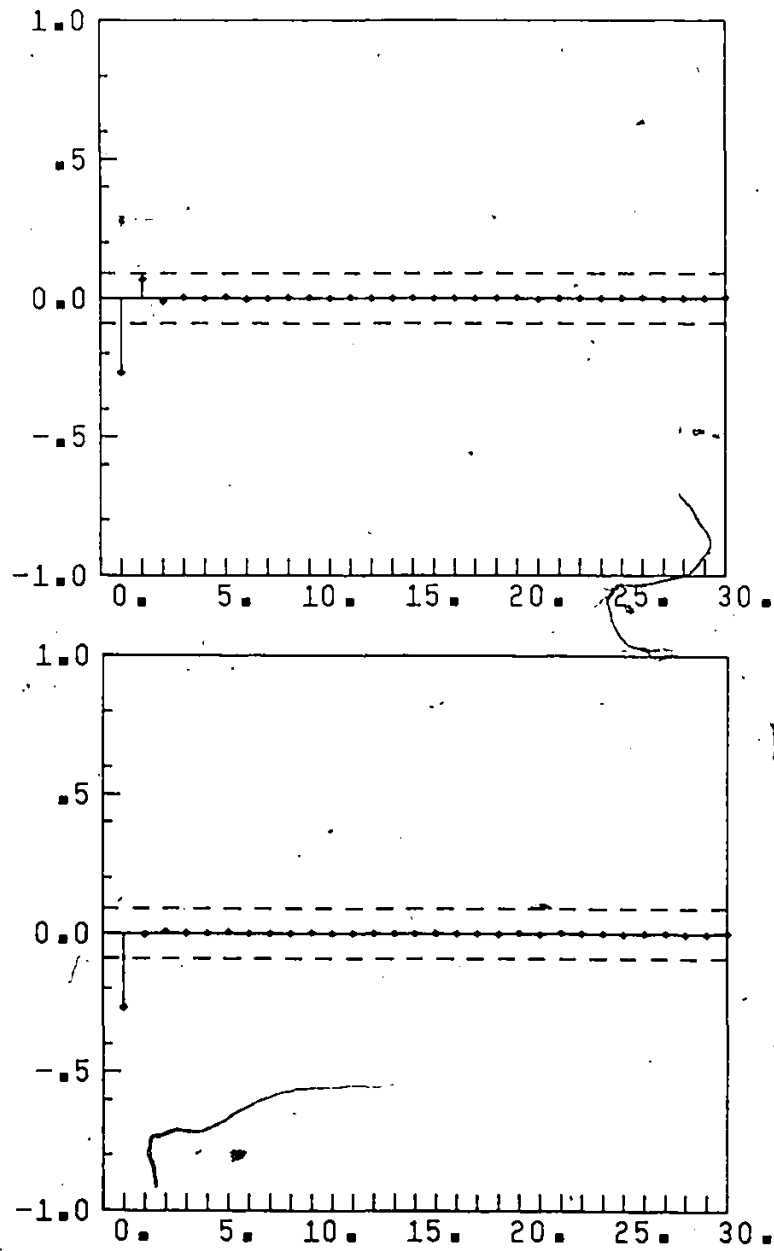


Figure 3.2 Autocorrelations and partial autocorrelations of the residuals after modelling  $e_t$

The accuracy of the estimated parameters and the correlations existing between the residuals are compared before and after using the suggested approach to illustrate the improvement. We take 1000 input-output data sampled at the rate of 0.007 sec. and the cubic splines as approximating functions. The least squares algorithm converged to the true parameter values. By examining the autocorrelations and partial autocorrelations of the residuals [see Figure 3.3] we find that the residuals are correlated.

After extensive simulations the residuals were modelled with the Integrated Moving Average model IMA (0,3,3) which improved the variance of the residuals as presented in Table 3.7, and decorrelated the residuals [see Figure 3.4].

IMA (0,3,3) is described by the following equation:

$$\nabla^3 e_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \theta_3 a_{t-3} \quad (3.21)$$

$$\text{i.e., } (1-B)^3 e_t = a_t + \sum_{i=1}^3 \theta_i a_{t-i} \quad (3.22)$$

The results prove that a model for the errors in the approximation is necessary even when no noise is superimposed on the system.

Table 3.7 Variance of the residuals with and without modelling  $e_t$

Without Modelling $e_t$	With Model IMA (0,3,3)
$0.5805 \times 10^{-11}$	$0.2054 \times 10^{-17}$

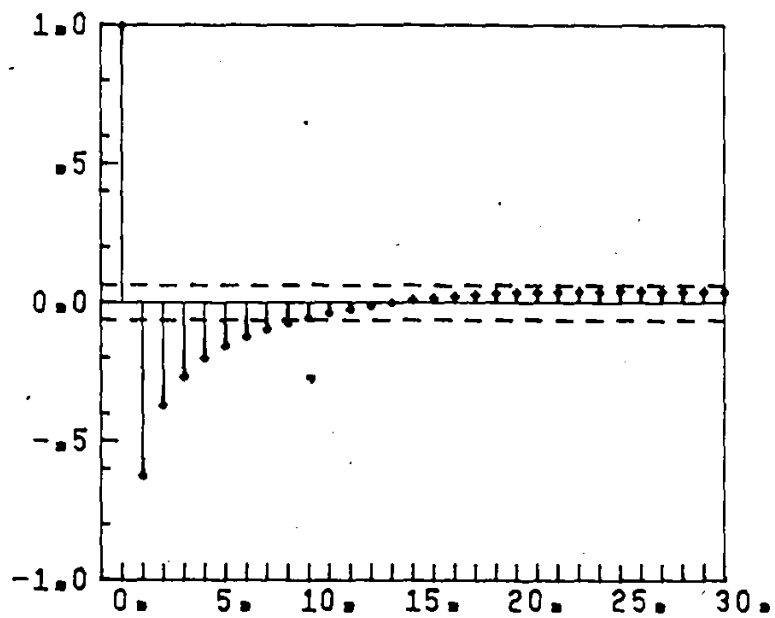
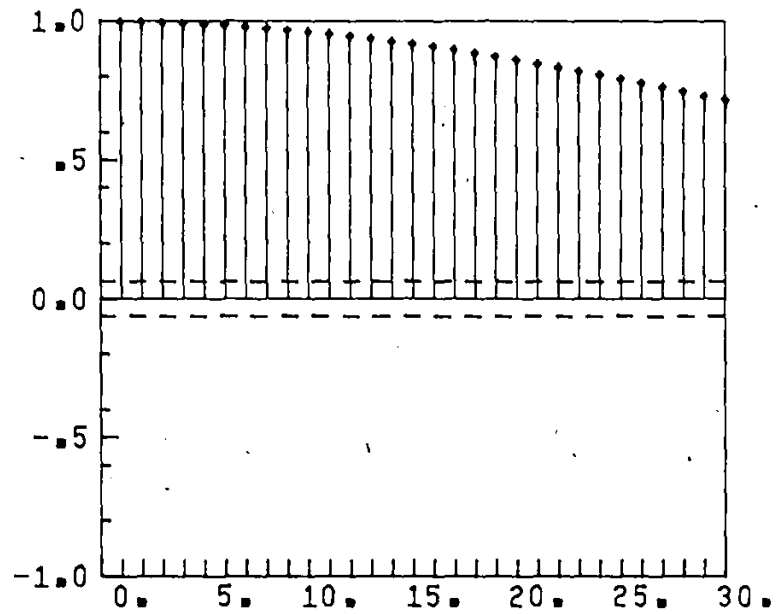


Figure 3.3 Autocorrelations and partial autocorrelations of the error term without modelling. Noise-free case

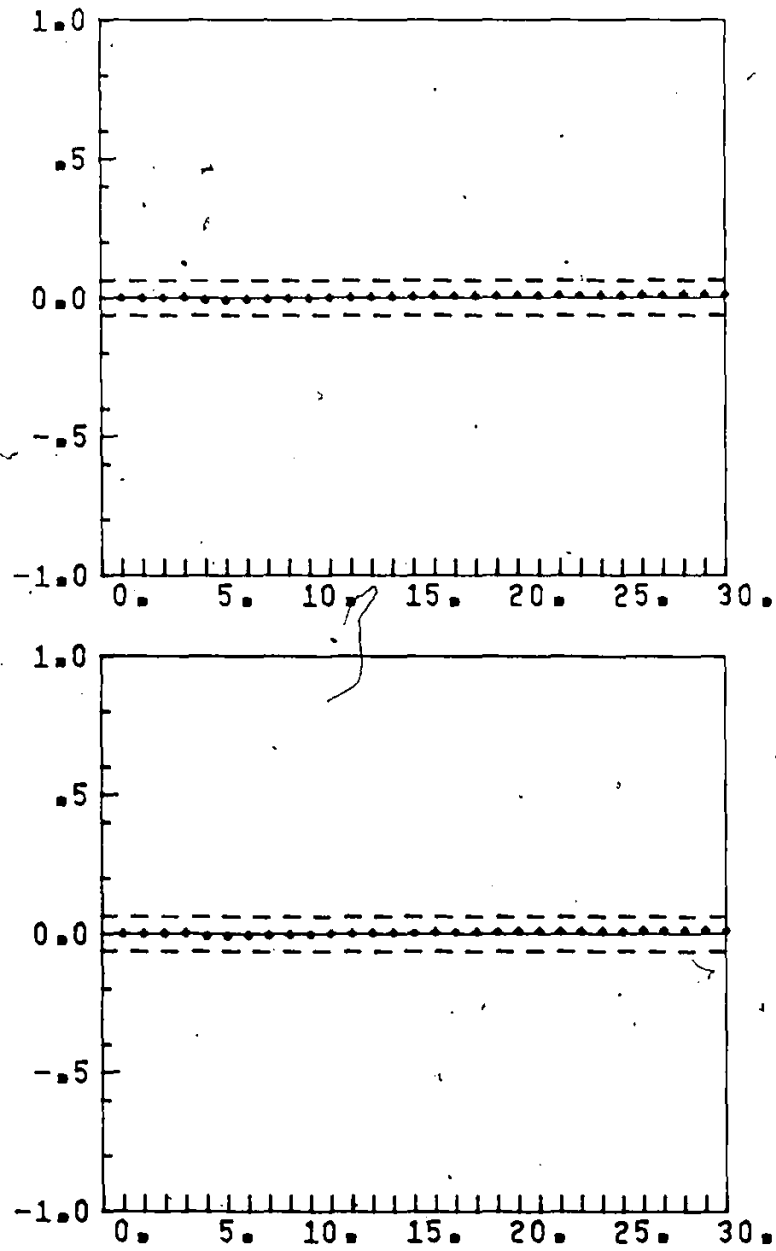


Figure 3.4 Autocorrelations and partial autocorrelations of the residuals with IMA (0,3,3) for the error term. Noise-free case

### 3.4.2. The Noisy Case

Two white noise sequences were generated by the computer and added to the output of the simulated third order system to examine the effect of the combined error term on the identification. The parameter estimates, when no modelling for the error term was used, were presented in section 3.2.1.2.

The noise sequences had zero means and standard deviations equal to ten and twenty percent of the output signal, respectively. When no model for the residuals was used the residuals were correlated. Figure 3.5 shows the autocorrelations and partial autocorrelations of the residuals when the system is contaminated with twenty percent noise level without modelling the residuals.

Several discrete-time models were tried, the best model found to fit the error series  $w_t$  in both noise levels was ARMA(1,1):

$$w_t + \phi_1 w_{t-1} = a_t + \theta_1 a_{t-1} \quad (3.23)$$

In Table 3.8 the variance of the modelled residuals and the corresponding parameter error norm are compared for each noise level with those obtained when no modelling of  $w_t$  was used.

By examining the residuals correlations in Figure 3.6 it becomes clear that the modelling of  $w_t$  did cause the residuals whiteness.



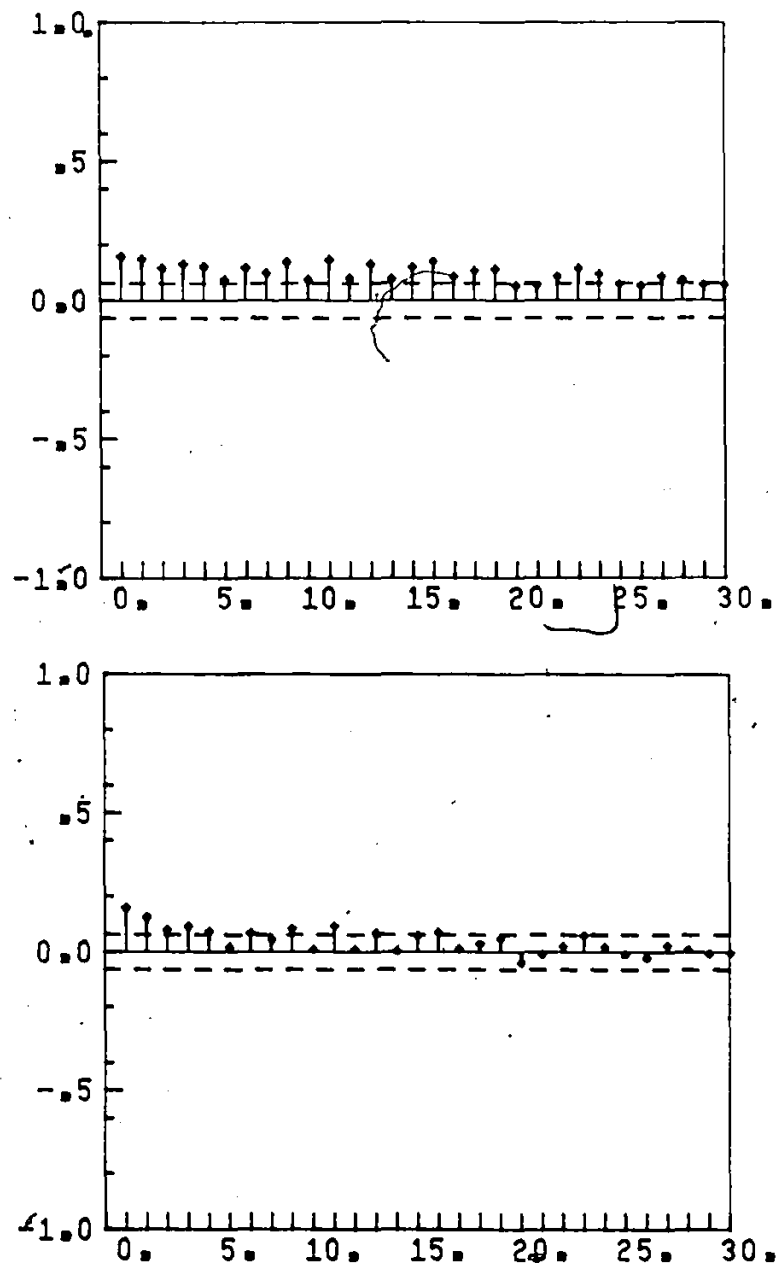


Figure 3.5 The autocorrelations and partial autocorrelations of  $w_t$  without modelling at 20% noise level

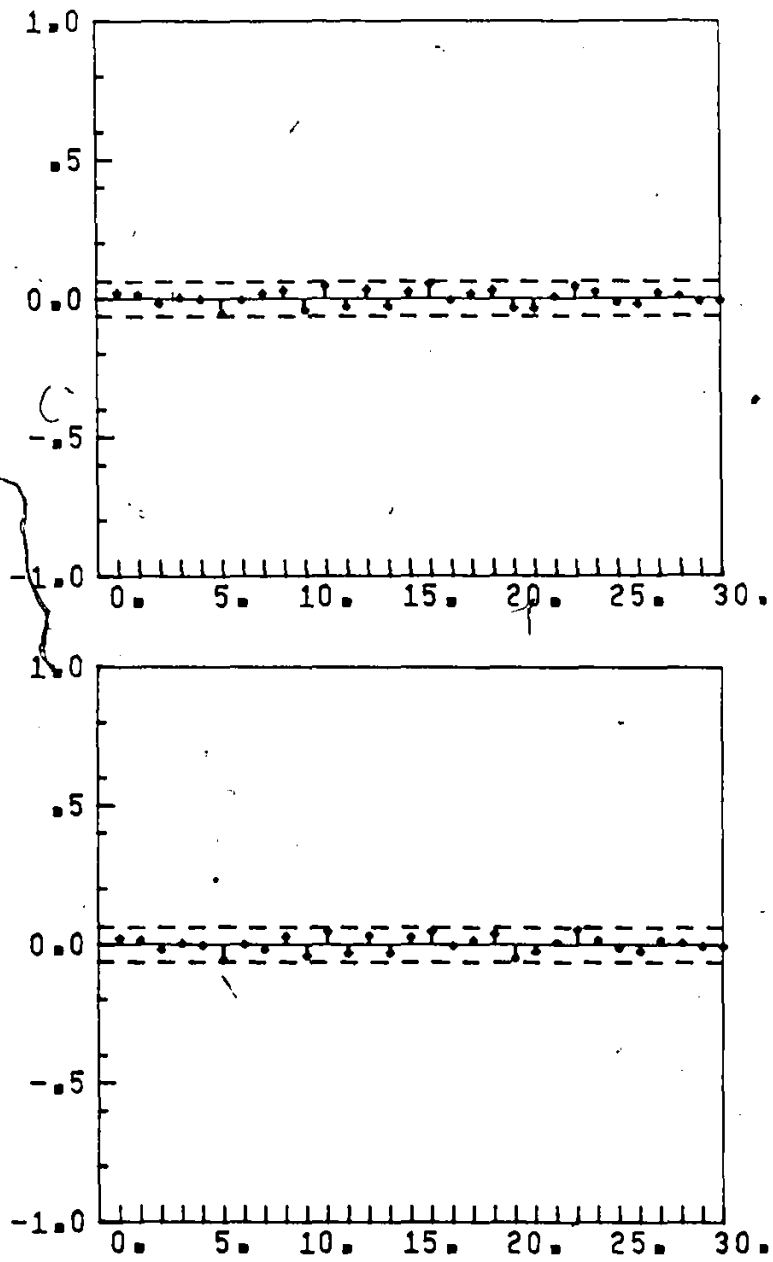


Figure 3.6 The autocorrelations and partial autocorrelations of the residuals when modelling  $w_t$  with ARMA(1,1) at 20% noise level

Table 3.8 Variance of the residuals and parameter error norm for different noise levels with and without modelling  $w_t$

	10% Noise $w_t$ Not Modelled	10% Noise ARMA(1,1)	20% Noise $w_t$ Not Modelled	20% Noise ARMA(1,1)
Variance of the Residuals	$0.9838 \times 10^{-3}$	$0.9083 \times 10^{-3}$	$0.3941 \times 10^{-2}$	$0.3627 \times 10^{-2}$
Parameter Error Norm	0.0686147	0.0635743	0.1501634	0.1390276

If we compare between different sampling intervals we find that the suggested approach does enhance the performance of the least squares algorithm when the sampling interval is not adequate for the system. This is clearly demonstrated in Tables 3.9 and 3.10.

Table 3.9 Third order system with  $T = 0.02$  sec. 10% noise - 1000 input-output data

True Parameters	$T = 0.02$ Without Error Model	$T = 0.02$ With Error Model IMA(0,1,1)
11.1	3.091	9.211
11.1	2.187	8.104
1.	0.221	0.8299
1.	0.4391	0.892
2.	0.7406	1.726
5.	0.9864	3.718
Residuals Variance	0.01164	$0.431 \times 10^{-2}$
Parameter Error Norm	0.764	0.2269

Table 3.10 Third order system with  $T = 0.01$  sec. 10% noise - 1000 input-output data

True Parameters	$T = 0.01$ Without Error Model	$T = 0.01$ With Error Model ARMA(1,1)
11.1	9.514	9.776
11.1	8.846	9.508
1.	0.8431	0.8887
1.	0.9213	0.946
2.	1.794	1.803
5.	3.969	4.246
Variance of Residuals	$0.4775 \times 10^{-2}$	$0.3917 \times 10^{-2}$
Parameter Error Norm	0.177301	0.132952

### 3.5 Concluding Remarks

Differential equations can be converted to linear algebraic models convenient for a least squares solution when using the direct approach. The signals and their successive integrals are approximated and substituted in the differential equations which then can be solved.

Three direct methods using the block pulse functions, the trapezoidal pulse functions and the cubic spline functions as approximating functions are thoroughly compared in the case of single-input single-output systems.

The Comparison showed the superiority of the cubic splines as approximating functions followed by the trapezoidal pulse functions method in the noise-free case. When the observations were contaminated with noise the cubic spline method did not show any superiority for second and third order systems. But probably in the case where the

order of the system will exceed say 6 or 7, the cubic splines will give better results because they are more accurate for high order integrations. All three direct methods gave almost the same parameter estimates. But for practical cases it is advisable to use the trapezoidal pulse functions method. It has the advantage of being easily incorporated in any on-line algorithm. It uses the actual output data unlike the block pulse functions method and hence utilizes the least computational time.

Several illustrative examples were presented throughout this chapter. It has been shown that the choice of the sampling interval affects the accuracy of the estimated parameter. A heuristic rule has been suggested to properly choose a suitable sampling interval. It was also noticed that as the noise level increases the parameter estimates deteriorate even with the sampling interval picked up carefully.

The effect of the error in the approximations and additive noise are then studied. The study showed that the errors in the approximation have a non-stationary behaviour and result in correlated residuals.

It has been also demonstrated that when the continuous-time system is contaminated with white noise, the integrated noise is added to the error term, hence causing identification problems. A new approach is proposed in section 3.4 to overcome these difficulties. This approach consists of modelling the error term which forms the stochastic part of the problem. The modelling has been done with the time series technique.

The proposed approach showed superior results in improving the estimates of the parameters and providing uncorrelated residuals. It also improves the sensitivity of the estimation algorithm to the choice of sampling intervals.

CHAPTER 4  
PARAMETER ESTIMATION OF MULTIVARIABLE  
CONTINUOUS-TIME SYSTEMS

4.1 Introduction

The identification of multivariable systems from the samples of input-output data was extensively treated in the literature for the discrete-time case [22-24]. Little was done in the field of estimating the parameters of a multivariable continuous-time model from the samples of the input-output data. This will be the topic of this chapter and the work of Chapter 3 will be extended to the multivariable case.

In sections 4.2 and 4.5 three direct methods will be used to identify multivariable systems from the available data. Comparisons will be drawn based on the accuracy of the parameter estimates in noise free and noisy cases. The three direct methods are those considered in Chapter 3 for the single-input single-output case.

The identification problems associated with the approximate integration methods are discussed for the multivariable case in section 4.3.

In section 4.4 the error modelling approach is proposed to overcome those problems and to improve parameter estimates of the least squares solution.

#### 4.1.1 Statement of the Problem

Consider an  $n$ th-order linear time-invariant system with  $m$  inputs and  $p$  outputs. The outputs of the system are assumed to be contaminated with additive noise.

The system can be described by the following equations

$$\dot{\underline{x}}(t) = A'\underline{x}(t) + B'\underline{u}(t)$$

$$\underline{z}(t) = C'\underline{x}(t) \tag{4.1}$$

$$\underline{y}(t) = \underline{z}(t) + \underline{w}(t)$$

where  $\underline{x}(t) \in R^n$ ,  $\underline{u}(t) \in R^m$  and  $\underline{z}(t) \in R^p$ . The noise vector  $\underline{w}(t)$  is assumed to be a zero-mean random noise vector of dimension  $p$ .

The problem of system identification may be stated as the determination of the matrices  $A'$ ,  $B'$  and  $C'$  from records of samples of  $\underline{u}(kT)$  and  $\underline{y}(kT)$ , where  $k$  is an integer, and  $T$  is the sampling



interval. For convenience, these sampled observations will be denoted as  $u(k)$  and  $y(k)$ , respectively. It will be assumed that the sampling interval,  $T$ , has been selected carefully. It will also be assumed that the order,  $n$ , of the model is known a priori.

It is well known that the matrices  $A'$ ,  $B'$  and  $C'$  are not unique and for any given input-output description, many such matrices can be obtained through a linear transformation of the state. Alternatively, one may utilize the transfer function matrix description of equation (4.1), which is unique.

$$Y(s) = G(s) U(s) + W(s) \quad (4.2)$$

where

$$G(s) = \begin{bmatrix} g_{11}(s) & g_{12}(s) & \dots & g_{1m}(s) \\ g_{12}(s) & g_{22}(s) & \dots & g_{2m}(s) \\ \vdots & \vdots & \ddots & \vdots \\ g_{p1}(s) & g_{p2}(s) & \dots & g_{pm}(s) \end{bmatrix} \quad (4.3)$$

is the transfer function matrix of the system. If necessary, it is always possible to obtain the state equations from the transfer function matrix. Another advantage of using the transfer function matrix is that one can decompose the multivariable system into  $p$

subsystems, each with one output and  $m$  inputs, corresponding to each row of  $G(s)$ . Hence, each output may be calculated in the following form:

$$Y_i(s) = \sum_{j=1}^m g_{ij}(s) u_j(s) + W_i(s) \quad (4.4)$$

$$i = 1, 2, \dots, p$$

Another commonly used representation of equation (4.1) is the input-output description which can be identified directly from the available input and output samples.

$$(D^n + A_{n-1} D^{n-1} + \dots + A_1 D + A_0) \underline{y}(t) = \quad (4.5)$$

$$(B_{n-1} D^{n-1} + \dots + B_1 D + B_0) \underline{u}(t) + \underline{w}(t)$$

where  $D$  corresponds to the operator  $d/dt$  and  $A$  &  $B$  are matrices of appropriate dimensions.

In the coming sections these two types of models, the transfer matrix representation and the input-output description, will be studied.

#### 4.2 Identification in the Transfer Function Form

The general Transfer Matrix representation given in equation (4.3) can be presented in another form

$$G(s) = \begin{bmatrix} \frac{B_{11}(s)}{D_1(s)} & \frac{B_{12}(s)}{D_1(s)} & \dots & \frac{B_{1m}(s)}{D_1(s)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{B_{p1}(s)}{D_p(s)} & \frac{B_{p2}(s)}{D_p(s)} & \dots & \frac{B_{pm}(s)}{D_p(s)} \end{bmatrix} \quad (4.6)$$

where  $D_i(s)$  is defined as the least common denominator of the  $i$ th row of  $G(s)$  of equation (4.3) having the degree  $n_i$  (less than or equal to  $n$ ) and  $B_{ij}(s)$ 's are polynomials of  $s$  of maximum degree  $n_i - 1$ . This form has been considered for identification by El-Sherief and Sinha [79] and Sinha & Zhou Qijie [25].

The structural parameters for  $G(s)$  of equation (4.6) are the orders  $n_i$ 's of each row of  $G(s)$  and the number of parameters to be estimated is  $\sum_{i=1}^P n_i (n_i + 1)$ . It can be noticed that the form of equation (4.3) for  $G(s)$  is unique and minimal while the form of equation (4.6) is unique but not minimal. The  $i$ th differential equation of the system represented by equation (4.6) can be written as:

$$\begin{aligned}
 (D^{n_1} y_1)_k &= \sum_{j=1}^m \sum_{\ell=1}^{n_1} b_{1j}(n_1-\ell) (D^{(n_1-\ell)} u_j)_k - \\
 &\sum_{\ell=1}^{n_1} d_{1\ell}(n_1-\ell) (D^{(n_1-\ell)} y_1)_k
 \end{aligned} \tag{4.7}$$

where  $b_{1j}(n_1-\ell)$  and  $d_{1\ell}(n_1-\ell)$  are the parameters of the polynomials  $B_{1j}(s)$  and  $D_1(s)$ , respectively, and  $D^{n_1} y_{1k}$  is the  $k^{\text{th}}$  sample of the  $n_1^{\text{th}}$  time derivative of the  $i^{\text{th}}$  output. Using the direct approach, i.e., integrating equation (4.7)  $n_1$  times, (4.7) becomes

$$y_{1k} = \sum_{j=1}^m \sum_{\ell=1}^{n_1} b_{1j}(n_1-\ell) I_{\ell}(u_j)_k - \sum_{\ell=1}^{n_1} d_{1\ell}(n_1-\ell) I_{\ell}(y_1)_k \tag{4.8}$$

where  $y_{1k}$  is the  $k^{\text{th}}$  sample of the  $i^{\text{th}}$  output without any integration

$I_{\ell}(y_1)_k$  is the  $k^{\text{th}}$  sample of the  $\ell^{\text{th}}$  integral of the  $i^{\text{th}}$  output

$I_{\ell}(u_j)_k$  is the  $k^{\text{th}}$  sample of the  $\ell^{\text{th}}$  integral of the  $j^{\text{th}}$  input

Now the given system has been decomposed into  $p$  subsystems where each subsystem corresponds to one row of the TFM (Transfer

Function Matrix) and can be regarded as a single-output multi-input system of order  $n_1$ .

In Chapter 5 three order determination methods will be applied to continuous-time multivariable systems in the TFM form to identify the order of each subsystem. The next subsections will deal with the identification of the parameters  $b_{ij}(n_1-l)$ 's and  $d_i(n_1-l)$ 's of each row with the direct approach for the noise-free and the noisy data cases.

#### 4.2.1 Identification in the Noise Free Case

The three direct approach methods: block pulse functions, trapezoidal pulse functions and the cubic spline method, discussed in Chapter 2 and used in Chapter 3 with single-input single-output systems will now be applied to the multivariable case. The parameter estimates obtained with each method and the corresponding normalised error value will be tabulated and compared.

Consider the following two-input two-output system with the following transfer function matrix:

$$G(s) = \begin{bmatrix} \frac{1}{s^2+3s+2} & \frac{2}{s+1} \\ \frac{3}{s+2} & \frac{2s+1}{s^2+3s+2} \end{bmatrix}$$

The inputs used to excite the system are:

$$u_1(t) = 1.5 \cos 0.9871t + 2.5 \cos 0.2137t - 4 \cos 5.8763t$$

$$u_2(t) = 2 \cos 0.4769t + 2 \cos 3.83t - 4 \cos 2.317t$$

With zero initial conditions. The exact outputs are given by:

$$y_1(t) = 1.63123e^{-2t} - 5.295588e^{-t} + 0.478639 \cos(0.9871t - 1.237807) + 1.21548 \cos(0.2137t - 0.31698) - 0.108106 \cos(5.8763t - 2.64498) + 3.6104 \cos(0.4769t - 0.4449) + 1.0105 \cos(3.83t - 1.3154) - 3.17 \cos(2.317t - 1.1633)$$

$$y_2(t) = 1.1286e^{-t} - 5.8133e^{-2t} + 2.01764 \cos(0.9871t - 0.4584) + 3.7287 \cos(0.2137t - 0.106446) - 1.9332 \cos(5.8763t - 1.2427) + 1.2133 \cos(0.4769t + 0.0826) + 0.9033 \cos(3.83t - 0.9639) - 2.4549 \cos(2.317t - 0.6637)$$

This example was used by Sinha and Zhou [25] to identify the parameters of the transfer function model with the block pulse functions method. In the noisy case they applied only a noise to signal ratio (N.S.R.) of 1.5% to the outputs to get reasonable estimates due to the identification problems discussed earlier in Chapter 3. The parameter estimates obtained by the three direct methods for the first and second subsystems are listed in Tables 4.1 and 4.2. The output and input signals have been sampled at two different sampling rates and the results are included in the same tables.

**Table 4.1 Comparison of three direct methods in the noise-free case. First subsystem**

Sampling Interval	Direct Method	Number of Samples	$b_{11}(0)$	$b_{12}(1)$	$b_{12}(0)$	$d_1(1)$	$d_1(0)$	Parameter Error Norm
0.1	BPF		1.003	2.008	4.013	3.011	2.006	0.00342
	TPF	200	1.004	2.009	4.017	3.014	2.008	0.00435
	cub.sp.		1.000	2.000	4.000	3.000	2.000	0.
0.05	BPF		1.001	2.002	4.002	3.002	2.001	0.00064
	TPF	400	1.001	2.002	4.003	3.003	2.001	0.00084
	cub.sp.		0.9997	2.000	3.999	2.999	1.999	0.00030

**Table 4.2 Comparison of three direct methods in the noise-free case. Second subsystem**

Sampling Interval	Direct Method	Number of Samples	$b_{21}(1)$	$b_{21}(0)$	$b_{22}(1)$	$b_{22}(0)$	$d_2(1)$	$d_2(0)$	Parameter Error Norm
0.1	BPF		3.03	2.998	2.012	0.9974	3.02	1.999	0.0063
	TPF	200	3.031	2.994	2.011	0.9957	3.019	1.996	0.0064
	cub.sp.		2.997	3.002	1.998	0.9999	2.998	2.002	0.0008
0.05	BPF		3.007	3.000	2.003	3.005	3.005	2.000	0.0015
	TPF	400	3.008	2.999	2.003	3.005	3.005	1.999	0.0016
	cub.sp.		3.000	3.000	2.000	3.000	3.000	2.000	0.00

The results of simulation indicated that, as expected, the cubic spline method in the noise free case is superior to the other two direct methods.

#### 4.2.2 Identification in the Presence of Noise

The outputs of the simulated system were contaminated with two different white noise sequences. The sequences had zero means and standard deviations equal to ten percent (10%) and twenty percent (20%) of the output signals, respectively.

In Table 4.3 the three approximate integration methods are compared with respect to the accuracy of the estimated parameters at a sampling interval of 0.05 sec. for the two subsystems.

The results show no difference in performance between the three methods when noise is applied to the multivariable system. The parameter estimates obtained are nearly identical. Also, the estimates are far from good especially when the noise level increases. This is due to the combined effect of the errors in the approximation plus the integrated noise term as discussed previously in Chapter 3 for single-input single-output systems. The analysis is extended to the multivariable case in the next section.

#### 4.3 Effect of Approximate Integration on the Identification in the Multivariable Case

Consider an  $n$ th order multivariable system described in the input-output form by the following differential equation:



Table 4.3 Comparison of three direct methods in the noisy case

Parameters	True Values	BPF	TPF	Cub. Sp.	BPF	TPF	Cub. Sp.
		N.S.R. = 10%			N.S.R. = 20%		
$d_1(1)$	3.	3.134	3.135	3.135	3.245	3.247	3.247
$d_1(0)$	2.	2.075	2.075	2.08	2.136	2.137	2.145
$b_{11}(0)$	1.	1.062	1.062	1.063	1.119	1.119	1.121
$b_{12}(1)$	2.	1.988	1.988	1.981	1.957	1.958	1.944
$b_{12}(0)$	4.	4.146	4.148	4.152	4.262	4.264	4.272
$d_2(1)$	3.	2.603	2.603	2.597	2.253	2.253	2.247
$d_2(0)$	2.	1.529	1.528	1.53	0.9098	0.9099	0.9109
$b_{21}(1)$	3.	2.973	2.974	2.965	3.007	3.007	2.998
$b_{21}(0)$	3.	2.285	2.285	2.287	1.355	1.355	1.356
$b_{22}(1)$	2.	1.821	1.821	1.817	1.647	1.647	1.643
$b_{22}(0)$	1.	0.6737	0.6734	0.6744	0.2893	0.2894	0.29
Parameter Error Norm I		0.038	0.038	0.039	0.069	0.069	0.071
Parameter Error Norm II		0.169	0.169	0.169	0.375	0.375	0.375

$$\frac{d^n}{dt^n} y(t) + A_{n-1} \frac{d^{n-1}}{dt^{n-1}} y(t) + \dots + A_0 y(t)$$

(4.9)

$$= B_{n-1} \frac{d^{n-1}}{dt^{n-1}} u(t) + \dots + B_0 u(t)$$

Integrating (4.9)  $n$  successive times, we get

$$y(t) + A_{n-1} I_1(y(t)) + \dots + A_0 I_n(y(t))$$

(4.10)

$$= B_{n-1} I_1(u(t)) + \dots + B_0 I_n(u(t))$$

Assuming an ideal case with no noise and taking into consideration the existing approximation in the integration equation (4.10) is modified into

$$\hat{y}(t) + A_{n-1} \hat{I}_1(\hat{y}(t)) + \dots + A_0 \hat{I}_n(\hat{y}(t)) =$$

(4.11)

$$B_{n-1} \hat{I}_1(u(t)) + \dots + B_0 \hat{I}_n(u(t)) + \underline{e}(t)$$

where  $\underline{e}(t)$  is the vector containing the errors in the approximation series associated with each output. For the  $i^{\text{th}}$  output equation (4.11) may be written explicitly as

$$\hat{y}_1(t) + \sum_{j=1}^p \sum_{\ell=1}^{n_1} a_{1j}(n_1-\ell) \hat{I}_\ell(y_j(t)) =$$

$$\sum_{j=1}^m \sum_{\ell=1}^{n_1} b_{1j}(n_1-\ell) \hat{I}_\ell(u_j(t)) + e_1(t) \quad (4.12)$$

If we compare equation (4.12) with equation (3.16) it becomes clear that in the multivariable case the errors in the approximation have a greater impact on the identification than in the single-input single-output case. We have additional accumulative errors due to the approximation of the integrals of the different outputs and inputs. In the transfer function form, the multi-input multi-output system is decomposed into  $p$  different systems with one output and  $m$  inputs. Hence the approximate integration of the other noisy outputs to the system does not influence the parameter estimation process contrary to the input-output form representation of equal order.

Let us add a measurement error term to equation (4.10), and assume that the errors in the approximation are negligible. Integrating (4.10)  $n$  successive times we get

$$y(t) + A_{n-1} I_1(y(t)) + \dots + A_0 I_n(y(t)) =$$

$$B_{n-1} I_1(u(t)) + \dots + B_0 I_n(u(t)) + I_n(n(t)) \quad (4.13)$$

where  $I_n(\underline{n}(t))$  represents the  $n$  fold time integral of the noise sequence associated with each output.

So if  $\underline{n}(t)$  are independent white noise sequences  $I_n(\underline{n}(t))$  is not guaranteed to be so. Since the errors in the approximation are not negligible as assumed, the new error term

$$\underline{w}(t) = \underline{e}(t) + I_n(\underline{n}(t)) \quad (4.14)$$

will be a vector of coloured noise sequences.

Hence we reach the same conclusion as in section 3.3. The error term composed of the errors generated by the approximate integration in addition to the integrated noise term are responsible for the non-satisfactory obtained parameter estimates as well as the correlated residuals.

#### 4.4 The Error Modelling Approach

To solve the problems associated with the identification of multivariable continuous-time systems, the error modelling approach presented in section 3.4 is proposed. The modelling of the coloured noise sequences is believed to provide a considerable improvement in the parameter estimates and gives accordingly uncorrelated residual series. In previous work, no modelling was used [37], [34] and the errors in the approximation problem never addressed, but it was noticed that the estimates deteriorate as the noise level increases [35]. The

time series approach of Box and Jenkins [70] presents a very good analytical approach which makes use of all the available information to identify a suitable model for the error series.

The error modelling approach recognizes the existence of a stochastic error series and shows the importance of modelling it. Once a model is identified for the error series, the parameters of both the multivariable system model and the noise model can be estimated at the same time recursively. This transforms the least squares algorithm into an approximate maximum likelihood method [78].

#### 4.4.1 Results of Simulation

In this section the error modelling approach will be applied to the system described in section 4.2. The plot of the autocorrelations and partial autocorrelations of the error series when the system is contaminated with 20% N.S.R. without modelling is given in Figure [4.1-2]. It is evident that the residuals are correlated. A second diagnostic test called the portmanteau criterion also shows an inflated  $Q$ , when compared with the  $\chi^2$  tables [70].

$$Q = n \sum_{k=1}^K r_k^2(\hat{a}) \quad (4.15)$$

where  $n = N-d$  is the number of  $w$ 's used to fit the model,  $d$  is the number of differencing and  $N$  is the total number of observations.  $r_k$  is the estimated autocorrelation at lag  $k$ . For  $n = 400$ ,  $d = 0$ ,  $K = 30$ , the corresponding values of  $Q$  are given in Table 4.4.

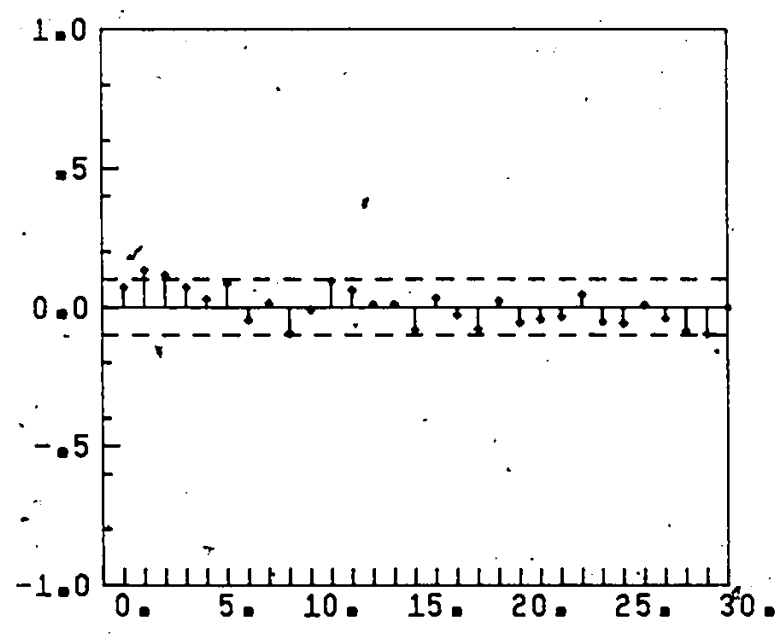
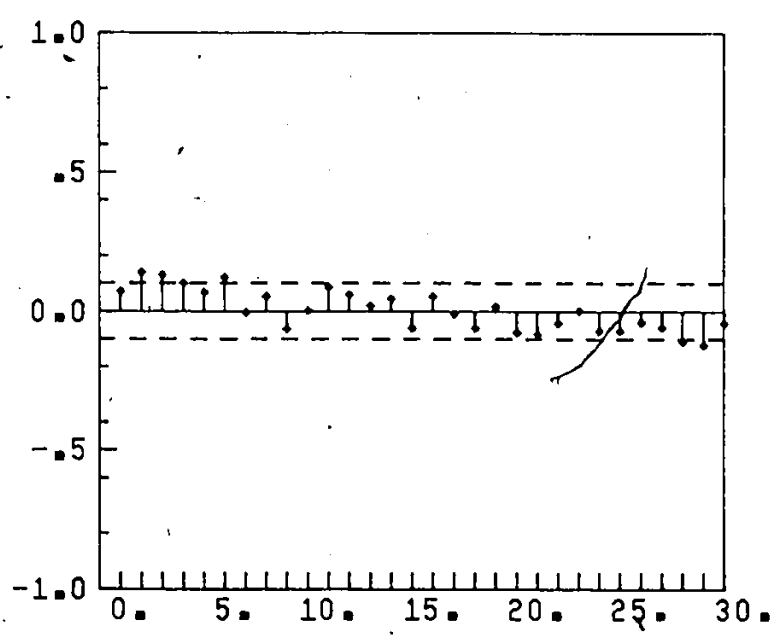


Figure 4.1 Autocorrelations and partial autocorrelations of the error series without modelling at 20% noise level. First subsystem

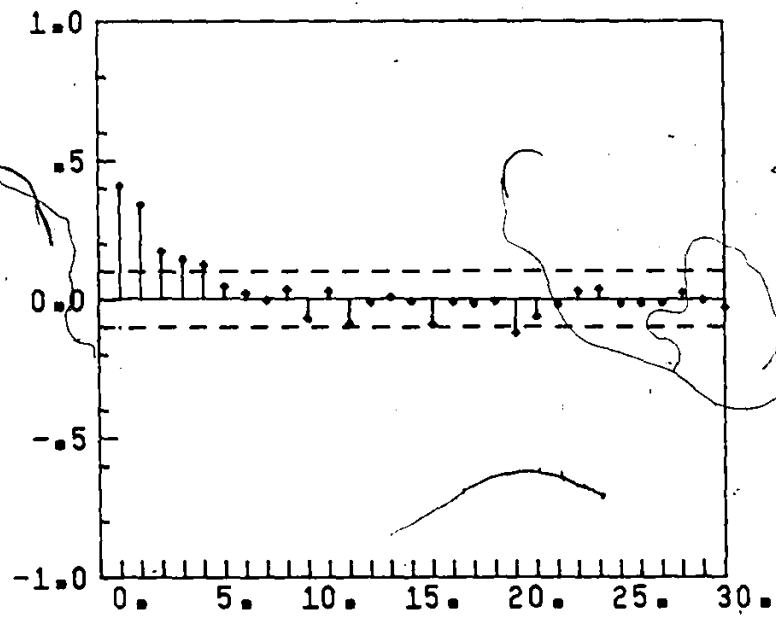
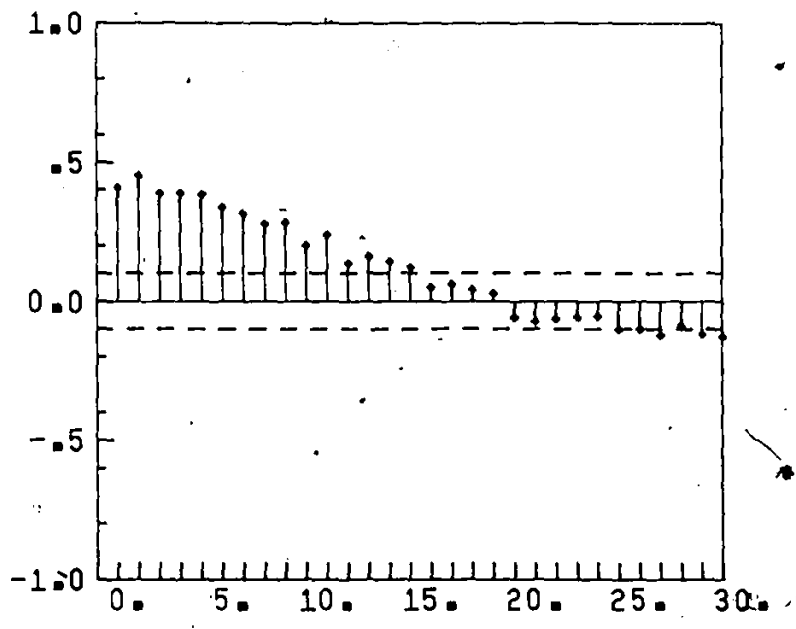


Figure 4.2 Autocorrelations and partial autocorrelations of the error series without modelling at 20% level. Second subsystem

Table 4.4 The values of Q when  $w_t$  is not modelled

	N.S.R. = 10%	N.S.R. = 20%
1st Subsystem	58.94	66.35
* 2nd Subsystem	1031.53	597.83

The time-series model identified for the error term of the first subsystem is an integrated moving average model of order one, IMA(0,1,1), equation (4.16), and the error series of the second subsystem was modelled by an autoregressive integrated moving average model, ARIMA (1,1,1), given in equation (4.17).

$$\nabla w_t = a_t + \theta_1 a_{t-1} \quad (4.16)$$

$$\nabla w_t - \phi_1 \nabla w_{t-1} = a_t + \theta_1 a_{t-1} \quad (4.17)$$

In Table 4.5 the parameter estimates obtained, before and after the modelling of the error series, are listed to show the improvement introduced by the error modelling approach.

Table 4.6 gives the values of Q after the modelling and Figures 4.3 - 4.4 show the estimated autocorrelations and partial autocorrelations of the residuals at 20% noise to signal ratio.

The autocorrelations of the residuals as well as the values of Q gives us a clear indication that the identified models do fit the



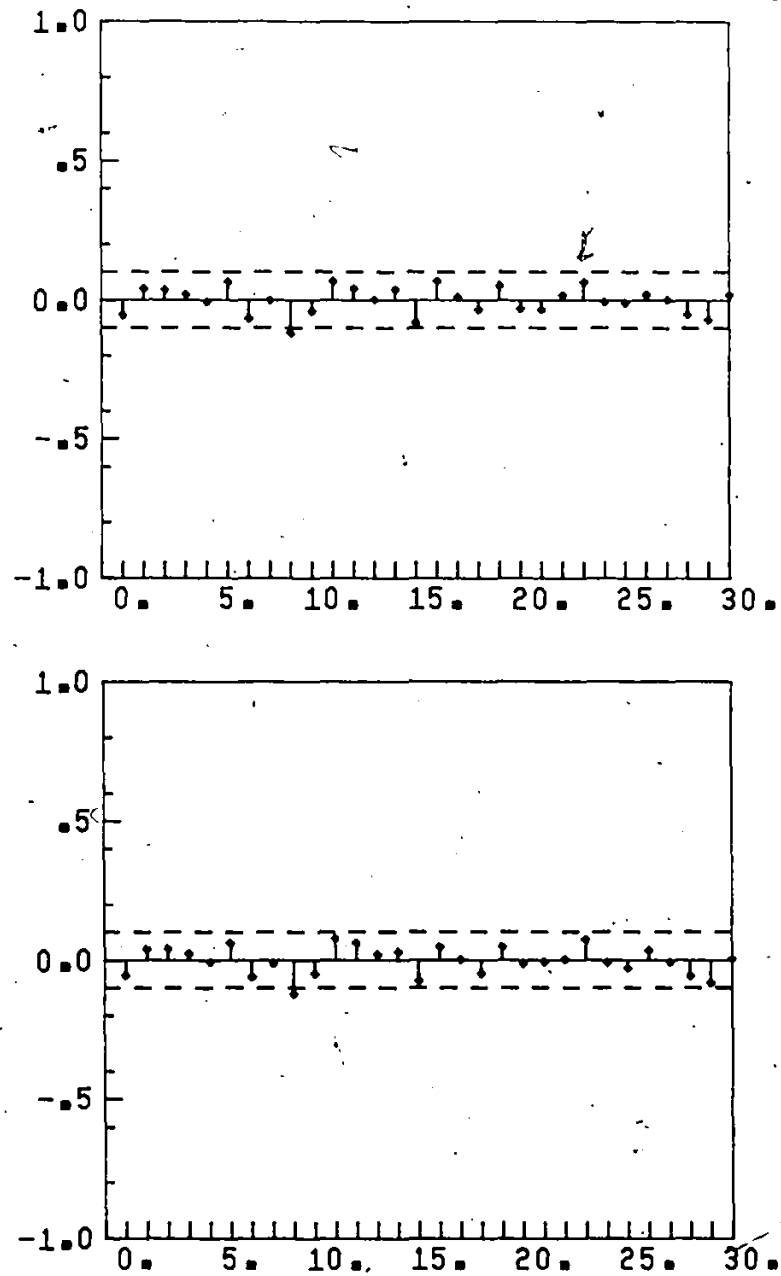


Figure 4.3 Autocorrelations and partial autocorrelations of the error series after modelling (20% noise). First subsystem

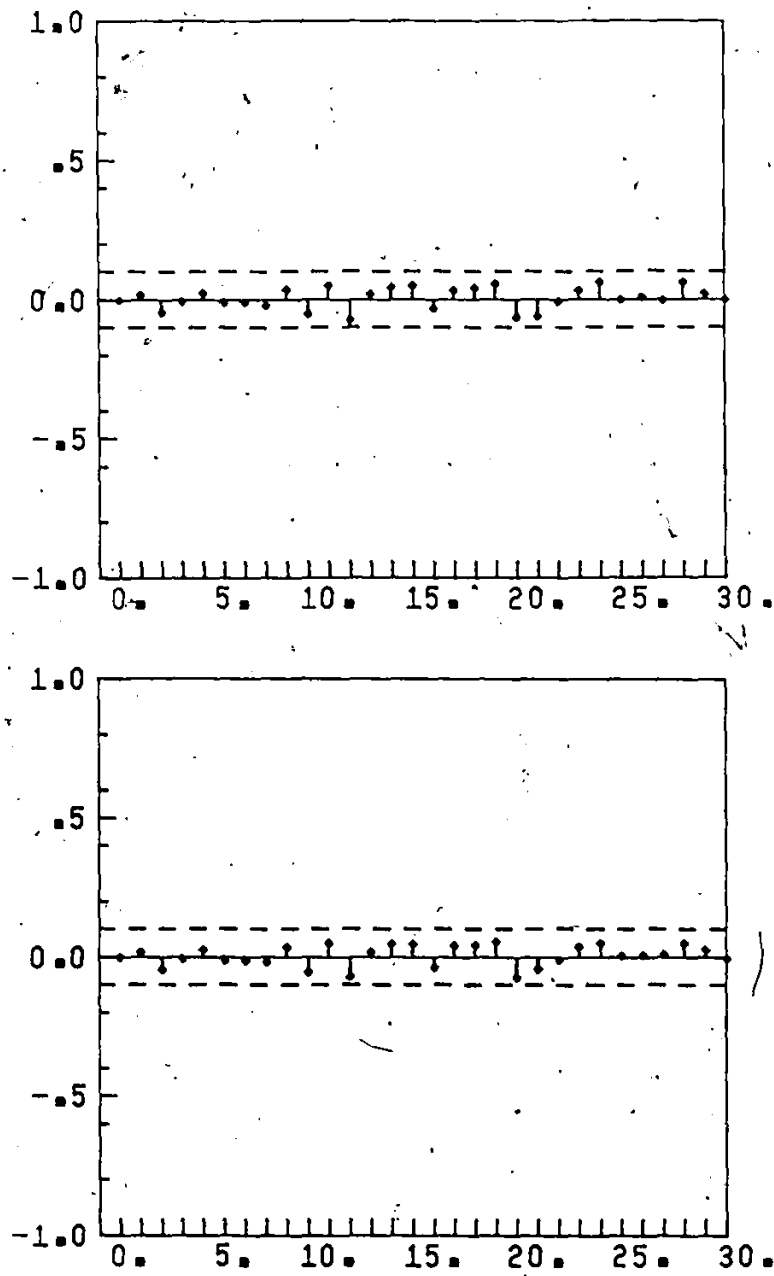


Figure 4.4 Autocorrelations and partial autocorrelations of the error series after modelling (20% noise). Second subsystem

Table 4.5 Results of simulation of the given example after 400 iterations at  
 $T = 0.05$  sec

Parameters	True Values	$w_t$ Not Modelled	Proposed Approach	$w_t$ Not Modelled	Proposed Approach
		N.S.R. = 10%		N.S.R. = 20%	
$d_1(1)$	3.	3.135	3.083	3.247	3.009
$d_1(0)$	2.	2.075	2.035	2.137	1.933
$b_{11}(0)$	1.	1.062	1.044	1.119	1.018
$b_{12}(1)$	2.	1.988	2.016	1.958	2.02
$b_{12}(0)$	4.	4.148	4.077	4.264	3.875
$d_2(1)$	3.	2.603	2.819	2.253	2.555
$d_2(0)$	2.	1.528	1.842	0.9099	1.461
$b_{21}(1)$	3.	2.974	2.996	3.007	2.992
$b_{21}(0)$	3.	2.285	2.768	1.355	2.197
$b_{22}(1)$	2.	1.821	1.977	1.647	1.92
$b_{22}(0)$	1.	0.6734	0.8489	0.2894	0.5673
Parameter Error Norm I		0.038	0.02185	0.069	0.0248
Parameter Error Norm II		0.169	0.06122	0.375	0.117

error series. It is also evident by examining Table 4.5 that the modelling of  $w_t$  improves the parameter estimates.

Table 4.6 Estimates of Q when  $w_t$  is modelled

	N.S.R. = 10%	N.S.R. = 20%
1st Subsystem IMA (0,1,1)	28.65	28.35
2nd Subsystem ARIMA (1,1,1)	19.7	18.41

#### 4.5 Identification in the Input-Output Form

Another popular form in which we can identify a multivariable system is the input-output form. This form plays an important role in system identification since the input-output models are linking directly the input and output observations. Guidorzi [24] considered for identification a canonical input-output difference equation representation for the discrete-time case which can be generalized to the continuous-time case as follows:

$$P(D) y(t) = Q(D) u(t), \quad (4.18)$$

$$\text{with } P(D) = \begin{bmatrix} P_{11}(D) & \dots & P_{1p}(D) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ P_{p1}(D) & & P_{pp}(D) \end{bmatrix} \quad (4.19)$$

$$Q(D) = \begin{bmatrix} q_{11}(D) & \dots & q_{1m}(D) \\ \vdots & & \vdots \\ q_{p1}(D) & \dots & q_{pm}(D) \end{bmatrix} \quad (4.20)$$

where  $p_{ij}(D)$  and  $q_{ij}(D)$  are polynomials in  $D (= \frac{d}{dt})$  of the following form

$$p_{ii}(D) = D^{n_i} - a_{ii}(n_i-1) D^{n_i-1} - \dots - a_{ii}(0) \quad (4.21)$$

$$p_{ij}(D) = -a_{ij}(n_{ij}-1) D^{n_{ij}-1} - \dots - a_{ij}(0) \quad (4.22)$$

$$q_{ij}(D) = b_{ij}(n_i-1) D^{n_i-1} + \dots + b_{ij}(0) \quad (4.23)$$

and  $n_i$ 's are the observability subindices of the system [24]. The canonical form of equations (4.19) - (4.23) has been also considered for identification by El-Sherief and Sinha [81]. Guidorzi also established [24] a relationship between the state space representation in a certain canonical form and the input-output representation of equations (4.19) - (4.23). This made it attractive for identification purposes and made the transformation to the state space form an easy task.

The canonical form described by Guidorzi will be used in the identification of continuous-time multivariable systems in this section and throughout Chapter 6.

#### 4.5.1 Results of Simulation

For the sake of completing our comparison of the three direct approach methods, the block pulse functions, the trapezoidal pulse functions and the cubic spline method, their performance will be also investigated when the identification is performed in the input-output form.

Consider the following two-input two-output system

$$D^2 \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} D \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad (4.24)$$

The outputs of the system were calculated for the following inputs

$$u_1(t) = \sin 0.5t + \sin t + \sin 1.5t$$

$$u_2(t) = \sin 1.2t + \sin 2.5t$$

Assuming zero initial conditions the exact output is given by:

$$\begin{aligned} y_1(t) = & -0.2677 \cos 0.5t + 0.773 \sin 0.5t - 0.6408 \cos t + 0.6248 \\ & \sin t - 0.8197 \cos 1.5t + 0.10587 \sin 1.5t + 1.7282 \\ & \cos 1.3856te^{-0.8t} + 0.15329 \sin 1.3856te^{-0.8t} \end{aligned}$$

$$\begin{aligned}
 y_2(t) = & 0.1222 \cos 0.5t - 0.1474 \sin 0.5t + 0.228 \cos t + 0.02375 \\
 & \sin t + 0.0232 \cos 1.5t + 0.18238 \sin 1.5t - 0.87775 \cos 1.2t \\
 & + 0.39458 \sin 1.2t - 0.23698 \cos 2.5t - 0.252788 \sin 2.5t - \\
 & 1.86781 \cos 1.299te^{-0.75t} - 1.55427 \sin 1.299te^{-0.75t} + 2.609 \\
 & \cos 1.3856te^{-0.8t} + 1.90567e^{-0.8t} \sin 1.3856t
 \end{aligned}$$

Table 4.7 gives the values of the parameters estimated with the block pulse functions, trapezoidal pulse functions and cubic spline methods. The results are obtained with 200 samples using the recursive least squares method with no modelling of the error sequences at a sampling interval of 0.1 sec. The noise to signal ratio utilized in the comparison is varied from 10% to 20% by adjusting the variance of the random sequences.

#### 4.5.2 Parameter Estimation with the Error Modelling Approach

By examining Table 4.7 we see that the results are not satisfactory. The errors in the parameters are large, and if we plot the autocorrelations and partial autocorrelations of the residual series as in Figures 4.5 and 4.6 (N.S.R. = 20%), it becomes evident that the residuals fail the whiteness test. The second diagnostic test, (Table 4.8), also shows that the estimated values of  $\theta$  are very large.

The error modelling approach will be now applied to the multi-variable system.

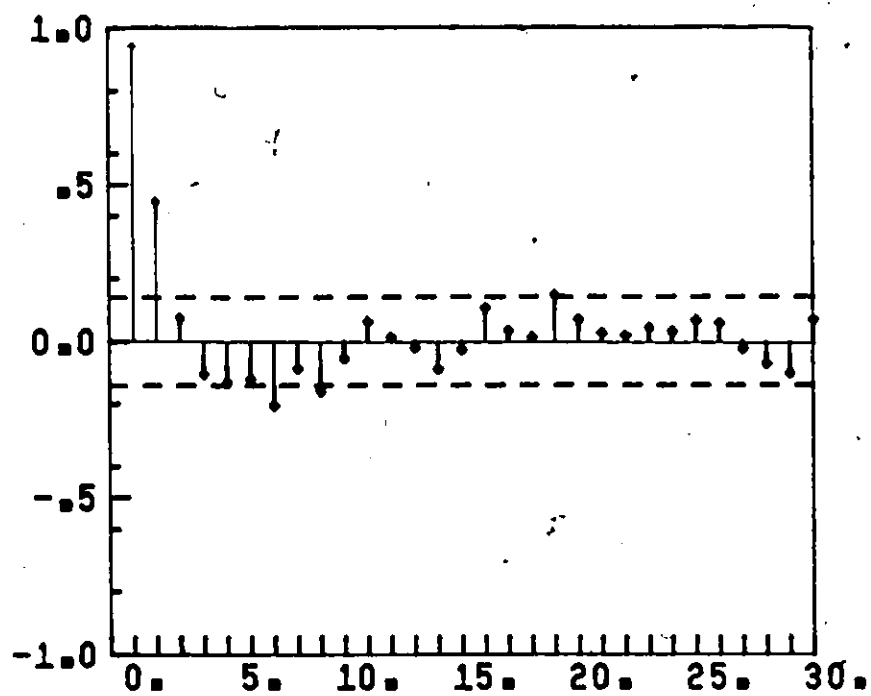
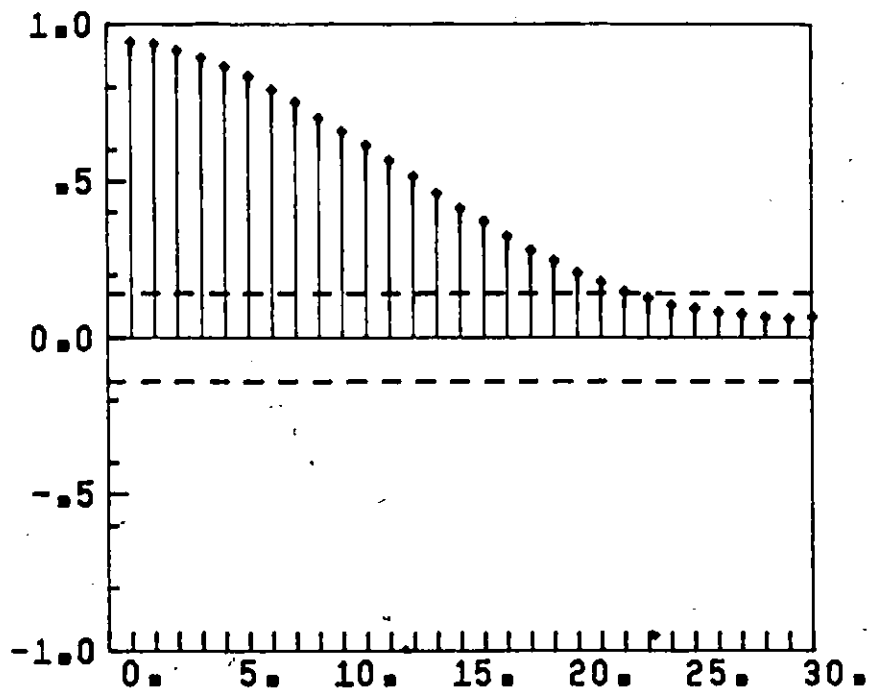


Figure 4.5 Autocorrelations and partial autocorrelations of the error series without modelling at 20% noise level. First subsystem



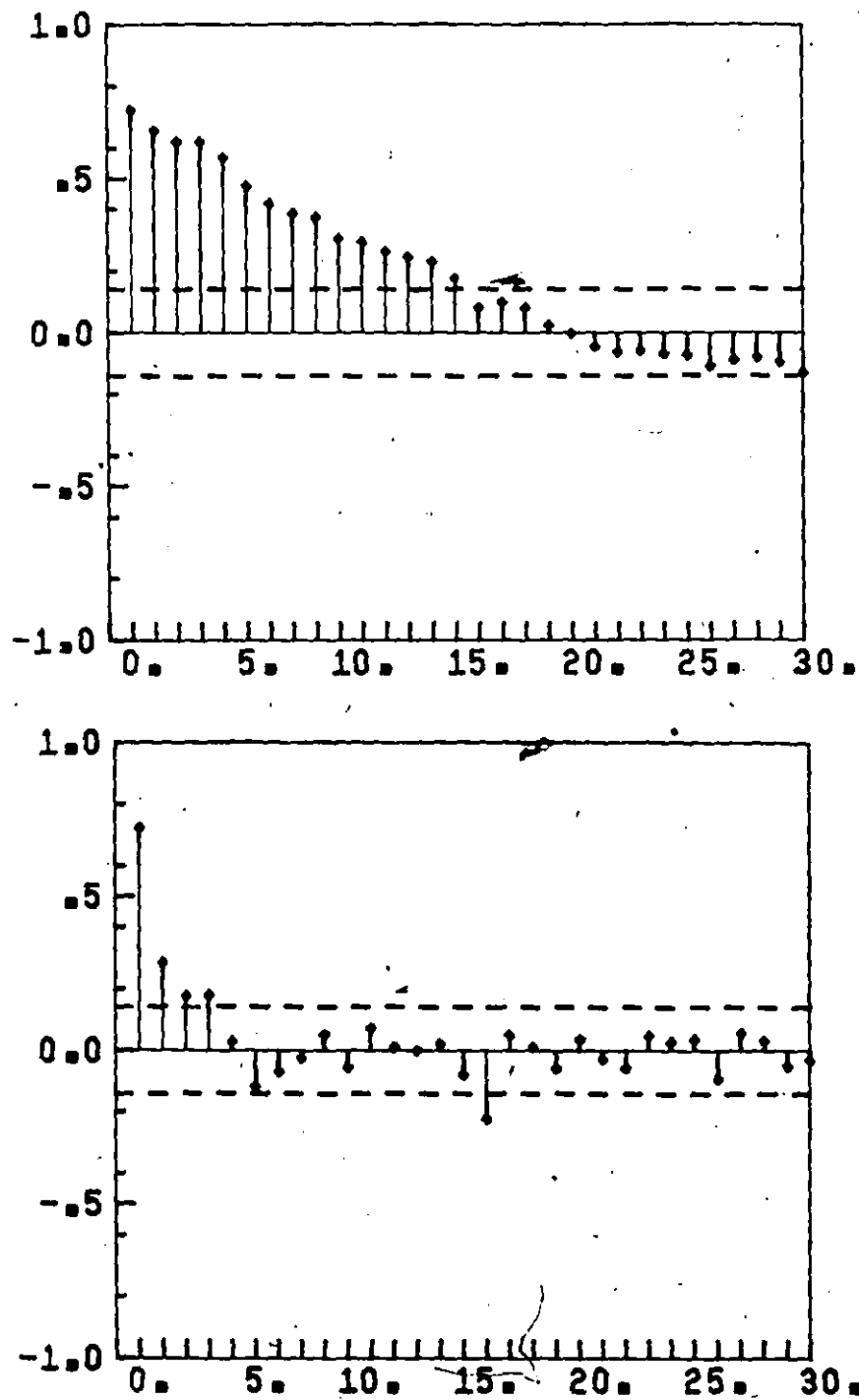


Figure 4.6 Autocorrelations and partial autocorrelations of the error series without modelling at 20% noise level. Second subsystem

Table 4.7 Comparison of three direct methods in the noisy case

Parameters	True Values	BPF	TPF	Cubic. Sp.	BPF	TPF	Cubic. Sp.
		N.S.R. = 10%			N.S.R. = 20%		
$a_{11}(1)$	1.6	1.473	1.473	1.469	0.7164	0.7173	0.7257
$a_{11}(0)$	2.56	2.158	2.157	2.15	1.018	1.019	1.029
$b_{11}(0)$	2.	1.711	1.711	1.702	0.8174	0.8181	0.8238
$a_{21}(0)$	1.9	1.567	1.565	1.555	1.195	1.193	1.187
$a_{22}(1)$	1.5	1.155	1.155	1.15	0.8045	0.8043	0.8015
$a_{22}(0)$	2.25	1.981	1.979	1.974	1.642	1.64	1.639
$b_{22}(0)$	0.5	0.3657	0.3651	0.3641	0.2368	0.2365	0.2363
Parameter Error Norm I		0.141	0.141	0.144	0.589	0.589	0.585
Parameter Error Norm II		0.169	0.171	0.173	0.356	0.357	0.358

Table 4.8 Estimated values of Q when  $w_t$  is not modelled

	N.S.R. = 10%	N.S.R. = 20%
1st Subsystem	2706.12	1716.93
2nd Subsystem	846.11	660.95

Several discrete models were investigated for the error series. The models that did pass the diagnostic tests were ARIMA (1,1,1) for the error term of the first subsystem and IMA (0,2,2) for the error term of the second subsystem. The Autoregressive Integrated Moving average model ARIMA (1,1,1) was presented in equation (4.16) and the Integrated Moving Average model IMA (0,2,2) can be written as:

$$\sqrt{2}w_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \quad (4.25)$$

Table 4.9 gives the estimated values of Q after the modelling, and Figures 4.7 and 4.8 show the estimated autocorrelations and partial autocorrelations of the residuals at 20% noise to signal ratio.

Table 4.9 Estimates of Q when  $w_t$  is modelled

	N.S.R. = 10%	N.S.R. = 20%
1st Subsystem ARIMA (1,1,1)	27.11	27.11
2nd Subsystem IMA (0,2,2)	27.83	28.06

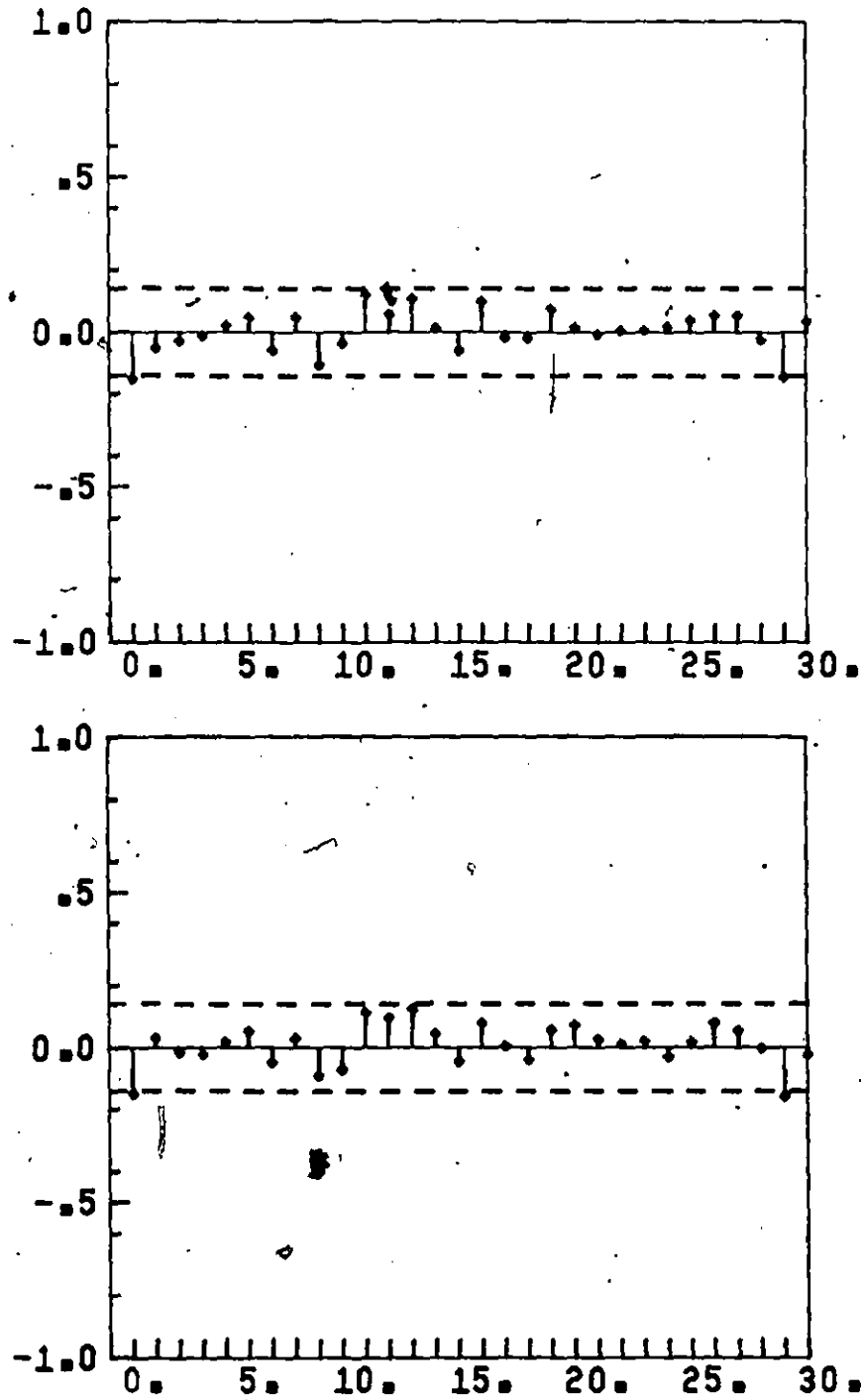


Figure 7.7 Autocorrelations and partial autocorrelations of the error series after modelling at 20% noise level. First subsystem

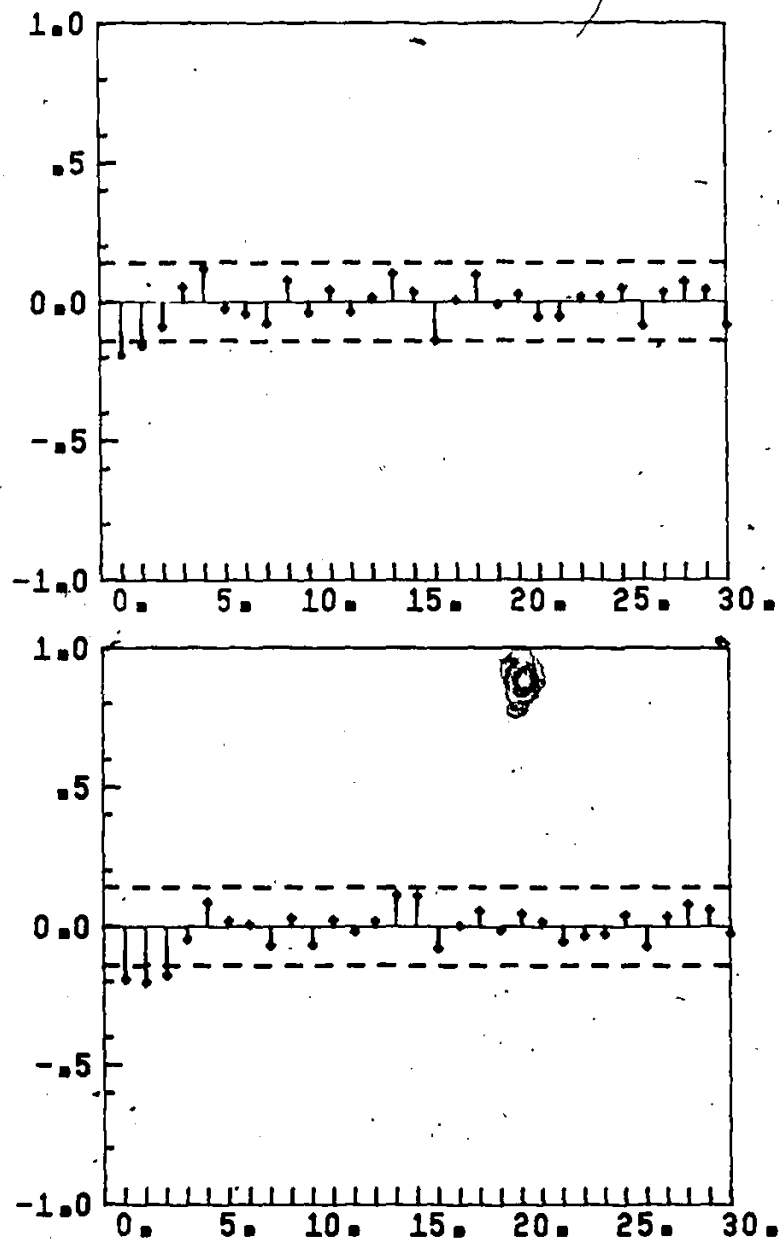


Figure 4.8 Autocorrelations and partial autocorrelations of the error series after modelling at 20% noise level. Second subsystem

Table 4.10 Results of simulations of the given example after 200 iterations at  $T = 0.1$  sec

Parameters	True Values	$w_t$ Not Modelled	Proposed Approach	$w_t$ Not Modelled	Proposed Approach
		N.S.R. = 10%		N.S.R. = 20%	
$a_{11}(1)$	1.6	1.473	1.603	0.7173	1.579
$a_{11}(0)$	2.56	2.157	2.543	1.019	2.494
$b_{11}(0)$	2.	1.711	1.965	0.8181	1.902
$a_{21}(0)$	1.9	1.565	1.852	1.193	1.786
$a_{22}(1)$	1.5	1.155	1.421	0.8043	1.377
$a_{22}(0)$	2.25	1.979	2.189	1.64	2.151
$b_{22}(0)$	0.5	0.3651	0.4511	0.2365	0.4491
Parameter Error Norm I		0.141	0.0107	0.589	0.033
Parameter Error Norm II		0.171	0.0362	0.357	0.06

In Table 4.10 the parameter estimates are listed, before and after the modelling of the error series.

#### 4.6 Concluding Remarks

In this chapter three different direct methods have been considered for identification of continuous-time multivariable systems. The comparison in noise-free and noisy environments in both the transfer function and the input-output representations showed that the cubic spline interpolation method is only superior in the noise free case. No distinct difference is noticeable when the level of noise superimposed on the system increases.

On the other hand, it has been noticed that when the noise is added to the multivariable system, identification problems arise such as inaccurate parameter estimates and correlated residuals. The analysis of this problem has been extended to the multivariable case. In section 4.4 an approach is proposed which recognizes the existence of a stochastic error series and points out the importance of modelling it.

As we can see from Tables 4.6 and 4.9, better estimates of the parameters were obtained after using the error modelling approach for different noise levels. The least squares algorithm is transformed into an approximate maximum likelihood method estimating at the same time the model of the system and the model of the corresponding error series.

## CHAPTER 5

### ORDER DETERMINATION OF MULTIVARIABLE

#### CONTINUOUS-TIME SYSTEMS

##### 5.1 Introduction

The choice of model structure is one of the basic factors in the formulation of the identification problem [23]. The choice will reflect on the computational effort, the number of parameters to be estimated, etc...

Order determination tests, in general, can be classified in two categories. First, there are the methods based upon testing the rank of the product moment matrix (PMM). They are fast and can serve as a rough determination of a range of orders. A major disadvantage of order tests based upon the (PMM) is that the rank condition is masked in the case of noisy data. This problem can be alleviated by enhancing the PMM [62], i.e., assume that a vector of disturbances is superimposed to the input and output signals and that the covariance matrix of the disturbances is known. However, enhancement involves additional computations, the instrumental PMM [63] is then used in place of the normal product matrix. The IPMM assumes the input to be noise-free and the measured output to be contaminated with disturbances. The rank condition test was exploited and used by many authors, notably Chow



[64], Tse and Weinert [65] and Guidorzi [24] who have applied the idea to autoregressive, moving average processes and to the structure determination of multivariable systems.

The second class of tests is based on examining the output signals for minimum prediction error. Akaike [66] proposed the final prediction error [FPE] test and is calculated from

$$\text{FPE} = \hat{\sigma}^2 \frac{(N + p + 1)}{(N - p - 1)} \quad (5.1)$$

where  $N$  is the number of data points,  $p$  is the number of parameters, and  $\hat{\sigma}^2$  is the variance of the residuals from the model estimation. A similar measure is Akaike's information criterion (AIC) [67], defined by:

$$\text{AIC} = (-2) \log_e (\text{maximum likelihood}) + 2 (\text{number of free parameters})$$

$$\text{or} \quad \text{AIC} = N \log_e \hat{\sigma}^2 + 2p \quad (5.2)$$

Another test based on the analysis of the residuals is the loss function test

$$v_1(\hat{n}_1) = \frac{1}{N - \hat{n}_1 - 1} \sum_{k=\hat{n}_1+1}^N [e_1^{(\hat{n}_1)}(k)]^2 \quad (5.3)$$

and the model with the minimum squared error has the true order  $n_1$ .

Diekman and Unbehauen [68] compared the PMM based tests, the loss function and the FPE tests against the polynomial test and said that the latest is the most accurate in determining the order of a multivariable system. The polynomial test does not examine the output signal but the parameters of the estimated transfer function. For each transfer function the poles and zeros will be calculated and plotted in the unit circle of the  $z$  domain. The most probable order of the system is equal to the number of poles not compensated by zeros.

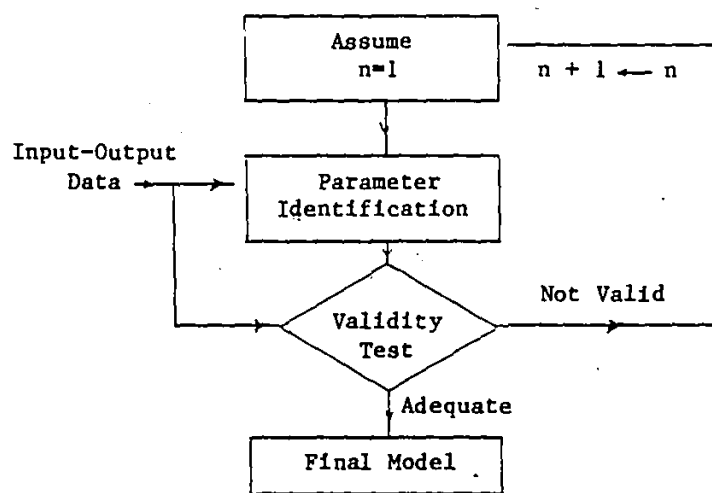
The Information Criterion (AIC) and FPE criteria [67], [68] appear to be very powerful practical approaches to the problem of model structure identification so long as the noise is normally distributed. When the noise distribution is not normal the information criteria tend to give larger values than the actual ones [69].

Suen and Liu [69] developed an algorithm 'the normalized residual technique' to estimate the structure of multivariable stochastic linear systems with the knowledge of only the signal to noise ratio. This algorithm have been successfully utilized by El-Sherief and Sinha [81].

Box and Jenkins [70] established a time series analysis method. A good model for the time series data is the one which is parametrically efficient (parsimonious) and simultaneously provides a low residual estimation error variance (i.e., a good explanation of the data) and low parametric estimation error variance. Box and Jenkins method has been discussed and utilized in Chapters 3 and 4 of this thesis. The most recent method in the order determination field is

Young et al.'s instrumental variable method [71] which will be presented in a coming section.

The difficulty with the tests based on analyzing the residuals is that they iterate on the system order and they require the calculation of the estimated parameter at every step [65] as we see in Figure (5.1).



**Figure. 5.1** Iterative procedure of tests based on analysing the residuals

In this chapter three order determination tests are proposed to be applied for multivariable continuous-time systems when approximate integration is performed on the data.

The methods under investigation are the instrumental variable method for model order identification, the residual error technique and

the instrumental determinant ratio test. From the three methods only the instrumental variable method was applied to continuous-time systems and in the simulation the direct approach with approximate integration was not utilized [71].

The continuous-time model of the multivariable system will be identified in the transfer function form for all three tests. The  $i^{\text{th}}$  differential equation representing the relation of the  $i^{\text{th}}$  output and the different inputs is given by:

$$D^{n_i} y_i(t) + a_{n_i-1} D^{n_i-1} y_i(t) + \dots + a_0 y_i(t) = B_{n_i-1} D^{n_i-1} \underline{u}(t) + \dots + B_0 \underline{u}(t) + e_i(t) \quad (5.4)$$

where  $e_i(t)$  is the error in the model of the  $i^{\text{th}}$  output.

The information matrices used in the three procedures are redefined to suit the new interpretation of the data. The proposed methods have been applied to a simulated three outputs two inputs continuous-time system. The different properties of each method are exposed based on the obtained results.

## 5.2 Reformulation of the Information Matrices

All system identification methods depend on the available input-output data to estimate the model order or parameters. We define by the information matrices the observation matrix  $G$  composed of the

input and output observations and the instrumental variable matrix  $W$ . The latter matrix has to satisfy the following conditions [2].

$$E[W^T \underline{v}] = 0 \quad (5.5)$$

$$E[W^T G] = R \quad (5.6)$$

where  $\underline{v}$  is the noise vector superimposed on the system.

An unbiased and consistent estimate of the parameters vector  $\underline{\theta}$  is obtained as

$$\underline{\theta}_{IV} = (W^T G)^{-1} W^T \underline{y} \quad (5.7)$$

Equations (5.5) and (5.6) imply that the instrumental variables are uncorrelated with the noise, and on the other hand strongly correlated with the inputs and outputs. To satisfy these conditions the IV sequence  $\underline{z}$  are chosen as the output of an auxiliary model with the same input as the system under investigation and provided that the filter which generates  $\underline{z}$  is stable and of order  $n_1$  or greater.

The IV method can be utilized on-line and off-line in a recursive form easy to implement on digital computers hence avoiding the matrix inversion in (5.7).

$$\underline{\theta}_k = \underline{\theta}_{k-1} + \frac{P_{k-1} \underline{\omega}_k \left[ \frac{d^{n_1}}{dt^{n_1}} y_k - \underline{g}_k^T \underline{\theta}_{k-1} \right]}{1 + \underline{g}_k^T P_{k-1} \underline{\omega}_k} \quad (5.8)$$

$$P_k = P_{k-1} - \frac{P_{k-1} \underline{\omega}_k \underline{g}_k^T P_{k-1}}{1 + \underline{g}_k^T P_{k-1} \underline{\omega}_k} \quad (5.9)$$

where

$$P_k = \left[ \begin{array}{c} k \quad T \\ \sum_{i=1} \underline{\omega}_i \underline{g}_i \end{array} \right]^{-1} \quad (5.10)$$

$\underline{\omega}_k$  and  $\underline{g}_k$  are defined by Young [61] for the single-input single-output case as:

$$\underline{g}_k = \left[ \frac{-d^{n_i-1}}{dt^{n_i-1}} y_k, \frac{-d^{n_i-2}}{dt^{n_i-2}} y_k, \dots, -y_k, \frac{-d^{n_i-1}}{dt^{n_i-1}} u_k, \dots, u_k \right] \quad (5.11)$$

$$\underline{\omega}_k = \left[ \frac{-d^{n_i-1}}{dt^{n_i-1}} z_k, \dots, -z_k, \frac{-d^{n_i-1}}{dt^{n_i-1}} u_k, \dots, u_k \right] \quad (5.12)$$

A new approach is now applied to identify continuous-time systems using approximate integration. So, let us redefine the information matrix  $G$  and the instrumental variables matrix  $W$  on the light of this new perspective for the system in (5.4).

let  $z_{n_1}^j =$  the  $n_1^{\text{th}}$  integral of  $z^j(t) = \int_0^t \int_0^t \dots \int_0^t z^j(t) dt$

$z_{n_1 k}^j =$  the  $k^{\text{th}}$  sample of the  $n_1^{\text{th}}$  integral of  $z^j(t)$

$u_{n_1}^1 =$  the  $n_1^{\text{th}}$  integral of  $u^1(t)$

$u_{n_1 k}^1 =$  the  $k^{\text{th}}$  sample of the  $n_1^{\text{th}}$  integral of  $u^1(t)$

$y_{n_1}^j$  and  $y_{n_1 k}^j$  are defined in the same manner.

For  $M$  observations  $W$  can be redefined as:

$$W = \begin{bmatrix} u_{11}^1 & u_{21}^1 & \dots & u_{n_1 1}^1 & u_{11}^2 & \dots & u_{n_1 1}^2 & \dots & u_{11}^m & \dots & u_{n_1 1}^m & -z_{11}^j & \dots & -z_{n_1 1}^j \\ u_{12}^1 & u_{22}^1 & & & & & & & & & & & & & \\ \cdot & \cdot & & \cdot & \cdot & & \cdot & & \cdot & & \cdot & & \cdot & & \\ \cdot & \cdot & & \cdot & \cdot & & \cdot & & \cdot & & \cdot & & \cdot & & \\ \cdot & \cdot & & \cdot & \cdot & & \cdot & & \cdot & & \cdot & & \cdot & & \\ u_{1M}^1 & u_{2M}^1 & & u_{n_1 M}^1 & u_{1M}^2 & & u_{n_1 M}^2 & & u_{1M}^m & \dots & u_{n_1 M}^m & -z_{1M}^j & \dots & -z_{n_1 M}^j \end{bmatrix}$$

(5.13)

G is defined in a similar manner by substituting  $z_n^j$  by  $y_n^j$ .

A slight change is also a necessity in equation (5.8) to become:

$$\underline{\theta}_k = \underline{\theta}_{k-1} + \frac{P_{k-1}\omega_k(y_k^j - \underline{g}_k^T \underline{\theta}_{k-1})}{1 + \underline{g}_k^T P_{k-1} \omega_k} \quad (5.14)$$

where  $y_k^j$  is the  $k^{\text{th}}$  sample of the  $j^{\text{th}}$  output without any integration.

We will now proceed to give a brief description about each of the order determination tests mentioned in the introduction.

### 5.3 The Instrumental Variable Method of Model Order Identification

This method was proposed by Young et al. [71] and the procedure can be summarized as follows:

1. Specify a range of model orders using all available information.
2. Use the IV recursive algorithm and calculate for each model order:
  - a)  $E\hat{V}N(\hat{n}_1)$  and
  - b)  $R_T^2(\hat{n}_1)$

where

$$E\hat{V}N(\hat{n}_1) = \frac{1}{\text{number of parameters}} \sum_{i=1}^{\text{number of parameters}} \sigma^2 P_{ii} \quad (5.15)$$



and  $\sigma^2$  is the variance of  $e(t)$  in (5.4) and

$$R_T^2 = 1 - \frac{J_0}{\sum_{i=1}^M y_i^2} \quad (5.16)$$

where  $J_0$  is the residuals sum of squares.

3. Select that model structure which yields the best combination of EVN and  $R_T^2$  [71]

- a) The EVN attains or is close to its minimum value.
- b)  $R_T^2$  should be consistent with the degree of model fit expected by the analyst.

Also for further incremental increase in model order, it should not increase substantially and should tend to "plateau". Other statistical factors were mentioned in [71] but EVN and  $R_T^2$  are the major indicators of  $n_1$ .

This procedure was applied after the necessary changes of the information matrices indicated in the previous section have been made.

#### 5.4 The Instrumental Determinant Ratio Test (IDR)

In his 1978 paper Wellstead [63] described the order determination methods based upon the ~~EM~~ as a rough estimate of the order of an unknown system. The speed with which these methods can assess an approximate order to the system is a very attractive feature. The IDR test developed by Wellstead has been applied to numerous discrete-time

applications. In this chapter, it is applied on multivariable continuous-time systems.

The instrumental product moment matrix (IPMM) can be defined as:

$$S(\hat{n}_1) = W^T G \quad (5.17)$$

and

$$IDR(\hat{n}_1) = \frac{\det S(\hat{n}_1)}{\det S(n_1+1)} \quad (5.18)$$

$W$  and  $G$  are the instrumental variable matrix and the observation matrix as defined in Section 5.2.

The rank of the IPMM matrix should collapse when  $\hat{n}_1 > n_1$ . Because of the presence of noise, the rank condition is not so clear but the value  $IDR(\hat{n}_1)$  should increase compared to the previous value of  $IDR(\hat{n}_1-1)$  if  $\hat{n}_1$  is the true order.

### 5.5 The Residual Error Technique (RET)

It is an off-line non-recursive method first introduced by Suen and Liu [69] to identify the structural indices of multivariable discrete-time systems in a certain canonical form. It was then utilized by El-Sherief and Sinha [81] to identify another canonical structure of discrete multivariable systems.

Considering the case of noisy data the RET theory is briefly stated [2]:

Let  $\underline{y}^*$  be the noise-free output vector and  $\underline{v}$  the zero mean noise vector then

$$\underline{y} = \underline{y}^* + \underline{v} \quad (5.19)$$

If  $\underline{y}^*$  is a linear combination of a set of other vectors where  $X = [\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n]$  then there exists a non zero vector such that

$$\underline{y}^* = X\theta \quad (5.20)$$

The optimal solution in the noise-free case should be:

$$\underline{\theta}^{\circ} = (X^T X)^{-1} X^T \underline{y}^* = X^+ \underline{y}^* \quad (5.21)$$

$$e(\theta^{\circ}) = \underline{y}^{*T} (I - X X^+) \underline{y}^* \quad (5.22)$$

then (5.19) may be written as:

$$\underline{y} = X\theta + \underline{v} \quad (5.23)$$

the residual error is thus obtained by substituting in (5.22) for  $\underline{y}^*$  by  $\underline{y}$  we get

$$e(\theta^0) = e^0 = \underline{y}^T(I-XX^+)\underline{y} \quad (5.24)$$

$$\begin{aligned} &= \underline{y}^{*T}(I-XX^+)\underline{y}^* + 2 \underline{y}^{*T}(I-XX^+)\underline{v} \\ &+ \underline{v}^T(I-XX^+)\underline{v} \end{aligned} \quad (5.25)$$

Lemma:

Assuming that  $\underline{v}$  is zero mean and uncorrelated with  $\underline{y}^*$  then if  $\underline{y}^*$  is a linear combination of  $\{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\}$  we have

$$E\{e^0\} = E \underline{v}^T(I-XX^+)\underline{v} \quad (5.26)$$

Otherwise

$$E\{e^0\} = E \{ \underline{v}^T(I-XX^+)\underline{v} \} + \underline{y}^{*T}(I-XX^+)\underline{y}^* \quad (5.27)$$

It is clear that the R.H.S. of equation (5.27) is greater than that of equation (5.26).

For more details, the reader could be referred to [2,69,81]. When the RET is applied to estimate the structural indices of continuous-time multivariable systems, it can be seen that the matrix  $X$  is simply the observation matrix  $G$ . So using the definitions stated in section 5.2, the residual error of the  $i^{\text{th}}$  subsystem can be expressed as:

$$e_i^0(n_1) \triangleq \underline{y}_1^T(M) [I-G_1(n_1;M) G_1^+(n_1;M)] \underline{y}_1(M) \quad (5.29)$$

El-Sherief and Sinha [81] indicated that in practice it is better to plot  $e_i^{o*}(n_i)$  against  $n_i$ , where

$$e_i^{o*}(n_i) = e_i^o(n_i) - e_i^o(n_i-1) \quad (5.30)$$

### 5.6 Results of Simulation

A 3 output 2 input multivariable continuous-time simulated system is sampled at the rate of 0.07 sec. according to the empirical rule  $\lambda T \leq 0.5$  [20] where  $\lambda$  is the farthest eigenvalue from the origin.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \frac{s+3}{(s+1)(s+2)} & \frac{-2}{(s+1)(s+2)} \\ \frac{s^2+10s+15}{(s+2)(s+3)(s+4)} & \frac{1}{(s+2)} \\ \frac{s+9}{(s+3)(s+4)} & \frac{s+10}{(s+3)(s+4)} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$n_1 = 2, n_2 = 3, n_3 = 2$$

$n_i$  is the structural index of the  $i^{\text{th}}$  subsystem

$u_1$  = unit step

$u_2$  =  $\sin(t)$

The three methods described in sections 5.3 to 5.5 have been applied to the simulated multivariable system with superimposed white noise ranging from low noise SNR = 20 db to medium noise SNR = 10 db.

The input-output data have been integrated using the cubic spline method [37], which provides good approximation to the original signals and their integrals.

Young et al.'s procedure has been applied to the simulated multivariable example with  $S = 500$  samples sets of input and output data. The instrumental variables sequence was generated with two IV schemes. First the IV sequence is taken as the output of a known fifth order system with the same input as the system under investigation, this is called the ordinary IV. The second scheme utilizes the system model as auxiliary model which is adaptively updated. The IV sequence and the output are adaptively integrated with the trapezoidal pulse function technique at each recursive step.

The results showed that regardless of the method used to generate the IV sequence the value of EVN in equation (5.15) in Young's method indicates a first order system in most of the studied cases. A better indicator of the true order of a system would be to compare the variance of the residuals generated by each model, see Table 5.1. The procedure with ordinary IV indicated the correct structure of the third subsystem  $n_3 = 2$  with  $S = 600$  samples, the first subsystem  $n_1 = 2$  with  $S = 800$  samples, see Figures 5.2 and 5.3, and failed to identify the structural index of the second subsystem  $n_2 = 3$  using  $S = 800$  samples at SNR = 20 db. We can notice that the procedure needs a sufficiently

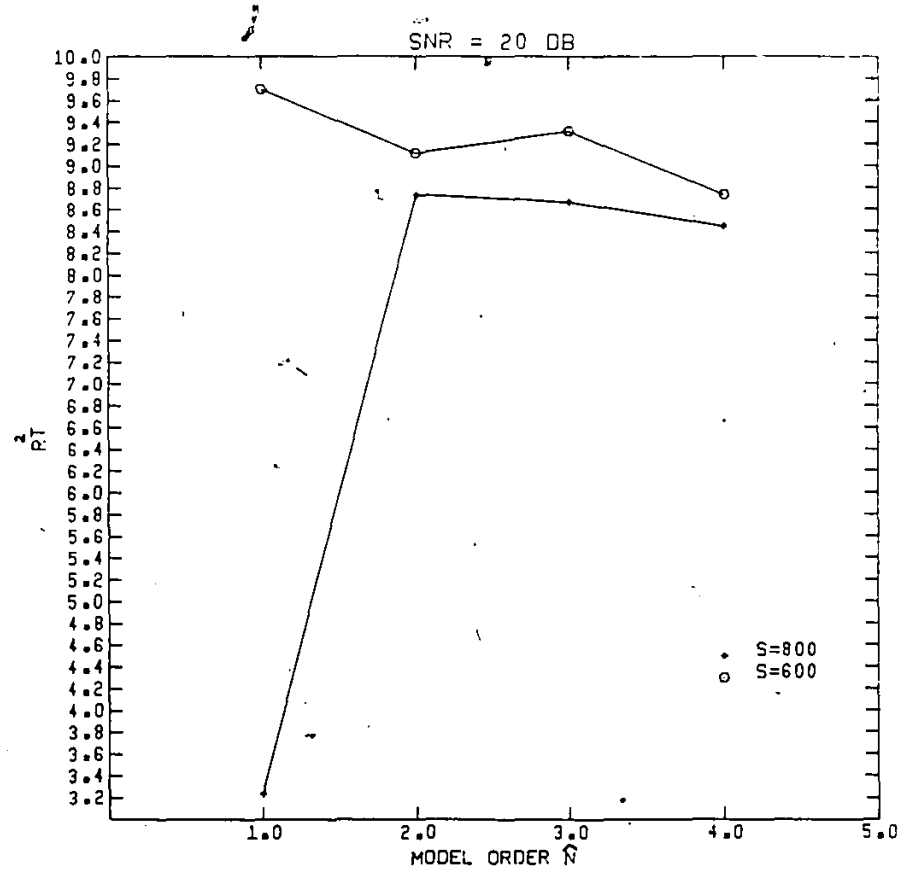


Figure 5.2 The estimation of  $n_1$  with the ordinary IV test

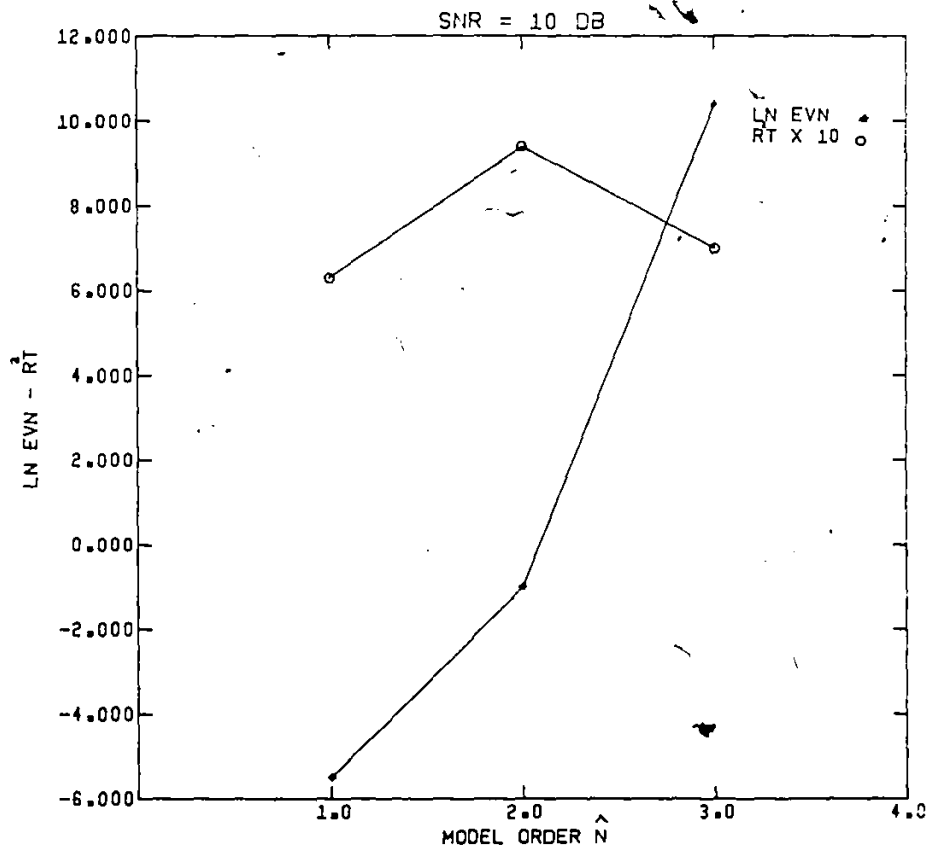


Figure 5.3 The estimation of  $n_3$  with the ordinary IV test



large amount of samples of data to precise the order of each subsystem when using an ordinary IV algorithm, which makes it computationally demanding for higher order systems. The same procedure with adaptively updated and integrated auxiliary model used only  $S = 500$  samples and did correctly identify the structural indices of the three subsystems as can be seen in Figure 5.4 from input-output data contaminated with white noise, SNR = 10 db.

**Table 5.1 Variance of the residuals white noise, SNR = 10 db, S = 500 samples adaptive auxiliary model**

Structure	$n_1$	$n_2$	$n_3$
1	49.8	1.85	11.88
2	3.88	7.5	0.202
3	82.	0.43	3.89
4	320.	6.44	1.5

**Table 5.2 Time comparison based on S = 600 samples**

Method	IDR	IV Method	RET
Time/iteration (sec)	0.8	7.4	3.6

The IDR test is a very fast test. It also carries the order determination part independently from estimating the parameters of the model. Figures 5.5 to 5.7 show the estimation of  $n_1$ ,  $n_2$  and  $n_3$  using  $S = 400-800$  samples. In presence of noise the IDR test tends to over estimate the value of  $n_1$  as can be seen in Figure 5.8. By raising the number of samples to 800 the test indicates in a vague way the correct

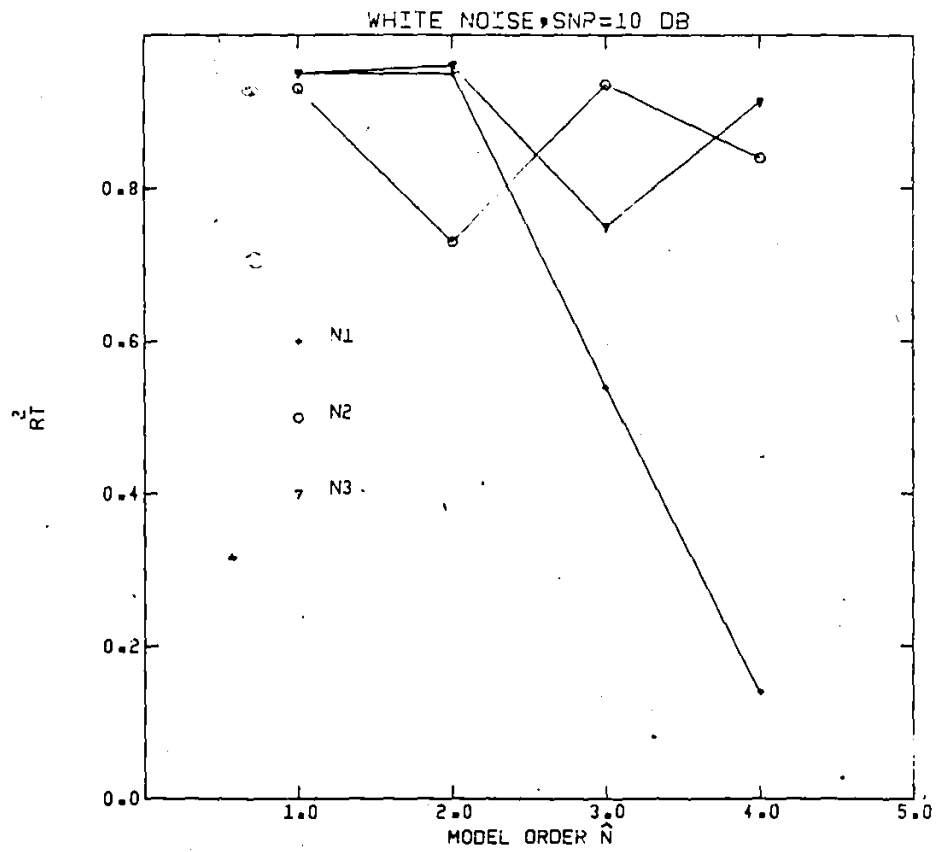


Figure 5.4 The estimation of the structural indices with the adaptive IV test

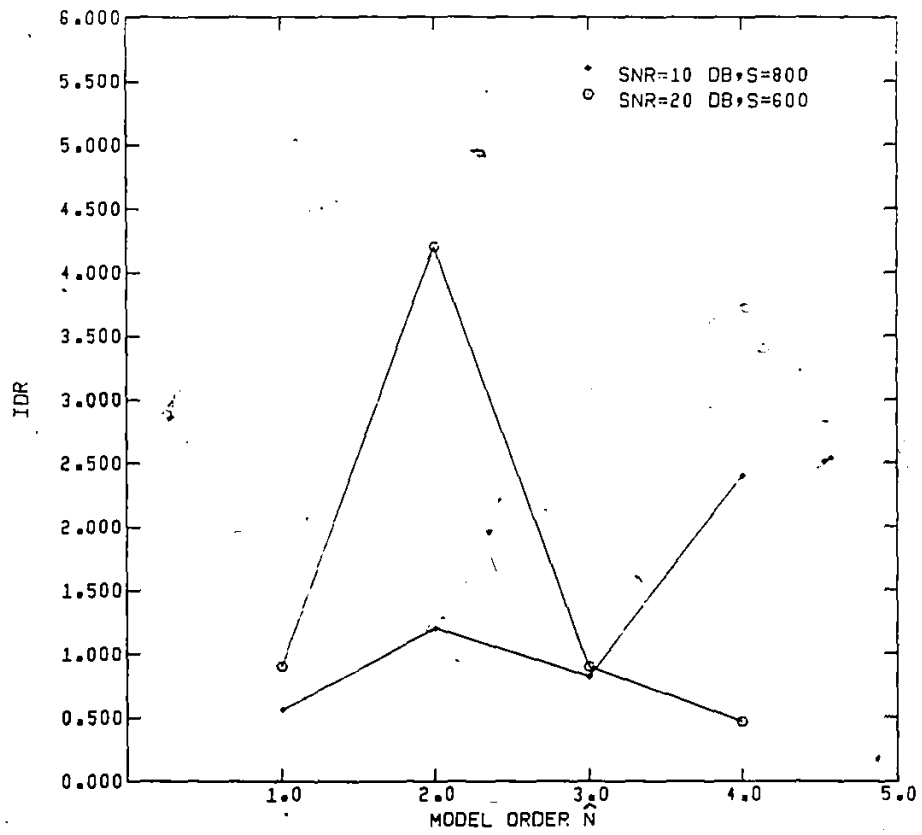


Figure 5.5 The determination of  $n_1$  with the IDR test

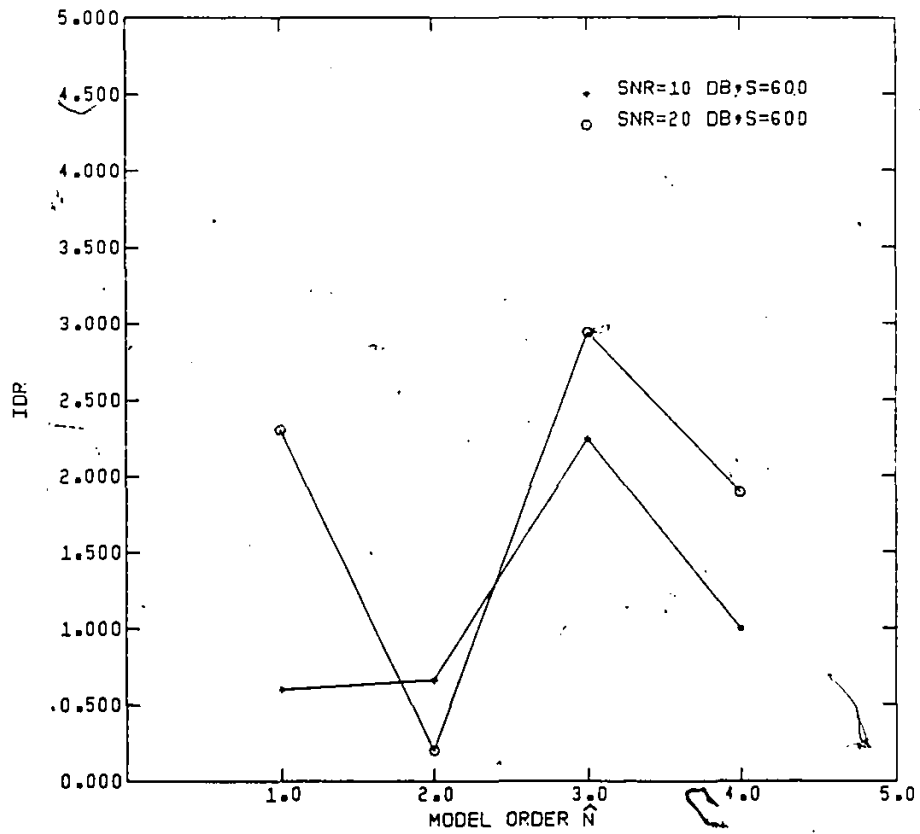


Figure 5.6 The determination of  $\hat{N}_2$  with the IDR test

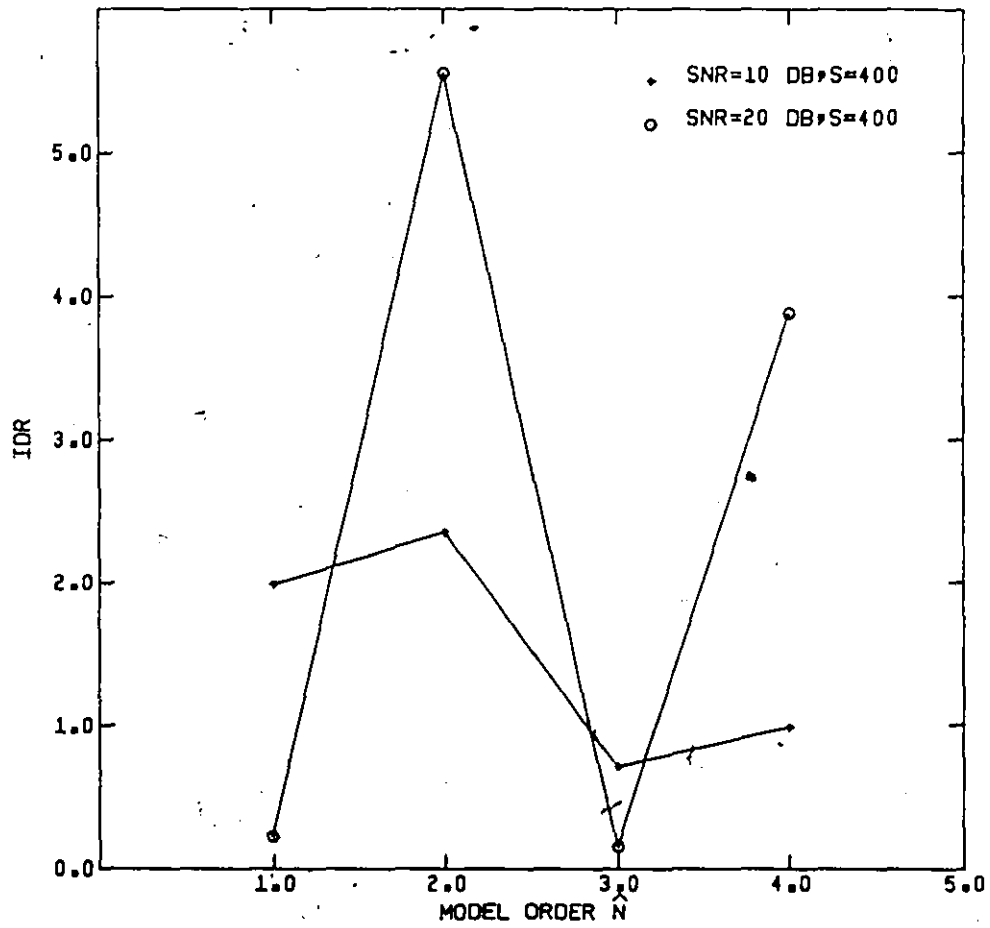


Figure 5.7 The determination of  $n_3$  with the IDR test

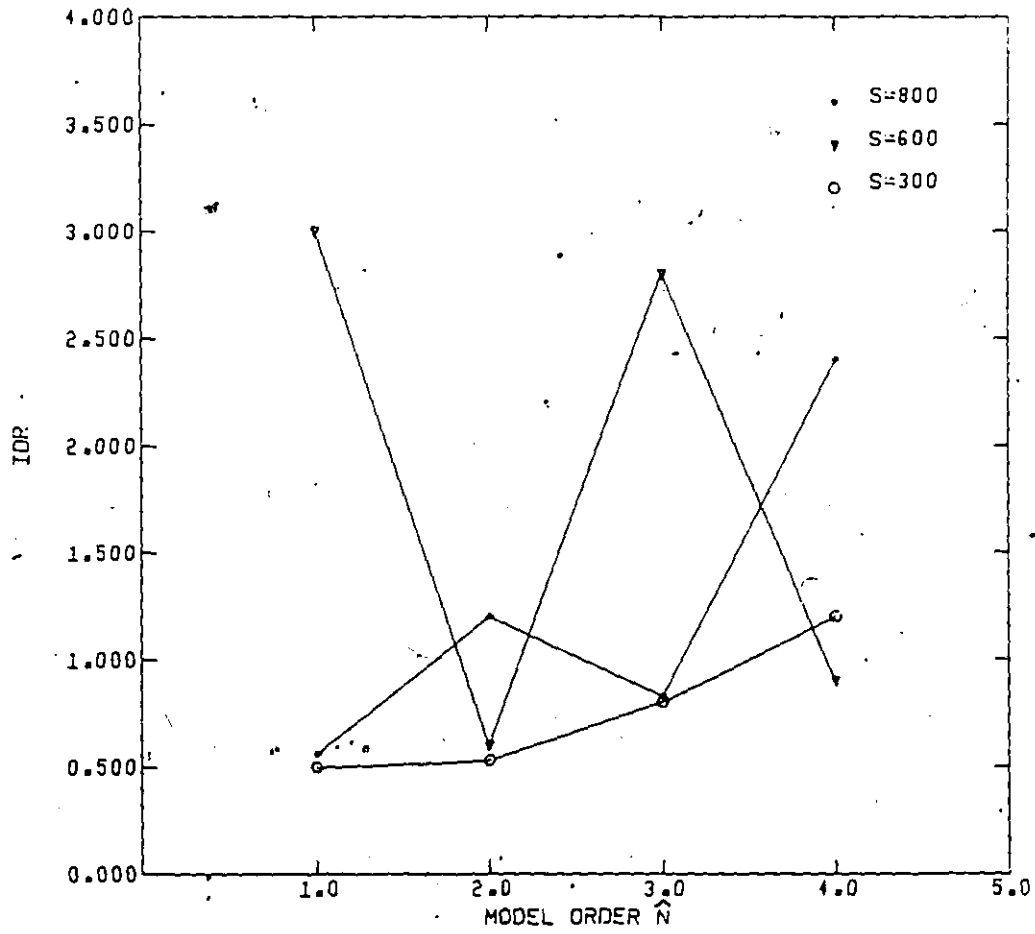


Figure 5.8 The determination of  $n_1$  with the IDR test at SNR = 10db

index  $n_1 = 2$  for SNR 10 db. This test does not have difficulty in estimating  $n_2 = 3$  in low noise and medium noise cases.

The RET requires large amount of storage. It operates on blocks of data non recursively. It is also slower than the IDR test because it does estimate the parameters of the model to obtain the value of the residuals. But the RET can detect the order of the structural indices in a short lengthed sequence of data which may alleviate some of the storage problem. Second order subsystems were correctly estimated with  $S = 150$  samples, but the third order subsystem could not be well identified until the algorithm was fed by  $S = 600$ . This result is clearly better than Young et al. procedure with ordinary IV which was not able to detect the correct index even with  $S = 800$ . (see Figures 5.9 to 5.12).

### 5.7 Concluding Remarks

In this chapter, the identification of the order or structural indices of multivariable continuous-time subsystems in the transfer function form from the samples of the input-output data has been proposed using three order determination methods. An important feature of the proposed methods is the operation of approximate integration of the sampled data.

The observation matrix and the instrumental variables matrix have been redefined to suit the new interpretation of the data. A simulated 3 output-2 input multivariable continuous-time system was used to illustrate the behaviour of each method. The three proposed

methods have given good results regardless of the noise superimposed on the data provided that they are fed with sufficient information. The proposed methods work well with various types of on-line and off-line algorithms for identification. The analyst can therefore choose the method which suits best the particular experiment's needs. But the residual error technique may be preferred as a good compromise between the accuracy and the used computation time.



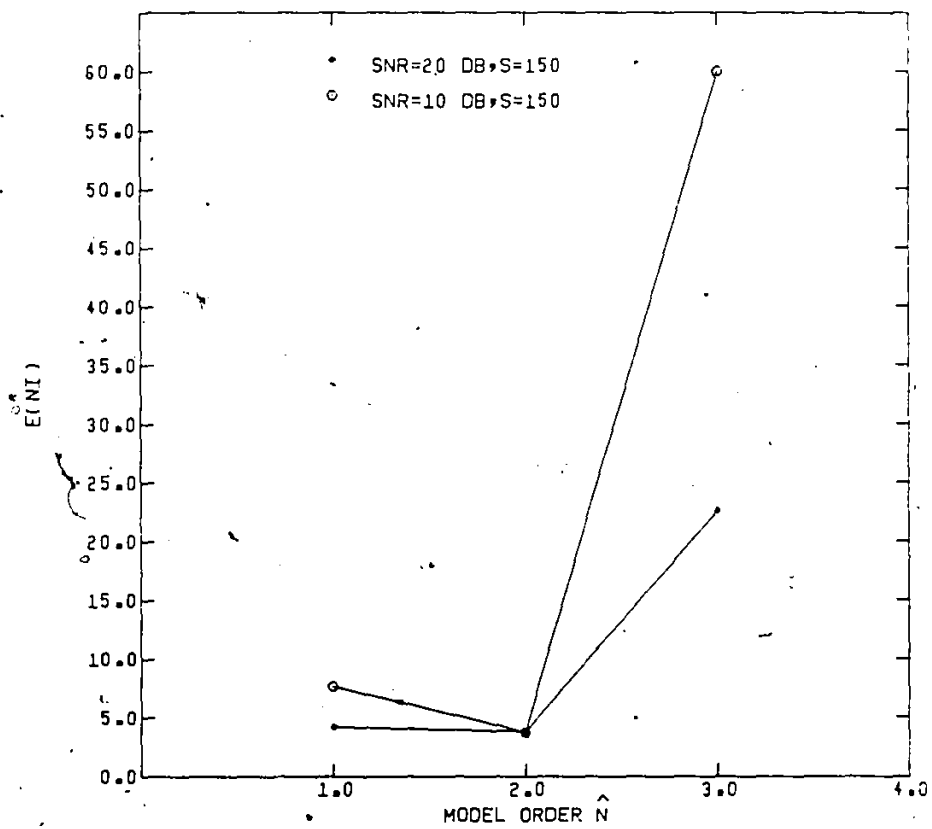


Figure 5.9 Residual difference plot of the first output

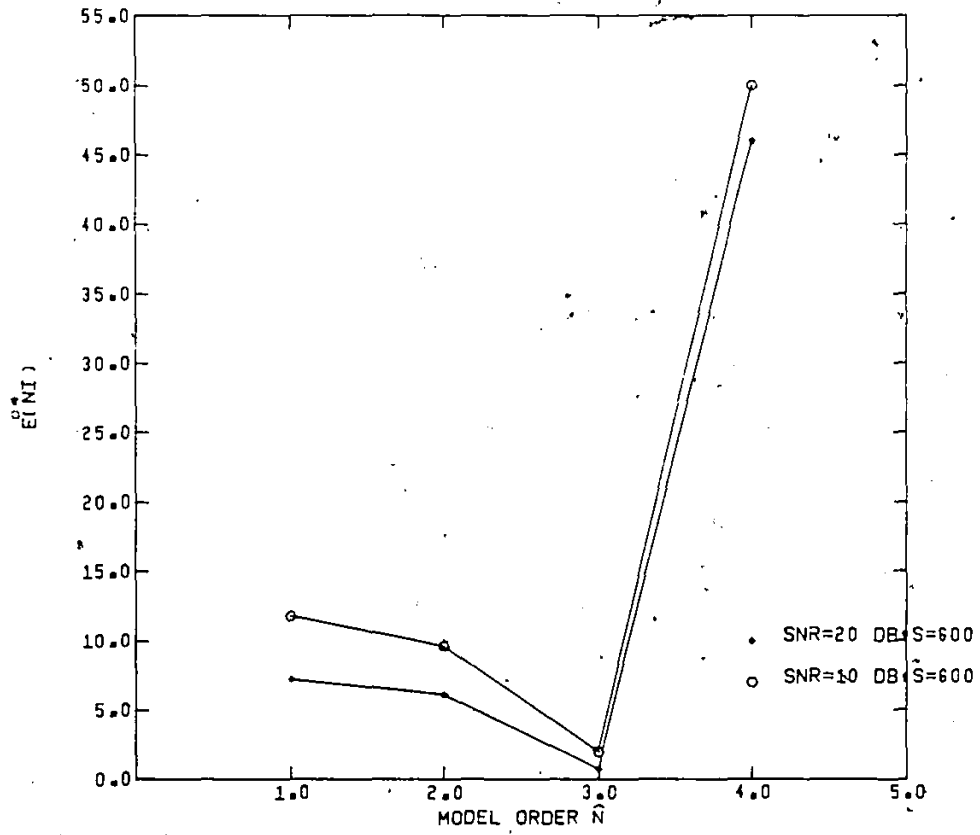


Figure 5.10 Residual difference plot of the second output

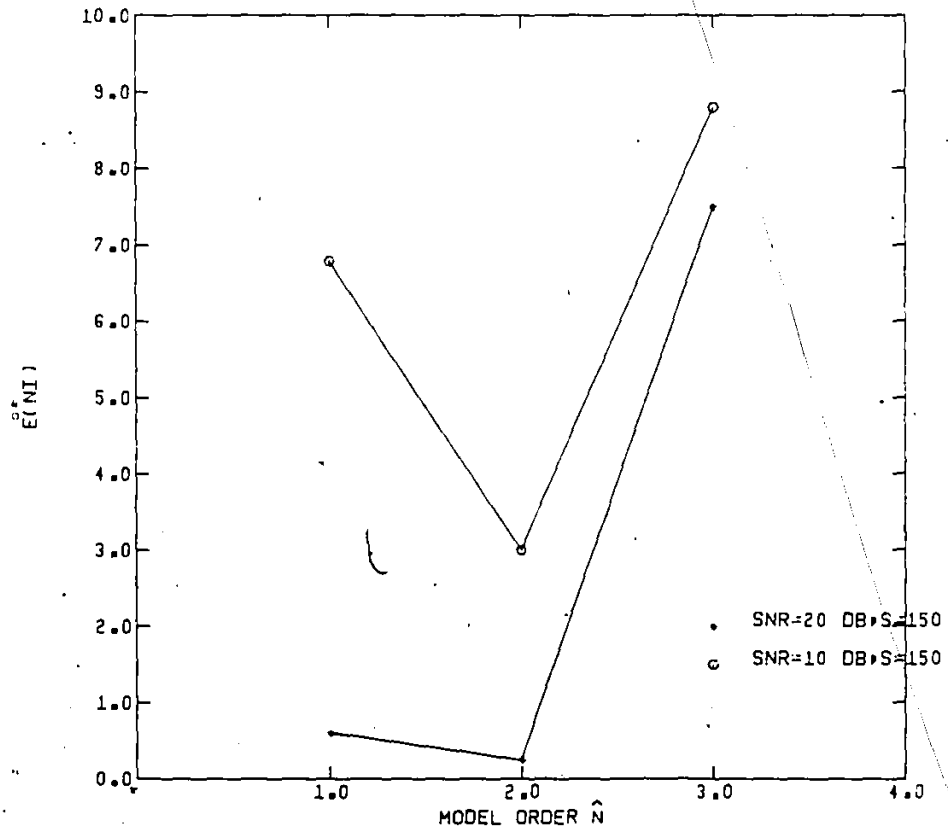


Figure 5.11 Residual difference plot of the third output

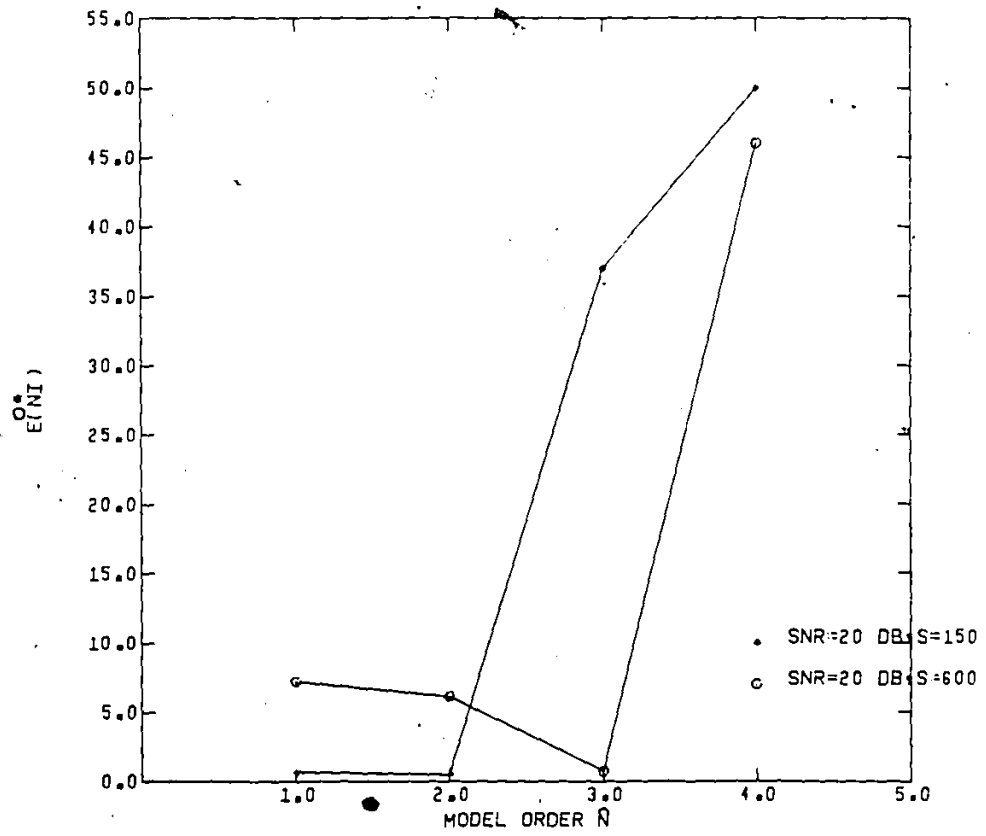


Figure 5.12 Residual difference plot of the second output at  $S = 150$  and  $600$  samples

## CHAPTER 6

### STRUCTURE SELECTION FOR MULTIVARIABLE CONTINUOUS-TIME SYSTEMS

#### 6.1 Introduction

The identification of single-input single-output systems consists of two parts: order determination and parameters estimation. In the case of multi-input multi-output systems a third and important step is added to the identification procedure: the selection of a model structure.

A multivariable system of known order can be represented by many different structures. The problem is to choose one which will give well-conditioned parameterization. Different approaches have been suggested and used in the past. The problem can be simplified by the use of a priori knowledge about the system, but often the system is considered as a black box and this approach cannot be used. Other researchers have reduced the multivariable system to several single-output systems by decoupling the outputs. This approach is also not realistic. The most popular approach during the past decade has been to describe specific systems within given equivalence classes by using suitable canonical forms [24], [72-73]. The canonical form implies that the system is represented with one unique model structure. The application of this approach in recursive off-line or on-line

identification scheme is critical when the system changes its structure and the model cannot track the changes. Hence, the principle of overlapping parametrization was introduced by Glover and Willems [74], Ljung and Rissanen [75]. Based on this idea Beghelli and Guidorzi [76], Van Overbeek and Ljung [82-83], have proposed procedures to solve the structure and identification problems of discrete-time multi-variable systems.

The overlapping models have a reduced but not minimal parametrization and can describe within a given order, several structures according to the actual parameter values. These parameterizations overlap, so that a change of parametrization can be made without loss of information. We have also the advantage of selecting, among different parametrizations, a well conditioned one.

In this chapter a new procedure which uses the overlapping parametrization approach is proposed to identify the structure of multivariable systems in the input-output form [84]. This procedure will enable the transformation from an ill-conditioned parametrization to a better conditioned one of the same order whenever it is necessary. There are several ways of defining a "best" parametrization for the representation of a stationary finite dimensional multivariable stochastic process. A logical method would be to select the structure that minimizes some scalar measure of the information matrix that corresponds to each parametrization. Wertz, Gevers and Hannan [85] showed that the determinants of all these information matrices are

asymptotically equivalent, and that this criterion is therefore unable to discriminate between different structures.

A criterion which discriminates between structures and decides when to switch to a different structure based on the principle of complexity is presented in section 6.4

In section 6.5 a simulated example is developed for both noise free and noisy cases to demonstrate the applicability of the proposed procedure.

## 6.2 A New Recursive Structure Selection Procedure for Multivariable Continuous-Time Systems in the Input-Output Form

We summarize the problem as follows:

Given a set of input-output data representing a continuous multivariable system of known order; select the structure that will give well-conditioned parametrization, i.e., that will not cause numerical problems during the minimization of the prediction error.

Input-output models of multivariable systems are very useful and practical in on-line identification of dynamic systems. These models consist of a direct link between the input and output samples and can thus be directly estimated from the process observations. Usually records of the noisy samples are available and the system is treated as a black box. Canonical input-output representations and their equivalence to canonical state space models have been investigated by Guidorzi [24], [86]. Multistructural input-output models were discussed by Gevers and Wertz [87], Beghelli and Guidorzi [88].

A multivariable continuous-time system is represented by the following equation in the input-output form:

$$P(D)\underline{y}(t) = Q(D)\underline{u}(t) + \underline{e}(t) \quad (6.1)$$

$\underline{y}(t)$  and  $\underline{u}(t)$  are the  $p$  dimensional output vector and the  $m$  dimensional input vector, respectively.  $P(D)$  and  $Q(D)$  are polynomial matrices in  $D$  (or  $d/dt$ ) the time derivative operator and  $\underline{e}(t)$  is a sequence of independent random,  $p$  dimensional vector, with zero mean value.

A new recursive procedure for structure selection is proposed [84], based on the overlapping parametrization approach. The main feature of this procedure, besides tackling this problem for continuous-time systems for the first time, is a criterion for ill conditionness. Hence, the transformation from an ill conditioned parametrization to a better conditioned one of the same order occurs only whenever it is necessary.

The proposed procedure is illustrated in the general flow chart in Figure 6.1. We can start with any pre-identified structure. The cubic spline technique is then used to approximate the various integrals of the inputs and the outputs utilizing the available data as explained in the earlier chapters (to minimize the time, trapezoidal or block pulse functions could be used instead). We then proceed to the identification of the system model by minimizing a recursive prediction error criterion.



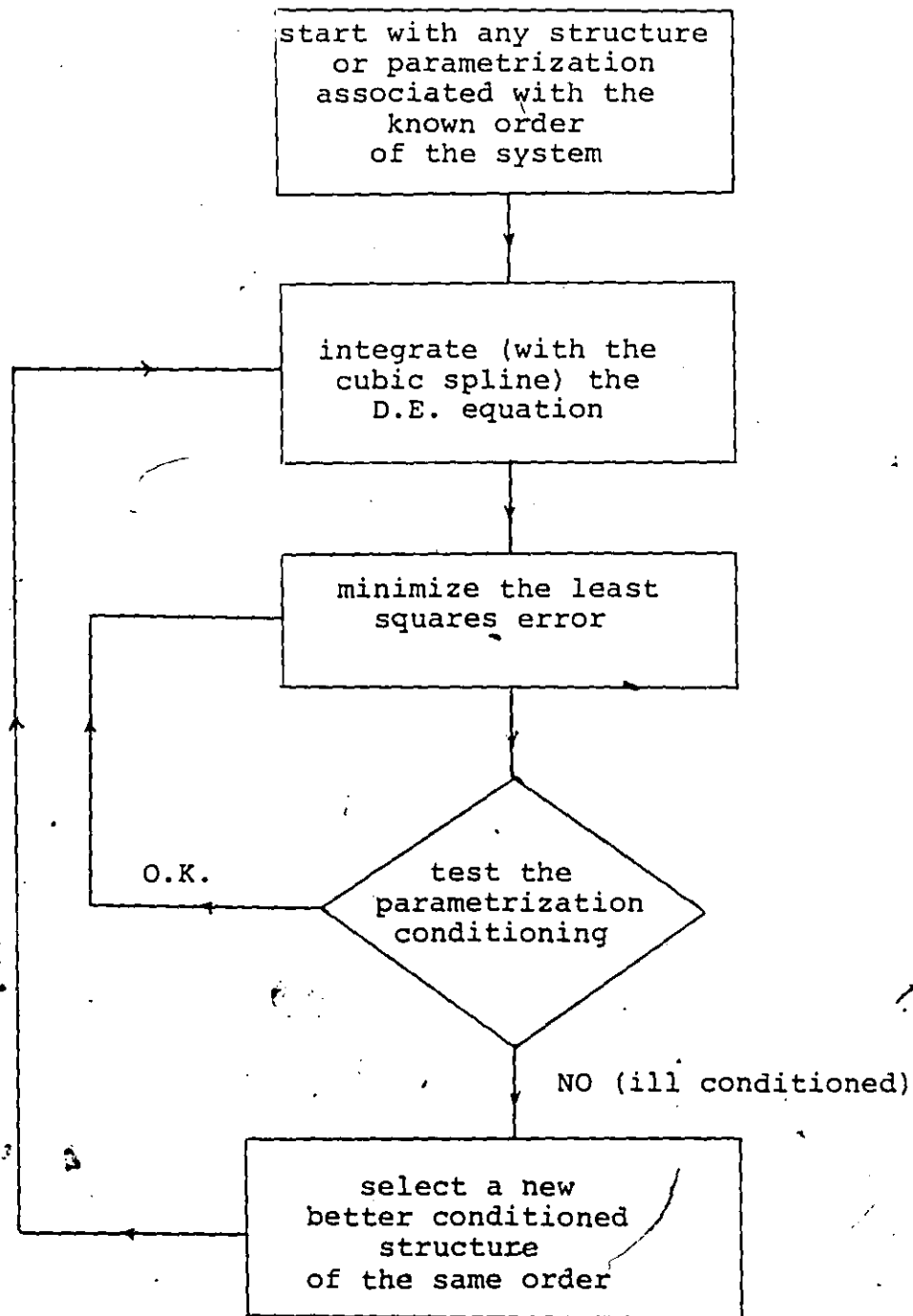


Figure 6.1 Flow chart of the structure selection procedure

After every  $M$  iterations a test is performed on the parametrization for ill-conditioning. If the result of the test is positive then the tested structure is good. If, on the other hand, the parametrization fails the test then it is starting to be ill-conditioned and the structure should be changed. Another structure for the system is selected which will have better conditioned parametrization. A change in the initial structure occurs once or several times during the procedure if we did not start with the true structure.

### 6.2.1 The Model Identification

The parameter estimation part is performed using the recursive least squares algorithm which is the simplest and most practical algorithm for on-line parameter estimation.

The algorithm in its basic recursive form is described by equations (6.2 and 6.3).

$$\hat{\theta}_k = \hat{\theta}_{k-1} - \frac{P_{k-1}g_k [g_k^T \hat{\theta}_{k-1} - y_k]}{1 + g_k^T P_{k-1} g_k} \quad (6.2)$$

$$P_k = P_{k-1} - P_{k-1}g_k [1 + g_k^T P_{k-1} g_k]^{-1} g_k^T P_{k-1} \quad (6.3)$$

$\hat{\theta}_k$  is the estimate of the parameters vector,  $g_k$  and  $y_k$  are the available information organized in an appropriate way [89] for the case

of identifying continuous-time systems using the direct approach (see Chapter 5).

$$P_k = \left[ \sum_{i=1}^k g_i g_i^T \right]^{-1} \quad (6.4)$$

Another more useful version of the algorithm [77] makes use of the statistical assumptions of the error term in (6.1). Provided that the following assumptions are satisfied: 1) the estimation error vector  $\tilde{\theta}_k = \hat{\theta}_k - \theta_k$  has zero mean value, 2)  $E\{e_k\} = 0$ , then the variance-covariance matrix  $P_k^* \triangleq E\{\tilde{\theta}_k \tilde{\theta}_k^T\}$  is related to the matrix  $P_k$  in (6.3) and (6.4) by  $P_k^* = \sigma^2 P_k$

Substituting  $P_k$  by  $\frac{P_k^*}{\sigma^2}$  in (6.2) and (6.3) yield

$$\hat{\theta}_k = \hat{\theta}_{k-1} - P_{k-1}^* g_k (g_k^T \hat{\theta}_{k-1} - y_k) (\sigma^2 + g_k^T P_{k-1}^* g_k)^{-1} \quad (6.5)$$

$$P_k^* = P_{k-1}^* - P_{k-1}^* g_k g_k^T P_{k-1}^{*T} (\sigma^2 + g_k^T P_{k-1}^* g_k)^{-1} \quad (6.6)$$

The variance of the residuals can be also computed recursively as in equation (6.7).

$$\sigma_k^2 = \sigma_{k-1}^2 - \frac{1}{k} [\hat{e}_k \hat{e}_k - \sigma_{k-1}^2] \quad (6.7)$$

This version of the least squares gives not only the parameter estimates and the variance of the residuals at each sampling instant, but also gives an indication of the accuracy of the estimates through the error covariance matrix  $P_k^*$ .

### 6.2.2 The Switching Criterion

Complexity is a measure of the interaction between the components of a random vector. The more the interaction there is, the larger the complexity.

Van Emden [90] showed that the complexity can be expressed using the covariance matrix of the random vector.

Maklad [91-92] suggested that a compromise between the whiteness of model residues and the accuracy of its estimated parameters be utilized. The derivation of this criterion is given in Appendix A. So Maklad computed the complexity of the joint random variables  $((\underline{e}|\hat{\underline{\theta}}), \hat{\underline{\theta}})$ , where  $\hat{\underline{\theta}}$  is the estimated parameters vector and  $\underline{e}$  are the model residuals as seen in Figure 6.2. It is noted that the information about  $\underline{e}$  is only attainable after the outcome of  $\hat{\underline{\theta}}$  is given. This is the reason for conditioning  $\underline{e}$  on  $\hat{\underline{\theta}}$ .

$$\text{COMP} = \left[ \frac{1}{N} \sum_{i=1}^N P_{ii}^{*2} - \left( \frac{\text{trace } P^*}{N} \right)^2 + \frac{2}{N} \sum_{i=1}^N \sum_{j=i+1}^N P_{ij}^{*2} \right] \quad (6.8)$$

$$+ \frac{2}{N_1} \sum_{i=1}^L (N_1 - 1) \hat{r}(i)$$

where  $P^*$  is the error covariance matrix,  $N$  the number of parameters,  $\hat{r}(i)$  the estimated autocorrelation of the residuals at lag  $i$ .  $N_1$  is the length of the available data and  $L$  is the first lag at which  $\hat{r}^2(L+1) < \text{var}(\hat{r}(1))$ . Maklad used his criterion in model order discrimination single-input, single-output systems and proved that it can detect the case of correlated residuals unlike AIC and FPE criteria [91]. An ill-conditioned parametrization will give a large complexity. So to monitor the conditioning of the parametrization of the tested structure the following procedure is suggested.

1. After every  $M$  iterations compute the complexity of the joint random variables  $((\underline{e}|\hat{\theta}), \hat{\theta})$  for the tested structure, the error covariance matrix  $P^*$  and the residuals are both supplied directly from the recursive least squares algorithm without any added computation.
2. Test the ratio of the old complexity over the new complexity value. If the ratio is greater than 1, then the structure is well-conditioned and the algorithm is trying to improve both the parameter estimates, and the whiteness of the residuals.
3. If the ratio is less than 1, then this parametrization is ill-conditioned and a change of structure is required.

#### Remarks

1. After every change in structure the first two or three computed complexities have in general high values because

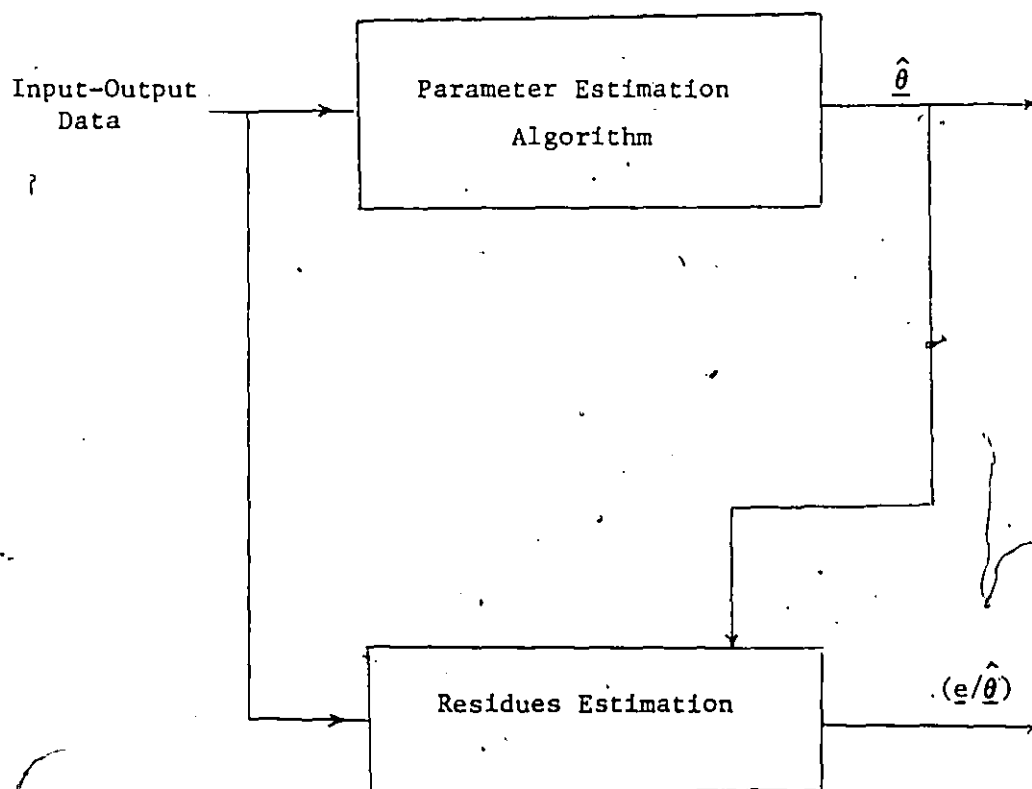


Figure 6.2 Parameter and Residues Estimation

the initial parameters to the identification algorithm are far from the true ones. As more data are fed to the algorithm the complexity starts to decrease very quickly for a suitable structure of the system.

2. It is better to choose  $M$  not very large so that if an ill parametrization is detected, and a change in structure is necessary, then the transformed parameters used as initial parameters will not constitute a bad start for the least-squares algorithm.

The switching criterion is important in determining when to switch to another better-conditioned structure. It eliminates the need to switch between parametrizations unnecessarily to look for a better structure like in [76], [83]. The presence of this criterion reduces considerably the amount of computations.

### 6.2.3 Selection and Transformation to a Better Structure

Guidorzi [93] has set in 1982 a definition for the adjacent pseudostructures associated with the same system (see Appendix B). He then developed an algorithm which transforms a system with pseudostructure  $\mu_i$  into an equivalent system with an adjacent pseudostructure  $\mu_j$ . The transformation algorithm consists of elementary row operations of the polynomial matrices  $P$  and  $Q$  and gives both the new pseudostructure and the new parameters of the system (Appendix B). To select between adjacent structure he proposed the following criterion.

Given an input-output multistructural model, consider the scalars  $\delta_i$  ( $i=1, \dots, p$ ) given by

$$\delta_i = \left( \sum_{j=1}^p \sum_{k=1}^{n_j} \alpha_{ijk}^2 \right)^{\frac{1}{2}} \quad (6.9)$$

where  $\alpha_{ijk}$  are elements of the  $P(D)$  polynomial matrix, and define the conditioning factor  $\rho = \delta_{\max}/\delta_{\min}$ . If we have two equivalent multistructural models, the best conditioned one is the model with the lower conditioning factor.

#### 6.2.4 Summary

We may summarize the structure selection procedure for continuous-time multivariable system [84] in the following steps:

1. Start with any pseudostructure.
2. Integrate the differential equations of the assumed input-output model. You may use as integration tools, the block pulse functions, trapezoidal pulse functions or cubic spline method, which approximates the original functions from the given data samples with various degrees of accuracy.
3. Arrange the observation matrix with the samples of the input-output integrals [89], and perform the least squares algorithm.



4. Compute, every M iterations, the complexity of the joint random variables  $((\underline{e}|\hat{\theta}), \hat{\theta})$  and follow the ill-conditioning test as demonstrated in Section 6.2.2.
5. If the structure passes the test go to 4.
6. If the test indicates ill conditioning, switch to an adjacent pseudostructure and compare with the selection criterion (6.9). If a better conditioned parametrization is found go to 3 to arrange the new information matrix. If the tested structure has the least conditioning factor then no better pseudostructure can be found. New analysis is advisable to change the order of the system.

### 6.3 Results of Simulation

To illustrate the applicability of the proposed procedure it has been tested on a simulated two-inputs two-outputs fourth-order continuous-time system. The input and output signals have been sampled at the rate of  $T = 0.07$  sec providing 500 input-output data samples.

The input-output model of the system is

$$\begin{bmatrix} D^2 + 1.6D + 0.6 & 0.5D + 0.2 \\ - & D^2 + 0.5D + 0.06 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} D + 2.6 & D + 1.1 \\ -1 & -D - 0.5 \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$$

The integrations have been done with the cubic spline integration method.

### 6.3.1 Noise-Free Case

Start the procedure with the three possible pseudostructures of the system (3,1), (2,2), (1,3) and see the ability of the procedure to decide when the tested parametrization is ill-conditioned and should be changed. The example illustrates the new selected structure which have a better conditioned parametrization as seen in Table 6.1.

**Table 6.1 Structure selection in noise-free case**

Pseudostructure	3,1	1,3	2,2
No. of iterations before change	100	40	500
No. of test calls	5	2	25
Complexity	$2.83 \times 10^{11}$	61188660.4	0.06
	0.00009858	$1.830766 \times 10^{11}$	$1.86 \times 10^{-9}$
C.F. (conditioning factor)	14.69	12.95	3.57
Decision	Switch to adjacent structure		No change
C.F. (2,2) (conditioning factor)	2.53	1.4536	
Decision	Use pseudostructure 2,2		
No. of iterations	400	460	
COMP	0.06629	17780.908	
	1.3673	7020.769	
Total execution time in sec.	15.006	15.819	14.761

### 6.3.2 The Noisy Case

To test the structure selection procedure in presence of noise, five percent zero mean white noise was added to the outputs of the tested system. The addition of noise did not affect the selection of a well conditioned parametrization for the system as seen in Table 6.2.

**Table 6.2 Structure selection in noisy case**

Pseudostructure	3,1	1,3	2,2
No. of iterations before change	60	60	500
No. of test calls	3	3	25
Complexity	$2.8377 \times 10^{11}$	0.06129	0.06
	0.84955	$1.339 \times 10^{11}$	0.00139
C.F. (conditioning factor)	2.218	54.97	3.425
Decision	Switch to adjacent pseudostructure		No change
C.F. (2,2) (conditioning factor)	1.66817	4.082	
Decision	Switch to pseudostructure 2,2		
No. of iterations	440	440	
COMP	0.06	20075.7	
	1.4	8120.63	

#### 6.4 Discussion

Overbeek and Ljung [82] proposed a computationally more expensive selection criterion in their identification in multistructural state space forms (6.10). It is based on computing the complexity of the state covariance matrix of the tested structure, then an organized search for a structure which will have more independent rows for the F matrix is conducted. The complexity of the new state covariance matrix is calculated and compared with the initial one until a better conditioned structure is found.

$$\begin{aligned} x(k+1) &= Fx(k) + G e(k) \\ y(k) &= Hx(k) + e(k) \end{aligned} \tag{6.10}$$

Their procedure was later criticized by Guidorzi [93]. He argued that the iterative computation necessary to obtain P (the state covariance matrix) can prove to be time-consuming in on-line identification particularly when the model with the best conditioning is requested. He then proposed the fast transformation and selection criterion described in section 6.2.3 and Appendix B and used in this chapter. However, the structure selection algorithm of Guidorzi lacked an important part and that is how to determine the suitable time to change the structure of the system? Otherwise, we have to test repeatedly all the possible structures to select the best conditioned one. This will also be time consuming specially when the order of the

system is high. To overcome this problem Beghelli and Guidorzi [76] imposed on the data more analysis. Model and structure identification was performed on 200 samples every 75 samples. Structural identification have been performed by the 'range error test' [76] which selects the most suitable ones for the system under consideration, and the switching is done only between a limited number of structures.

In the procedure proposed in this chapter a criterion which detects ill conditioning is presented in section 6.2.2 and hence we can save all the effort of repeated model and structure identification. When the chosen structure starts to give an ill-conditioned parameterized, then at this moment alone that the switching between the adjacent pseudostructure begins until the structure with a smaller conditioning factor than the initial one is found.

### 6.5 Concluding Remarks

A step-by-step procedure for identifying the structure of continuous-time multivariable system is presented for the representation in the input-output form. This procedure is suitable for both stationary and non-stationary systems when a change of structure occurs while the order of the system remains constant. It uses the overlapping multistructural parametrization approach to choose a better conditioned parametrization to the multivariable system whenever ill-conditioning is detected. The switching decision is based on the complexity principle which provides a good monitor for the conditioning of the parametrization as well as the suitability of the tested

structure. The inclusion of this criterion in the selection procedure answered the question of when to switch to another pseudostructure. This leads to the reduction of the time required for computations when compared with other structure selection procedures. The overlapping parametrization principle allows to obtain the new parameters values of the new selected structure with the aid of a similarity transformation with no loss of information.

## CHAPTER 7

### CONCLUSIONS

The major effort in this thesis has been directed towards the problem of identification of continuous-time systems from input-output data. The identification with the direct approach converts the original differential input-output model to a linear algebraic model convenient for a direct solution. It has been shown that based on the available samples, the input and output signals can be reconstructed with approximating functions. The system differential equations can be integrated using these approximations and the results used for estimating the parameters of the model. The system identification problem of multivariable systems consists of three main steps, order determination, structure selection and parameter estimation. In this thesis the three parts of the problem were addressed and studied. New approaches were developed for the structure selection and the parameter estimation of continuous-time systems. Order determination procedures previously used for discrete-time systems were applied for the first time for continuous-time systems. These algorithms find their direct applications in digital adaptive control, digital simulation of continuous-time systems as well as the identification of the system through samples of the input-output data.

A comprehensive survey is presented in Chapter 2 to discuss the identification of continuous-time systems from sampled data with the direct and the indirect approaches. In the indirect approach a discrete-time model is obtained from the samples of the observations then a corresponding continuous-time model is derived. The advantages and the problems of the two approaches were discussed.

Three approximating functions were proposed recently in the literature [37], [48] and [52]. They are the block pulse functions, trapezoidal pulse functions and cubic splines. It has been shown in Chapter 3 and Chapter 4 that the cubic splines method is superior to the other two methods in the noise-free case. When the observations were contaminated with noise the three methods gave identical parameter estimates. Since in practice both the accuracy of the estimates and the computation time are important, the trapezoidal pulse functions method is recommended. It has the advantage of being easily incorporated in any on-line algorithm. It also uses the actual data values unlike the block pulse functions method which requires an extra arithmetic operation.

It has been shown that the direct approach suffers some serious difficulties such as inaccurate parameter estimates and correlated residuals in presence of noise. A study was conducted in Chapter 3 for the single-input single-output case and then extended for the multi-variable case in Chapter 4. The study revealed that the errors in the approximation result in correlated residuals and that the approximate



integration give birth to an integrated noise term instead of the expected series of non-correlated random shocks.

Based on this new information, it was possible to deduce that the error term resulting from the combination of the two factors is not a white noise series as was implied in all the previous work conducted in this field.

A new approach is presented to overcome the combined effect of the errors in the approximation and additive white noise on the identification of continuous-time systems. The method consists of modelling the combined error term. The errors were absorbed by the error model, thus providing good parameter estimates and uncorrelated residuals of the system under consideration. Extensive simulations were conducted in order to illustrate the merits of the new procedure. The method worked successfully for different noise levels (10% and 20%) compared to 1.5% reported previously in the literature when no modelling was used. The modelling of the error in the proposed algorithm has been done using the time series method. Another advantage of the proposed method is its insensitivity to the choice of the sampling interval. It has been tested with systems sampled at different sampling intervals and the results showed a considerable improvement in the estimated parameters and their accuracy as well as the whitening of the residuals.

Order determination is the first part of the identification problem and it is generally done off-line. Three order determination tests have been studied in Chapter 5 and used to identify the

structural indices of multivariable continuous-time systems in the transfer function form. Two of those tests are applied for the first time for continuous-time systems as the author is aware. The three methods are: the instrumental variable method for model order identification, the instrumental determinant ratio test and the residual error technique. The sampled input-output signals have been integrated with the cubic spline technique and the information matrices have been reformulated to suit the new interpretation of the data. The three order determination tests have been compared according to the computation time, the number of input-output samples used to indicate the correct order and their robustness to added noise. The residual error technique was found to present a good compromise between all these factors.

Multivariable systems unlike single-input single-output systems can be represented within different structures. The problem is to choose one which will give a well conditioned parameterization. In Chapter 6 a new procedure is presented to identify the structure of continuous-time multivariable systems in the input-output form. This procedure is suitable for both stationary and non-stationary systems when a change in the structure occurs while the order remains constant. It uses the concept of overlapping parameterization to choose a better conditioned parameterization for the multivariable system whenever ill conditioning is detected.

A switching criterion is presented based on the complexity principle which provides a good monitor of the conditioning of the

parameterization as well as the suitability of the tested structure. The inclusion of this criterion in the selection procedure answered the question of when to switch to another pseudostructure. This leads to the reduction of time required for computations when compared with the other structure selection procedures [76]. The latter procedures all lack a switching criterion.

### 7.1 Suggestions for Further Research

1. The choice of a suitable sampling interval is an essential requirement in system identification. All identification procedures such as the direct and the indirect approaches are based on the fact that the sampling interval has been selected properly. But when the system is treated as a black box and there is no information about the system time constants we cannot predict a suitable sampling interval beforehand. Sinha and Puthenpura [21] have developed a criterion to determine the optimum sampling interval from the input and output data. It is based on the common rule that the sampling interval  $T$ , should be chosen such that

$$\lambda_f T \ll 0.5 \quad (7.1)$$

where  $\lambda_f$  is the magnitude of the largest eigenvalue of the continuous model. Since  $\lambda_f$  are not known a priori, they [21] used a bilinear transformation to map the area inside the unit circle so that it can be spread out to the entire left-half of the w-plane

$$w = \frac{z + 1}{z - 1} \quad (7.2)$$

so if  $z = e^{\alpha}$  the corresponding point in the w plane is given by

$$w = w_1 + j w_2 \quad (7.3)$$

$$\text{and } w_1 = \frac{e^{2\alpha} - 1}{e^{2\alpha} - 2e^{\alpha} \cos \beta + 1} \quad \text{and } w_2 = \frac{we^{2\alpha} \sin \beta}{e^{2\alpha} - 2e^{\alpha} \cos \beta + 1}$$

They proposed to place the poles at a distance

$$R = 5.00 = \frac{e^{-0.5} + 1}{e^{-0.5} - 1} \quad (7.4)$$

Simulations in Chapter 3 showed that the equality in equation (7.1) is not suitable for the identification with the direct approach by approximate integration. It resulted in a

large sampling interval value and hence inaccurate parameter estimates are obtained.

A more appropriate choice for this case is the rule-of-thumb proposed in this thesis.

$$0.05 < \lambda_f T < 0.1$$

(7.5)

If we use the upper limit the value of R becomes

$$R = \frac{e^{-0.1} + 1}{e^{-0.1} - 1} = 20.00$$

(7.6)

This will be more efficient than Sinha and Puthenpura [21] in the direct approach specially in the noisy case.

2. The proposed on-line algorithm for structure selection and the algorithms for parameter estimation of multivariable continuous-time systems are easy to use. Since the storage requirements for data are small and the identification algorithm with the trapezoidal pulse functions requires few arithmetic operations, they may be implemented in real time on a microcomputer. The direct application of this approach could be in the field of robotics and computer aided manufacturing.
3. Despite their attractive features for identification purposes, the input-output form and the transfer function form require

the approximate integration to be performed several times in order to solve the systems differential equations. This results in some deterioration in the parameter estimates. The direct identification in the state space form presents a promising alternative for higher order systems. As we see in equation (7.7)

$$\dot{\underline{x}} = \underline{A}\underline{x} + \underline{B}\underline{u} \quad (7.7)$$

the approximate integration is done only once.

Shrindhar et al. [35] were the first to use the spline technique in conjunction with the Kalman estimation procedure to generate required state variables and to minimize a performance index with respect to the system parameters.

However, the problem with the state space form is the additional burden of estimating the state variables. The convergence of bootstrap algorithms of the combined state and parameter estimation has not been theoretically justified. However, when each stage of these two stage algorithms is treated separately the convergence of each stage can be argued if the other one satisfies certain properties. Generally, the convergence of the overall algorithm is not obvious [22]. This aspect has to be further investigated.

## APPENDIX A

### THE COMPLEXITY CRITERION

#### 1. The Concept of Complexity

Definition: Complexity is a measure of the difference between a whole and the noninteracting composition of its components.

If we decompose a system  $S$  into two subsystems  $S_1$  and  $S_2$  which are the simplest possible components, we can consider the complexity of  $C_1(S)$  of the system  $S$  to be given by

$$C_1(S) = C_1(S_1) + C_1(S_2) + R(S_1, S_2) \quad (\text{A.1})$$

where  $R(\dots, \dots)$  signifies interaction between the arguments and where we considered  $\sum_{i=1}^n C_1(S_i) = 0$ ,  $n$  is the number of the simple components of the system.

Consider the two discrete random variables  $X$  and  $Y$  with outcomes  $(x_1, x_2, \dots, x_m)$  and  $(y_1, y_2, \dots, y_n)$ , respectively. Let the joint probabilities be given by

$$\Pr[X = x_i, Y = y_j] = \gamma_{ij} \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n$$

The marginal probabilities are then

$$\Pr[X = x_i] = p_i = \sum_{j=1}^n \gamma_{ij}, \quad i = 1, 2, \dots, m,$$

$$\Pr[Y = y_j] = q_j = \sum_{i=1}^m \gamma_{ij}, \quad j = 1, 2, \dots, n,$$

Define two independent random variables  $X^*$ ,  $Y^*$  with the same outcomes as  $X$  and  $Y$ , respectively, then

$$\Pr[X^* = x_i, Y^* = y_j] = \Pr[X^* = x_i] \cdot \Pr[Y^* = y_j]$$

$$= \Pr[X = x_i] \cdot \Pr[Y = y_j]$$

$$= p_i q_j$$

$$i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n$$

There would have been no interaction between  $X$  and  $Y$  if they had the same joint probabilities as  $X^*$  and  $Y^*$ . Thus a measure of the interaction between  $X$  and  $Y$  may be the degree of fit (or discrimination) between the probability distribution of  $(X, Y)$  and that of  $(X^*, Y^*)$ .



For this purpose the Kullback-Leibler's information for discrimination will be employed [94]. This yields

$$\begin{aligned}
 R(X, Y) &= \sum_{i=1}^m \sum_{j=1}^n \gamma_{ij} \ln \frac{\gamma_{ij}}{p_i q_j} \\
 &= \sum_{i=1}^m \sum_{j=1}^n \gamma_{ij} \ln \gamma_{ij} \\
 &\quad - \sum_{i=1}^m \sum_{j=1}^n \gamma_{ij} [\ln p_i + \ln q_j] \\
 &= \sum_{i=1}^m \sum_{j=1}^n \gamma_{ij} \ln \gamma_{ij} \\
 &\quad - \sum_{i=1}^m p_i \ln p_i - \sum_{j=1}^n q_j \ln q_j \\
 &= -I(X, Y) + I(X) + I(Y) \\
 &= I(X) - I(X/Y) \\
 &= I(Y) - I(Y/X) \tag{A.2}
 \end{aligned}$$

where  $I(\cdot)$  is the entropy of a random variable (or of a probability distribution) defined by Shannon [95].

In general for  $k$  random variables  $X_1, X_2, \dots, X_k$ :

$$R(X_1, X_2, \dots, X_k) = \sum_{i=1}^k I(X_i) - I(X_1, X_2, \dots, X_k) \quad (A.3)$$

is a measure of the interaction between these random variables.

## 2. Derivation of Maklad's Complexity Criterion

In system identification the desirable model has properties that are related to the parameter estimates  $\hat{\theta}$  and model residues  $\underline{\epsilon}_N$ . Maklad considered the complexity of the joint random variables  $((\underline{\epsilon}_N/\hat{\theta}), \hat{\theta})$ .

Using (A.1)

$$C_1((\underline{\epsilon}_N/\hat{\theta}), \hat{\theta}) = C_1(\underline{\epsilon}_N/\hat{\theta}) + C_1(\hat{\theta}) + R((\underline{\epsilon}_N/\hat{\theta}), \hat{\theta})$$

From (A.2)

$$\begin{aligned} R((\underline{\epsilon}_N/\hat{\theta}), \hat{\theta}) &= I(\underline{\epsilon}_N/\hat{\theta}) - I((\underline{\epsilon}_N/\hat{\theta})/\hat{\theta}) \\ &= I(\underline{\epsilon}_N/\hat{\theta}) - I(\underline{\epsilon}_N/\hat{\theta}) \\ &= 0. \end{aligned}$$

Regarding the complexities of  $(\epsilon(1)/\hat{\theta})$  and  $\hat{\theta}_j$ ;  $i = 1, 2, \dots, N$ , and  $j = 1, 2, \dots, k$  where  $k = \dim \hat{\theta}$ , as zeros, we get,

$$\begin{aligned}
 C_1((\underline{\epsilon}_N/\hat{\theta}), \hat{\theta}) &= R(\epsilon(1), \dots, \epsilon(N)/\hat{\theta}) + R(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k) \\
 &= \sum_{i=1}^N I(\epsilon(1)/\hat{\theta}) - I(\underline{\epsilon}_N/\hat{\theta}) \\
 &\quad + \sum_{j=1}^k I(\hat{\theta}_j) - I(\hat{\theta})
 \end{aligned} \tag{A.4}$$

$\underline{\epsilon}_N$  has a zero mean and the empirical covariance matrix

$$\hat{R} = [r_{ij}] \quad i, j = 1, 2, \dots, N$$

where  $r_{ij}$  are the estimated autocorrelations of the residuals

$$I(\underline{\epsilon}_N/\hat{\theta}) = \frac{N}{\lambda} [\ln 2\pi + 1] + \frac{1}{2} \ln \det \hat{R} \tag{A.5}$$

$$I(\epsilon(t)/\hat{\theta}) = \frac{1}{2} [\ln 2\pi + 1] + \frac{1}{2} \ln \hat{r}(o) \tag{A.6}$$

Using the estimated covariance  $Q$  for the actual covariance, we get

$$I(\hat{\theta}) = \frac{k}{2} (\ln 2\pi + 1) + \frac{1}{2} \ln \det Q \quad (\text{A.7})$$

and

$$I(\hat{\theta}_1) = \frac{1}{2} (\ln 2\pi + 1) + \frac{1}{2} \ln \hat{q}_{11} \quad (\text{A.8})$$

Using (A.7), (A.8)

$$C_1(\underline{\epsilon}_N/\hat{\theta}) = \frac{1}{2} \sum_{i=1}^N \ln \hat{r}(0) - \frac{1}{2} \ln \det \hat{R} \quad (\text{A.9})$$

and

$$C_1(\hat{\theta}) = \frac{1}{2} \sum_{i=1}^k \ln \hat{q}_{11} - \frac{1}{2} \ln \det Q \quad (\text{A.10})$$

where  $\hat{q}_{11}$  is the  $i$ th diagonal element of  $Q$ .

But  $C_1(\hat{\theta})$  given by (A.10) is not valid for discrimination another one value which depends only on  $Q$  and would not change under orthogonal transformations is given by:

$$C_2(\hat{\theta}) = \frac{k}{2} \ln ((\text{trace } Q)/k) - \frac{1}{2} \ln \det Q \quad (\text{A.11})$$

A simpler measure was suggested by Van Emden [90] and has the form:

$$C(\hat{\theta}) = \frac{1}{k} \sum_{i=1}^k \ln \hat{q}_{ii}^2 - \left( \frac{\text{trace } Q}{k} \right)^2 + \frac{2}{k} \sum_{i=1}^k \sum_{j=i+1}^k \hat{q}_{ij}^2 \quad (\text{A.12})$$

and

$$\begin{aligned} C(\underline{\varepsilon}_N / \hat{\theta}) &= \frac{1}{N} \sum_{i=1}^N \ln \hat{r}^2(0) - \left( \frac{\text{trace } \hat{R}}{N} \right)^2 + \frac{2}{N} \sum_{i=1}^k \sum_{j=i+1}^k r_{ij} \\ &= \frac{2}{N} \sum_{j=1}^{N-1} (N-j) \hat{r}^2(j). \end{aligned} \quad (\text{A.13})$$

To compute  $C(\underline{\varepsilon}_N / \hat{\theta})$ , Maklad did not consider all  $\hat{r}(i)$  up to  $i = N - 1$ . This is due to the stationarity of  $(\underline{\varepsilon}(t) / \hat{\theta})$ , where  $\hat{r}(i)$  should decay with increasing  $i$ . We now determine an upper bound for  $i$ .

Employing Bartlett's results [96] concerning the variance of the estimates of the autocorrelations of normally distributed, stationary time series:

$$\begin{aligned} \text{var } [r(i)] &= \frac{1}{N} \sum_{j=1}^N \{ r^2(j) + r(j-1)r(j+1) \\ &\quad - 4r(1)r(j)r(1-j) + 2r^2(1)r^2(j) \} \end{aligned}$$

Since  $\{r^2(1)\}$  is decaying, the dominant term in the above expression is

$$\text{var} [\hat{r}(1)] = \frac{1}{N} \sum_{j=-\infty}^{\infty} r^2(j) \quad (\text{A.14})$$

$$= \frac{1}{N} \left[ \sigma^4(N, \hat{\theta}) + 2 \sum_{j=1}^{N-1} \hat{r}^2(j) \right].$$

In calculating  $C(\underline{\varepsilon}_N/\hat{\theta})$ , Maklad considered the first  $L$  correlations, where  $L$  is the first lag at which

$$\hat{r}^2(L+1) < \text{var} [\hat{r}(1)].$$

Thus, Maklad rule will be

$$\text{COMP} = \left[ \frac{1}{k} \sum_{i=1}^k \hat{q}_{1i}^2 - \left( \frac{\text{trace } Q}{k} \right)^2 + \frac{2}{k} \sum_{i=1}^k \sum_{j=i+1}^k \hat{q}_{1j}^2 \right] \quad (\text{A.15})$$

$$+ \left[ \frac{2}{N} \sum_{i=1}^L (N-1) \hat{r}^2(1) \right]$$

APPENDIX B

MULTISTRUCTURAL INPUT-OUTPUT MODELS

1. Equivalence Between Input-Output Models

A linear time-invariant discrete-time system is described in the input-output representation by [76]

$$P(z) y(t) = Q(z) u(t) \quad (\text{B.1})$$

where  $t \in \mathbb{Z}$ ,  $y(t)$  is the  $m$  dimensional output vector,  $u(t)$  is the  $r$ -dimensional input vector,  $P(z)$  is a square non singular polynomial matrix in  $z$  (unitary advance operator) and  $Q(z)$  is an  $(m \times r)$  polynomial matrix in  $z$ .

For identification purposes two input-output models can be considered equivalent when they describe the same external behaviour [76]. If the additional condition of sharing the same order is added, the following well-known algebraic condition is obtained [97].

The input-output models  $\{P'(z), Q'(z)\}$  and  $\{P''(z), Q''(z)\}$  are equivalent if and only if

$$\begin{aligned} P'(z) &= M(z) P''(z) \\ Q'(z) &= M(z) Q''(z) \end{aligned} \quad (\text{B.2})$$

where  $M(z)$  is an  $(m \times m)$  non singular unimodular polynomial matrix.

As already noted, all the models (B.1) belonging to the same equivalence class with respect to the equivalence relation (B.2) have the same order and describe the same system.

## 2. Input-Output Multistructural Models

According to Beghelli and Guidorzi [76] the pair  $\{\hat{P}(z), \hat{Q}(z)\}$

$$\hat{P}(z) = [p_{ij}(z)] \quad (i, j = 1, \dots, m)$$

$$\hat{Q}(z) = [q_{ij}(z)] \quad (i = 1, \dots, m; j = 1, \dots, r)$$

defines an input-output multistructural model if and only if

$$\deg \{p_{ii}(z)\} > \deg \{p_{ji}(z)\} \quad \text{for } i \neq j \quad (\text{B.3a})$$

$$\deg \{q_{ij}(z)\} \leq \rho_i \quad (\text{B.3b})$$

where  $\rho_i$  is the degree of the  $i$ th row of  $\hat{P}(z)$ .

The entries of a multistructural pair  $\{\hat{P}(z), \hat{Q}(z)\}$  will be denoted in the following way

$$p_{ii}(z) = z^{\mu_i} - \alpha_{i1\mu_i} z^{(\mu_i-1)} - \dots - \alpha_{i11}, \quad \mu_i > 0 \quad (\text{B.3c})$$



$$p_{ij}(z) = -\alpha_{ij\mu_j} z^{(\mu_j-1)} - \dots - \alpha_{ij1} \quad (\text{B.3d})$$

$$q_{ij}(z) = \beta_{ij(\rho_i+1)} z^{\rho_i} + \dots - \beta_{ij2} z + \beta_{ij1} \quad (\text{B.3e})$$

The integers  $\mu_i (i = 1, \dots, m)$  define the output pseudostructure for the considered multistructural input-output model.

### 3. Adjacent Pseudostructures Definition (Guidorzi [93])

The pseudostructures  $\{\mu'_i\}$  and  $\{\mu''_i\}$  associated with the same system are called adjacent if there exist two integers  $\lambda$  and  $l$  ( $1 \leq \lambda, l \leq m$ ) such that

$$\mu''_\lambda = \mu'_\lambda + 1$$

$$\mu''_l = \mu'_l + 1$$

$$\mu''_i = \mu'_i \quad \text{for } i \neq \lambda, l$$

### 4. Transformation Algorithm Between Adjacent Models

Denote by  $\{\hat{P}'(z), \hat{Q}'(z)\}$  a multistructural model characterized by the pseudostructure  $\{\mu'_i\}$ , and by  $\{\hat{P}''(z), \hat{Q}''(z)\}$  an equivalent adjacent multistructural model characterized by the pseudostructure  $\{\mu''_i\}$ .

The transformation from the first model to the second can be obtained by means of the following elementary steps.

Step 1. Add to the  $\lambda$ th row of  $\hat{P}'(z)$  its  $\lambda$ th row multiplied by  $z/\alpha'_{\lambda\lambda}\mu'_{\lambda}$

Step 2. Exchange the  $\lambda$ th row of  $\hat{P}'(z)$  with the  $\lambda$ th one. Note that after these operations  $\deg \{p'_{\lambda\lambda}(z)\} = \mu'_{\lambda} + 1$  and  $\deg \{p'_{\lambda\lambda}(z)\} = \mu'_{\lambda} - 1$ .

Step 3. The entries  $p'_{ij}(z)$ ,  $i \neq j$ , are tested with respect to column condition (B.3b). When  $\mu'_{ij} = \deg \{p'_{ij}(z)\} < \mu'_j = \deg \{p'_{jj}(z)\}$ , no operations are performed. If after Step 2,  $\mu'_{ij} = \mu'_j$ , the degree of  $p'_{ij}(z)$  is lowered by subtracting from the  $i$ th row of  $\hat{P}'(z)$  its  $j$ th row multiplied by the ratio of the maximal degree coefficients in  $p'_{ij}(z)$  and  $p'_{jj}(z)$ . The previous operations are repeated until condition (B.3b) is achieved for all the polynomials  $p'_{ij}(z)$ .

Note that the operations performed in this step do not change the output pseudostructure obtained in Step 2.

Step 4. The  $i$ th row of  $\hat{P}'(z)$  ( $i=1, \dots, m$ ) is divided by the coefficient of  $z^{\mu'_i}$  in  $p'_{ii}(z)$ .

After Step 4 all the polynomials on the main diagonal of  $\hat{P}'(z)$  are monic. All the operations performed in the previous steps on the rows of  $\hat{P}'(z)$  must be simultaneously performed also on the rows of  $\hat{Q}'(z)$ .

Now denote by  $\{\hat{P}''(z), \hat{Q}''(z)\}$ , the polynomial pair obtained; the models  $\{\hat{P}'(z), \hat{Q}'(z)\}$  and  $\{\hat{P}''(z), \hat{Q}''(z)\}$  are equivalent according to

conditions (B.2), and the multistructural model  $\{P''(z), Q''(z)\}$  exhibits the new adjacent pseudostructure. The unimodular matrix  $M(z)$  describing the performed transformation can be obtained, if desired, as the product of the unimodular matrices describing the elementary row operations performed on  $P'(z)$  and  $Q'(z)$ .

The following corollary can be immediately deduced from Step 1 of the above algorithm.

**Corollary 1**

A model with output pseudostructure  $\{\mu'_1\}$  can be transformed to an adjacent model with pseudostructure  $\{\mu''_1\}$  if and only if the parameter  $\alpha'_{\lambda_1 \mu'_2}$  in the first model is non-zero.

**Remark**

The transformation of an input-output multistructural model to a different one, characterized by a non-adjacent output pseudostructure but belonging to the same equivalence class, can be performed by means of repeated applications of the algorithm.

## REFERENCES

- [1] Bergstrom, A.R., 'The construction and use of economic models', English Universities Press, London, 1967.
- [2] Sinha, N.K., and Kuszta, B., 'Modeling and identification of dynamic systems', Van Nostrand Reinhold Company, 1983.
- [3] Graupe, D., 'Identification of systems', Van Nostrand Reinhold, New York, 1972.
- [4] Mendel, J.M., 'Discrete techniques of parameter estimation: the equation error approach', Marcell Dekker Inc., New York, 1973.
- [5] Eykhoff, P., 'System identification', John Wiley and Sons, Ltd., London, 1974.
- [6] Sage, A.P., and Melsa, J.L., 'System identification', Academic Press, New York, 1972.
- [7] Rajbman, N.S., and Sinha, N.K., 'Identification of multivariable systems - a critical survey', Proc. of 14th Annual Allerton Conf. on Circuits and Systems, 1976.
- [8] Beck, J.V., and Arnold, K.J., 'Parameter estimation in Engineering and Science', John Wiley and Sons, New York, 1977.
- [9] Eykhoff, P., 'Process parameter and state estimation', Automatica, Vol. 4, pp. 205-233, 1968.
- [10] Freeman, T.G., 'Selecting the best linear transfer function model', Automatica, Vol. 21, pp. 361-370, 1985.
- [11] Guillemin, E.A., 'Synthesis of passive networks', Wiley, 1957.
- [12] Kalman, R.E., and Declaris, N., 'Aspects of network and system theory', Holt, Rinehart & Winston, pp. 311-325, 1971.
- [13] Shieh, L.S., Chen, C.F., and Huang, C.J., 'On identifying transfer functions and state equations for linear systems', IEEE Trans., AES-8, pp. 811-820, 1972.
- [14] Sinha, N.K., 'Estimation of transfer function of continuous system from sampled data', Proc. IEE, Vol. 119, pp. 612-614, 1972.

- [15] Strmcnik, S., and Bremsak, F., 'Some new transformation algorithms in the identification of continuous-time multivariable systems using discrete identification methods', Preprints 5th IFAC Symp. on Identification and System Parameter Estimation, Darmstadt, F.R. of G., pp. 397-407, 1979.
- [16] Hsia, T.C., 'On sampled data approach to parameter identification of continuous linear systems', IEEE Trans., AC-17, pp. 247-249, 1972.
- [17] Smith, F.W., 'System Laplace - transform estimation from sampled data', IEEE Trans, AC-13, pp. 37-55, 1968.
- [18] Mehra, R.K., 'Synthesis of optimal inputs for MIMO systems with process noise' in 'System identification: Advances and Case Studies', Academic Press: New York, 1976.
- [19] Sargan, J.D., 'Some discrete approximations to continuous-time stochastic models', Int. Royal Stat. Soc, Series B, 36, pp. 74-90, 1974.
- [20] Haykin, S.S., 'A unified treatment of recursive digital filtering, IEEE Trans, AC-17, pp. 113-116, 1972.
- [21] Sinha, N.K., and Puthenpura, S., 'Choice of the sampling interval for the identification of continuous-time systems from samples of input-output data', Proc IEE on control theory and applications, Vol. 132, Pt. D, No. 6, pp. 263-267, 1985.
- [22] El-Sherief, H.E., 'Identification of linear multivariable discrete-time systems', Ph.D. thesis, McMaster University, Hamilton, Ontario, Canada, 1979.
- [23] Astrom, K.J., and Eykhoff, P., 'System identification - a survey', Automatica, Vol. 7, pp. 123-162, 1971.
- [24] Guidorzi, R., 'Canonical structures in the identification of multivariable systems', Automatica, Vol. 11, pp. 361-374, 1975.
- [25] Sinha, N.K., and Zhou, Q., 'Identification of multivariable continuous-time systems from samples of input-output data', Proc. of the 21st IEEE conf. on Decision and Control, Orlando, Florida, pp. 1189-1194, 1982.
- [26] Sinha, N.K. and Zhou, Q., 'Discrete-time approximation of multivariable continuous-time systems, IEE Proc., Vol. 130, Pt. D, pp. 103-110, 1983.
- [27] Sinha, N.K. and Lastman, G.T., 'Transformation algorithm for identification of continuous-time multivariable systems from discrete data', Electronics Letters, Vol. 17, pp. 779-780, 1981.

- [28] Lanczos, C., 'Applied analysis', Englewood Cliffs, N.J., Prentice-Hall, 1956.
- [29] Rao, G.P., and Sivakumar, L., 'Identification of deterministic time-lag systems', IEEE Trans. on Auto. Control, AC-21, pp. 527-29, 1976.
- [30] Shinbrot, M., 'On the analysis of linear and non-linear systems', Trans. ASME, Vol. 79, pp. 547-552, 1957.
- [31] Loeb, J.M., and Cahen, G.M., 'More about process identification', IEEE Trans. Autom. Control, AC-10, pp. 359-61, 1965.
- [32] Fairman, F.W., and Shen, D.W.C., 'Parameter identification for linear time varying dynamic processes', Proc. IEE, Vol. 117, pp. 2025-29, 1970.
- [33] Diamessis, J.E., 'A new method of determining the parameters of physical systems', Proc. IEEE, Vol. 53, pp. 205-206, 1965.
- [34] Tzafestas, S.G., 'Walsh series approach to lumped and distributed system identification', Journal of the Franklin Inst., Vol. 305, pp. 199-220, 1978.
- [35] Shridhar, M., Miller, W.C., and Balatoni, N., 'System identification using cubic splines', Proceedings of the 7th annual Princeton conf. on Information and Systems, pp. 149-154, 1973.
- [36] Bellman, R., and Roth, R.S., 'The use of splines with unknown end points in the identification of systems', J. Math. Anal. and Appl., 34, pp. 26-30, 1971.
- [37] Sinha, N.K., and Zhou, Q., 'Identification of continuous-time systems from sampled data: a comparison of 3 direct methods', 7th IFAC Symposium on Identification and System parameter estimation, pp. 1575-78, 1985.
- [38] Paley, R.E.A.C., 'A remarkable series of orthogonal functions', Proc. London Math. Soc., 2nd series, Vol. 34, pp. 241-279, 1932.
- [39] Chen, C.F., and Hsiao, C.H., 'Time domain synthesis via Walsh functions', Proc. IEE, Vol. 122, pp. 565-570, 1975.
- [40] Chen, C.F., Tsay, Y.T., and Wu, T.T., 'Walsh operational matrices for fractional calculus and their application to distributed systems', Jnl. of the Franklin Inst., Vol. 303, pp. 267-284, 1977.
- [41] Rao, G.P., and Sivakumar, L., 'System identification via Walsh functions', Proc. IEE, Vol. 122, pp. 1160-1161, 1975.

- [42] Sannuti, P., 'Analysis and synthesis of dynamic systems via block-pulse functions', Proc. IEE, Vol. 124, pp. 569-571, 1977.
- [43] Shieh, L.S., Yeung, C.K., and McInnis, B.C., 'Solutions of state-space equations via block pulse functions', Int. Jnl. of Control, Vol. 28, pp. 283-392, 1978.
- [44] Jan, Y.G. and Wong, K.M. 'Bilinear system identification by block pulse functions', J. of the Franklin Inst., Vol. 312, No. 5, pp. 349-359, November 1981.
- [45] Kwong, C.P. and Chen, C.F., 'The convergence properties of block pulse series', Int. J. Systems Sci., Vol. 12, pp. 745-754, 1981.
- [46] Rao, G.P. and Srinivasan, T., 'Analysis and synthesis of dynamic systems containing time delays via block-pulse functions', Proc. IEE, Vol. 125, pp. 1064-1068, 1978.
- [47] Palanisamy, K.R., and Bhattacharya, D.K., 'System identification via block-pulse functions', Int. Jnl. of Systems Sci., Vol. 12, pp. 643-647, 1981.
- [48] Zhou, Q., and Sinha, N.K., 'System identification using block-pulse functions', Proc. 25th Midwest Symp. on Circuit and Systems, pp. 325-328, 1982.
- [49] Harmuth, H.F., 'Applications of Walsh functions in communications', IEEE Spectrum, pp. 82-91, 1969.
- [50] Huog, J.C., Liu, C.C., and Chou, P.Y., 'An algebraic method for system parameter identification', Proc. 14th Asilomar Conference on Circuits, Systems and Computers, pp. 277-279, 1980.
- [51] Sinha, N.K., and Lastman, G.J., 'Identification of continuous-time multivariable systems from sampled data', Int. J. Control, Vol. 35, pp. 117-126, 1982.
- [52] Prasad, T., and Sinha, N.K., 'Modelling of continuous-time systems from sampled data using trapezoidal pulse functions', Proc. IEEE Conference on System, Man and Cybernetics, Bombay, pp. 427-30, 1983.
- [53] Young, P.C., 'Some observations on instrumental variable methods of time-series analysis', Int. Jnl. Control, 23, pp. 593-612, 1976.
- [54] Sarris, G.N., 'A comparison of six on-line identification algorithms', Automatica, Vol. 10, pp. 69-79, 1974.
- [55] Weinert, H., and Anton, J., 'Canonical forms for multivariable system identification', Proc. of IEEE Conf. on Decision and Control, pp. 37-39, 1972.

- [56] Kwatny, H.G., 'A note on stochastic approximation algorithm in system identification', IEEE Trans. Aut. Control, Vol. AC-17, pp. 571-572, 1972.
- [57] Sen, A., and Sinha, N.K., 'On-line system identification algorithm combining stochastic approximation and pseudoinverse', Automatica, Vol. 11, pp. 425-429, 1975.
- [58] Hsia, T.C., 'On least squares algorithms for system parameters identification', IEEE Trans. Aut. Control, Vol. AC-23, pp. 104-108, 1976.
- [59] Puthenpura, S., and Sinha, N.K., 'A procedure for determining the optimal sampling interval for system identification using a digital computer', Canadian Elec. Eng. Jnl, Vol. 10, pp. 152-157, 1985.
- [60] Sinha, N.K., Lastman, G.J., and Zhou, Q., 'On the choice of sampling interval for the identification of continuous-time systems from sampled data', Proc. 25th Midwest Symp. on Circuit and Systems, pp. 328-332, 1982.
- [61] Young, P.C., 'Parameter estimation for continuous-time models - a survey', 5th IFAC Symp. on Identification and System parameter estimation, Darmstadt, F.R. of G., pp. 17-41, 1979.
- [62] Woodside, C.M., 'Estimation of the order of linear systems', Automatica, Vol. 7, pp. 727-733, 1971.
- [63] Wellstead, P.E., 'An instrumental product moment test for model order estimation', Automatica, Vol. 14, pp. 89-91, 1978.
- [64] Chow, J.C., 'On estimating the orders of an autoregressive moving average process with uncertain observations', IEEE Trans. Aut. Control, AC.17, pp. 707-709, 1972.
- [65] Tse, E., and Weinert, H.L., 'Structure determination and parameter identification for multivariable stochastic linear systems', IEEE Trans. Aut. Control, AC-20, pp. 603-613, 1975.
- [66] Akaike, H., 'Fitting autoregressive models for prediction', Ann. Inst. Statist. Math., 21, pp. 243-247, 1969.
- [67] Akaike, H., 'A new look at statistical model identification', IEEE Trans. AC-19, pp. 716-722, 1974.
- [68] Diekmann, K., and Unbehauen, H., 'Tests for determining the order of canonical models of multivariable systems', Proc. IFAC Symp. on Digital Control, India, pp. 13.1-13.7, 1982.



- [69] Suen, L.C., and Liu, R., 'Determination of the structure of multivariable stochastic linear systems', IEEE Trans. Aut. Control, AC-23, pp. 458-464, 1978.
- [70] Box, G.E.P., and Jenkins, G.M., 'Time series analysis - Forecasting and Control', Holden Day, San Francisco, 1976.
- [71] Young, P., Jakeman, A., and McMurtrie, R., 'An instrumental variable method for model order identification', Automatica, Vol. 16, pp. 281-294, 1980.
- [72] Rissanen, J., 'Basis of invariants and canonical forms for linear dynamic systems', Automatica, Vol. 10, pp. 175-182, 1974.
- [73] Luenberger, D.G., 'Canonical forms for linear multivariable systems', IEEE Trans. Aut. Control, AC-12, pp. 290-293, 1967.
- [74] Glover, K. and Willems, J.C., 'Parameterization of linear dynamical systems - canonical forms and identifiability', IEEE Trans. Aut. Control, AC-19, pp. 667-674, 1974.
- [75] Ljung, L., and Rissanen, J., 'On canonical forms, parameter identifiability and the concept of complexity', Proc. 4th IFAC Symposium on Identification and System Parameter Estimation, Tbilisi, U.S.S.R., pp. 58-69, 1976.
- [76] Beghelli, S., and Guidorzi, R.P., 'Transformations between input-output multistructural models: properties and applications', Int. J. Control, Vol. 37, No. 6, pp. 1385-1400, 1983.
- [77] Lee, R.C.K., 'Optimal estimation, identification and control', MIT Press, Cambridge, Massachusetts, 1964.
- [78] Young, P.C., and Jakeman, A.J., 'Refined instrumental variables methods of recursive time-series analysis', Int. J. Control 29, pp. 1-30, 1979.
- [79] El-Sherief, H., and Sinha, N.K., 'Identification of multivariable systems in the transfer-function matrix forms', Journal of Cybernetics, Vol. 9, pp. 113-125, 1979.
- [80] Abu-El-Magd, Z., and Sinha, N.K., 'Effect of approximate integration on the identification of continuous-time systems from sampled data', Submitted for publication, Canadian Electrical Engineering Journal.
- [81] El-Sherief, H., and Sinha, N.K., 'Determination of the structure of a canonical model for the identification of linear multivariable systems', 5th IFAC Symp. on Identification and System Parameter Estimation, Darmstadt, F.R. of G., pp. 569-574, 1979.

- [82] Van Overbeek, A.J.M., and Ljung, L., 'On-line structure selection for multivariable state space models', Proc. 5th IFAC Symp. on Identification and System Parameter Estimation, Darmstadt, pp. 387-396, 1979.
- [83] Van Overbeek, A.J.M., and Ljung, L., 'On-line structure selection for multivariable state space models', Automatica, Vol. 18, No. 5, pp. 529-543, 1982.
- [84] Abu-El-Magd, Z., and Sinha, N.K., 'On-line structure selection for multivariable continuous-time input-output models', IASTED International Symp. Applied Simulation and Modeling, Vancouver, 1986, accepted for publication.
- [85] Wertz, V., Gevers, M., and Hannan, E.J., 'The determination of optimum structures for the state space representation of multivariable stochastic processes', IEEE Trans. on Aut. Control, AC-27, No. 6, pp. 1200-1211, 1982.
- [86] Guidorzi, R.P., 'Invariants and canonical forms for systems structure and parametric identification', Automatica, Vol. 17, pp. 117-133, 1981.
- [87] Gevers, M., and Wertz, V., 'On the problem of structure selection for the identification of stationary stochastic processes', 6th IFAC Symp. on Identification and System Parameter Estimation, pp. 287-292, 1982.
- [88] Beghelli, S., and Guidorzi, R.P., 'Input-output multistructural models in multivariable systems identification', 6th IFAC Symp. on Identification and System Parameter Estimation, pp. 461-465, 1982.
- [89] Abu-El-Magd, Z.H., and Sinha, N.K., 'Three order determination methods for multivariable continuous-time systems', 27th Midwest Symp. on Circuits and Systems, pp. 820-823, 1984.
- [90] Van Emden, M.H., 'An analysis of complexity', Mathematical Centre Tracts, vol. 35. Amsterdam: Mathematical Centrum, 1971.
- [91] Maklad, M.S. and Nichols, S.T., 'A new approach to model structure discrimination', IEEE Trans. on Systems, Man and Cybernetics, Vol. SMC-10, No. 2, pp. 78-84, 1980.
- [92] Maklad, M.S., 'Complexity and model structure discrimination', Proc. Int. Conf. on Cybernetics and Society, Denver, Colorado, 1978.
- [93] Guidorzi, R.P., 'Multistructural model selection', Proc. 6th European Meetings on Cybernetics and Systems Research, Wien, pp. 135-140, 1982.

- [94] Kulback, S., 'Information Theory and Statistics', New York: Wiley, 1959.
- [95] Shannon, C.E., 'A mathematical theory of communications', Bell System Tech. J., vol 27, pp. 379-423, 623-656, 1948.
- [96] Bartlett, M.S., 'On the theoretical specification and sampling properties of autocorrelated time series', J. Royal Stat. Soc., vol. B8, pp. 27-41, 1946.
- [97] Wolovich, W.A., 'Linear Multivariable Systems, Berlin: Springer-Verlag, 1974.