Linear systems identification using near-Markov methods.

Mohamed. Rafiquzzaman

University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation


https://scholar.uwindsor.ca/etd/2619

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
LINEAR SYSTEM IDENTIFICATION

USING NEAR-MARKOV METHODS

by

MOHAMED RAFIQUZZAMAN

A Dissertation submitted to the Faculty of Graduate Studies through the Department of Electrical Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Windsor, Windsor, Ontario, Canada.
ABSTRACT

This thesis develops a computer-aided identification technique for multidimensional linear systems using near-Markov methods. These systems when excited by wideband Gaussian noise generate state variables which are near-Markov in nature.

Sampled-data information about the state variables is used to characterize the conditional incremental moment functions of the state vector. The coefficient matrices appearing in the state equation description of the system are then identified using these incremental statistics and the system vectors.
ACKNOWLEDGEMENTS.

The author would like to express his appreciation to his supervisor Dr. W. C. Miller for many valuable discussions and constructive criticisms.

The author is indebted to the National Research Council of Canada and the Government of the Province of Ontario for the financial support of the research in the form of scholarships.

Thanks are also due to all those who helped the author in any form, and to Mrs. Shirley Kurpel for typing this thesis.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>CHAPTER I INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 System Identification</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Purpose and Contents</td>
<td>8</td>
</tr>
<tr>
<td>CHAPTER II MARKOV AND NEAR-MARKOV PROCESSES</td>
<td>10</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>10</td>
</tr>
<tr>
<td>2.2 Continuous Markov Processes</td>
<td>12</td>
</tr>
<tr>
<td>2.3 Processes with Independent Increments</td>
<td>13</td>
</tr>
<tr>
<td>2.4 Stochastic Differential Equations</td>
<td>14</td>
</tr>
<tr>
<td>2.5 Ordinary and Stochastic Differential Equations</td>
<td>18</td>
</tr>
<tr>
<td>2.6 The Estimation of a Physical Model</td>
<td>21</td>
</tr>
<tr>
<td>2.7 Summary</td>
<td>24</td>
</tr>
<tr>
<td>CHAPTER III LINEAR SYSTEM IDENTIFICATION USING NEAR-MARKOV METHODS</td>
<td>26</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>26</td>
</tr>
</tbody>
</table>
### CHAPTER III

3.2 Derivation of the Conditional Incremental Moment Functions for Multidimensional Linear Systems ............... 27

3.3 The Relationships of the [A] and [B] matrices to the Conditional Incremental Moment Functions and System Vectors .............. 38

3.4 The Identification Procedure
   (a) Estimation of the Conditional Incremental Moment Functions ............... 50
   (b) Choice of the Sampling Interval .......... 51
   (c) Noise source Properties .......... 52

3.5 (a) An Example of System Identification .......... 57
   (b) A Second Identification Example .......... 86

### CHAPTER IV

DISCUSSION AND RESULTS ............... 94

4.1 Introduction ............... 94

4.2 Discussion of Theoretical Developments ............... 94

4.3 Simulation Studies .......... 95

4.4 Practical Applications .......... 104

### CHAPTER V

CONCLUSIONS .......... 106

### REFERENCES .......... 110
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-1</td>
<td>Geometrical Representation of Equations (3.46) and (3.47)</td>
<td>46</td>
</tr>
<tr>
<td>3-2</td>
<td>Gaussian Distributed Piecewise Constant Noise, its Autocorrelation and Power Spectrum</td>
<td>56</td>
</tr>
<tr>
<td>3-3</td>
<td>R.M.S. Error ( x_1 ) vs. Differencing Interval ( H )</td>
<td>62</td>
</tr>
<tr>
<td>3-4</td>
<td>R.M.S. Error ( x_2 ) vs. Differencing Interval ( H )</td>
<td>63</td>
</tr>
<tr>
<td>3-5</td>
<td>The R.M.S. Error of the Estimate of the Incremental Static of ( x_1(t) ) given as a Function of the Number of Samples used in Making the Estimates</td>
<td>65</td>
</tr>
<tr>
<td>3-6</td>
<td>The R.M.S. Error of the Estimate of the Incremental Statistic of ( x_2(t) ) Given as a Function of the Number of Samples Used in Making the Estimates</td>
<td>66</td>
</tr>
<tr>
<td>3-7</td>
<td>The effect of the Total Number of Samples on the Identification of ( a_k ) Coefficients</td>
<td>68</td>
</tr>
<tr>
<td>3-8</td>
<td>The effect of the Total Number of Samples on the Identification of ( b_k ) Coefficients</td>
<td>69</td>
</tr>
<tr>
<td>3-9</td>
<td>The R.M.S. Error Associated with the Identification of the ( a_k ) Coefficients Given as a Function of the Total Number of Samples Used</td>
<td>72</td>
</tr>
<tr>
<td>3-10</td>
<td>The R.M.S. Error Associated with the Identification of the ( b_k ) Coefficients Given as a Function of the Total Number of Samples Used</td>
<td>73</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3-11</td>
<td>The Effect of the Number of Increments on the Identification of $a_{lk}$ Coefficients</td>
<td>74</td>
</tr>
<tr>
<td>3-12</td>
<td>The R.M.S. Error Associated with the Estimation of $a_{lk}$ Coefficients Plotted as a Function of the Number of Increments used</td>
<td>75</td>
</tr>
<tr>
<td>3-13</td>
<td>The R.M.S. Error Associated with the Identification of $a_{lk}$ Coefficients Given as a Function of the Sampling Interval</td>
<td>76</td>
</tr>
<tr>
<td>3-14</td>
<td>The R.M.S. Error Associated with the Identification of $b_{lk}$ Coefficients Given as a Function of the Sampling Interval</td>
<td>77</td>
</tr>
<tr>
<td>3-15</td>
<td>The Effect of Changing $\frac{\tau \text{Cor}(x_1)}{\tau \text{Cor}(z_1)}$ on the Identification of $a_{lk}$ Coefficients</td>
<td>84</td>
</tr>
<tr>
<td>3-16</td>
<td>The R.M.S. Error Associated with the Identification of $a_{lk}$ Coefficients plotted as a function of $\frac{\tau \text{Cor}(x_1)}{\tau \text{Cor}(z_1)}$</td>
<td>85</td>
</tr>
<tr>
<td>3-17</td>
<td>The Effect of Changing $\frac{\tau \text{Cor}(x_2)}{\tau \text{Cor}(z_1)}$ on the Identification of $a_{lk}$ Coefficients</td>
<td>87</td>
</tr>
<tr>
<td>3-18</td>
<td>The R.M.S. Error Associated with the Identification of $a_{lk}$ Coefficients Given as a Function of $\frac{\tau \text{Cor}(x_2)}{\tau \text{Cor}(z_1)}$</td>
<td>88</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3-19</td>
<td>The Effect of the Total Number of Samples on the Identification of $a_k$ Coefficients</td>
<td>90</td>
</tr>
<tr>
<td>3-20</td>
<td>The R.M.S. Error Associated with the Identification of $a_k$ Coefficients Plotted as a Function of the Total Number of Samples Used</td>
<td>91</td>
</tr>
<tr>
<td>3-21</td>
<td>The Effect of the Total number of Samples on the Identification of $b_k$ Coefficients</td>
<td>92</td>
</tr>
<tr>
<td>3-22</td>
<td>The R.M.S. Error Associated with the Identification of $b_k$ Coefficients Plotted as a Function of the Total Number of Samples Used</td>
<td>93</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

1.1 System Identification

In many engineering applications it is desirable to be able to model physical systems using measured data.

System identification has been the interest of many researchers. In order to show the varied nature of the identification techniques available, a brief survey of some of the techniques will be given in this section.

Saridis and Stein (1968) considered the parameter identification problem in the theory of adaptive control systems from the point of view of stochastic approximation. In their technique, a generalized algorithm for on-line identification of a stochastic linear discrete-time system using noisy input and output measurements is developed and shown to converge in the mean-square sense. The algorithm requires knowledge of the noise variances involved. It is shown that this requirement is a disadvantage associated with on-line identification schemes based on minimum mean-square-error criteria. Two off-line identification schemes are presented which utilize measurements obtained from repeated runs of the system's transient response and do not require explicit knowledge of the noise variances. These algorithms
converge with probability one to the true parameter values.

Real-time identification and control of linear discrete systems with Gauss-Markov random parameters and conditioned quadratic cost function are considered by Farison, Graham and Shelton (1967). In this technique, algorithms for explicit calculation of the identification and control are given for systems with perfect state measurement and are illustrated by example applications. The control strategy generated by the cost function is an identification-adaptive controller which is continually revised as new data are received. The identification equations develop an explicit model of the system which is available for other purposes, if desired and can be used apart from the control context.

In most practical applications, dynamic systems can be modelled by differential equations with prescribed forms but unknown parameters. Hsia and Vimalvanich (1969) developed an analogue method for identifying these unknown parameters. The method is based on the learning model concept and is capable of performing on-line identification in the presence of noise. Single variable and multivariable linear systems can be identified by this technique.

Djorovic and Bingulac (1972) made an attempt to solve the identification problem of general multivariable linear
continuous dynamic systems - determination of order and calculation of parameters of a minimal realization of the system. The suggested identification procedure starts from normal input-output data corrupted by zero-mean measurement noises. It is necessary for the system to be initially at rest. Particular attention was paid to the computational aspects of the identification problem. An illustrative example exhibited satisfactory results. The algorithm has still to be tested for efficiency in more realistic situations.

The problem of identifying the plant dynamics is an important part of adaptive control. Kumar and Sridhar (1964) dealt with the application of the quasi-linearization method to identify the differential equation governing the dynamical behaviour of a linear or a nonlinear stationary plant. It has been shown how the method is well-tailored to the use of a digital computer. Irrespective of whether the plant to be identified is linear or non-linear, the multipoint boundary value problem to be solved is nonlinear and the method remains exactly the same. Very little is known about the existence and uniqueness of solutions of nonlinear multipoint boundary value problem. The numerical method suggested in this paper seems to hold promise as a useful tool. Undoubtedly much experimentation and experience is necessary before this method
can be extended to the identification of time-varying plants.

Pazdersa and Pottinger (1969) applied Liapunov design
techniques to the identification problem. A scheme for
identifying slowly varying linear systems with accessible
state vector has been derived. The identification method is
stable and convergence of the adjustable parameters to the
correct values is guaranteed whenever the input is general
enough to disturb the system in a manner which permits
complete identification. The 'general enough input' problem
can be alleviated by adding a small noise signal to the
regular system input. The Liapunov derived identification
scheme was simulated and found very effective. Rapid and
accurate convergence of the adjustable parameters was noted.
The rapid and accurate convergence evidenced during this
study indicates that satisfactory tracking of more general
time-varying linear systems should be possible by the
identification scheme presented in their paper. Also,
addition of a small signal to the regular system input should
permit identification of an approximate linear model for certain
classes of non-linear systems. The approximate linear model
would vary depending upon the present operating point of the
nonlinear system.
Saridis and Lobbia (1972) considered the parameter-adaptive self-organizing control of linear discrete-time systems by designing dynamic feedback controllers which depend on the estimates of the parameters provided by an appropriate identifier. Two stochastic approximation algorithms for consistent identification of feedback systems are investigated and a condition of identification of identifiability is presented. Then, two controllers, one based on 'overall' and another based on 'per-interval' optimization, both depending on the output of the identifier, are discussed and their evaluations relative to the optimal are compared in illustrative examples.

Furuta and Ha (1972) proposed a procedure for identification of the weighting function of a system. Their procedure uses a new type of criterion involving a Priori information about the weighting function as well as the integral of the squared response error. The optimal weighting function applied in this criterion can be determined by solving the optimal tracking problem, provided that the output of a linear, free dynamical system is supplied to the plant as an input. The criterion function is also valid from the estimation theory. The computational algorithm of
the procedure is also presented. Numerical computations of several examples have shown that the proposed procedure gives satisfactory results even for the short-time observed data.

The problem of identification of linear dynamic systems and adaptive Kalman filtering are considered by Mehra (1971). Kalman gave a set of recursive equations for estimating the state of a linear dynamic system. However the Kalman filter requires a knowledge of all the system and noise parameters. Here it is assumed that all these parameters are unknown and therefore must be identified before use in the Kalman filter. A correlation technique which identifies a system in its canonical form is presented. The estimates are shown to be asymptotically normal, unbiased, and consistent. The scheme is capable of being implemented on-line and can be used in conjunction with the Kalman filter. A technique for more efficient estimation by using higher order correlations is also given. A recursive technique is given to determine the order of the system when the dimension of the system is unknown. The results are first derived for stationary processes and are then extended to non-stationary processes which are stationary in the qth increment. An application
of the results to a practical problem is also presented. Lee (1960) has developed an identification technique for linear systems where white Gaussian noise excitation and crosscorrelation techniques are used. Later on, Lee and Schetzen (1962) have extended this concept to nonlinear systems. Their representation of a nonlinear system is based on an orthogonalized form of a Volterra integral series. The terms in the series are determined by applying a white Gaussian noise input simultaneously to the nonlinear system and to multidimensional delay circuit. By estimating the crosscorrelation functions involving the outputs from the nonlinear system and the delay circuit it is possible to determine the functions appearing in the Volterra series.

A brief survey of identification of lumped parameter system is given so far. Several methods are also available for identifying the distributed parameter systems. These techniques normally involve the reduction of the distributed system to lumped parameter form or to a system of linear algebraic equations; identification is obtained by solving these equations by numerical or analog methods. However a detailed treatise of the techniques available for the identification of distributed parameter systems will not be given here.
1.2 **Purpose and Contents**

Identification techniques involving random processes can be advantageous in that they have the potential to be implemented while the system is operating.

In previous stochastic identification schemes as shown by Lee (1960), the application of input-output cross-correlation theory has been used to characterize a linear system in terms of its impulse response by exciting the system with noise. Usually the noise used has a Gaussian probability density function and a power density spectrum that is considerably wider than the bandwidth of the system under investigation. This type of noise is referred to as Gaussian "White" noise.

For linear systems described by state equations of the form \( [X] = [A][X] + [B][U] \), an identification technique is developed in this thesis that does not require a cross-correlation operation. These systems when excited by wideband Gaussian noise produce an output which is near-Markov in nature. It is possible to estimate the incremental statistics associated with the near-Markov processes from sampled-data information about the state variables. These estimates and the sampled data information of the system
vectors can be used to compute the \([A]\) and \([B]\) matrices. Thus, this thesis develops a technique whereby multivariable linear systems can be identified by exploiting their near-Markov rather than correlation properties.

As the theoretical basis for the identification technique developed in this thesis is intimately associated with diffusion processes, this topic is treated in Chapter II which describes the relationship between near-Markov physical processes defined by ordinary differential equations and diffusion processes. It also describes the interpretations of a stochastic integral derived by Ito (1951) and Stratonovich (1966). Next, the estimation of a physical or stochastic model for a physical process is considered.

In Chapter III, mathematical model for the identification technique has been derived. Simulation results, estimation, and identification of two second-order linear systems which are taken as examples for demonstrating the identification technique are also presented.

A discussion of the theoretical and experimental results is treated in Chapter IV. This chapter also contains a description of some of the practical applications of the identification technique.

The conclusions that can be reached as a result of this research are given in Chapter V.
CHAPTER II

MARKOV AND NEAR-MARKOV PROCESSES

2.1 Introduction

As the theoretical basis for the identification scheme is intimately associated with diffusion processes, this topic is treated in this chapter.

Many effective mathematical methods have been developed within the framework of Markov Process theory. There are thus advantages in the ability to obtain a diffusion model for a random process. The statistics of a diffusion process or continuous Markov Process, are described by the corresponding incremental moment functions.

If for a given continuous Markov Process the incremental moment functions are known, then the defining stochastic differential equations, in the sense of Ito, may be deduced. These incremental moment functions are exactly specified by equations (2.8) and (2.10) when the evaluating time interval approaches zero. For a finite time interval, the incremental moment functions may be approximated by utilizing equations (2.29) and (2.30).

In spite of the advantages of a diffusion approximation most random processes encountered in practice have an inherent
smoothness about them, whereas, diffusion processes do not have this property. Therefore, any physical process can be at most near-Markov. As such, the incremental properties of a near-Markov physical process do not exactly specify the statistics of the process.

It is possible to estimate quantities analogous to the incremental moment functions of a diffusion process from sampled-data information taken over a finite sampling interval from a near-Markov process. These quantities are referred to as the conditional incremental statistics. For a physical process, there will be a range of sampling intervals over which these incremental statistics can be evaluated and still be of useful accuracy.

Once these conditional incremental statistics have been computed, the stochastic differential equation, in the sense of Ito, can be formulated. The solution to this stochastic differential equation is a continuous Markov Process whose statistical properties are similar to those of the near-Markov Process being modelled.

In engineering work, diffusion processes are usually only of academic interest. Since it is often desirable to model the process on a computer, it is necessary to obtain
an ordinary differential equation whose solution is a near-Markov Process with statistical properties similar to the near-Markov Process being modelled. In order to estimate the specific functions appearing in the ordinary differential equation chosen as the model, it is necessary to relate the estimates of the incremental properties of the process to those functions.

2.2 Continuous Markov Processes

A continuous Markov Process is a special type of stochastic process and is often called a process without aftereffect. According to Papoulis (1965), a process is Markovian in nature, if the past has no influence on the statistics of the future under the condition that the present is known. This can be stated in terms of conditional probabilities. For any \( n \) and \( t_1 > t_2 > \ldots > t_n \), the condition

\[
P \{ x(t_1) \leq x_1 \mid x(t_2) = x_2, \ldots, x(t_n) = x_n \}
\]

\[
= P \{ x(t_1) \leq x_1 \mid x(t_2) = x_2 \}
\]

is true for a Markovian Process. Their statistics are completely determined if one knows merely their second order probability density function.

A continuous Markov Process is commonly called a diffusion process and is not differentiable. Any stochastic process encountered in nature has a certain inherent smoothness and thus is at best only near-Markov.
2.3 **Processes with independent increments**

If \( y(t) \) is a process with independent increments and 
\( y(t_1), y(t_2), \ldots, y(t_n); t_1 < t_2 < \ldots < t_n \ (n \geq 3) \) are 
realizations of the process at times \( t_1, t_2, \ldots, t_n \), the 
differences

\[
[y(t_2)-y(t_1)], [y(t_3)-y(t_2)], \ldots, [y(t_n)-y(t_{n-1})] \ldots (2.1)
\]

are mutually independent. If the distribution \( y(t+\Delta t)-y(t) \) 
depends only on \( \Delta t \), a process with independent increments 
is said to have stationary increments also.

Since processes with independent increments are 
specified by the distribution of the increments, the random 
variable of the process of interest is really the difference 
\( y(t_n) - y(t_{n-1}) \). To avoid this situation, \( y(0) \) is defined 
with probability one.

The Brownian motion process has independent increments 
and properties that \( y(t+\Delta t) - y(t) \) is real and normally 
distributed with

\[
E\{y(t+\Delta t) - y(t)\} = 0 \quad \ldots (2.2)
\]

\[
E\{[y(t+\Delta t) - y(t)]^2\} = \sigma^2 \Delta t \quad \ldots (2.3)
\]

where \( \sigma^2 \) is the variance parameter. \( E\{\ldots\} \) indicates the 
expectation of the quantities enclosed in brackets.
When the variance parameter of the Brownian motion process equals one, the process often is referred to as a Wiener Process. The Wiener Process is especially important in the theory of stochastic processes as a stochastic differential equation which uses this process has been defined by Ito (1951). Its solution is a continuous Markov Process.

Another feature of the Brownian motion process is that it is not differentiable. That is, \( \frac{dy(t)}{dt} \) cannot be finite for all \( t \) as Doob (1942) pointed out, since the variance of the increment is proportional to \( \Delta t \), the standard deviation is of the order of \( (\Delta t)^{1/2} \) and hence \( y(t + \Delta t) - y(t) \sim (\Delta t)^{1/2} \) \( \cdots \) (2.4)

This result is also important in the discussion of Ito's stochastic differential equation occurring in the next section.

2.4 Stochastic Differential Equations

Ito (1951) has formulated a stochastic differential equation which defines a continuous Markov Process. The scalar form of this equation is given by equation (2.5).

\[
dx(t) = f[x(t), t] \, dt + g[x(t), t] \, dy(t) \quad \cdots \quad (2.5)
\]

where \( y(t) \) is a Wiener Process with the following properties.

\[
E \{ y(t_2) - y(t_1) \} = 0 \quad \cdots \quad (2.6)
\]

\[
E \{ (y(t_2) - y(t_1))^2 \} = |t_2 - t_1| \quad \cdots \quad (2.6)
\]
It is tempting to try to solve equation (2.5) by dividing through by \( dt \) and using conventional calculus techniques; however, as shown in the previous section \( \frac{dy(t)}{dt} \) is not finite. Equation (2.5) can be solved by defining a special stochastic integral as outlined by Miller (1969). Since the solution to equation (2.5) is a continuous Markov Process, for certain conditions imposed upon the functions \( f(x(t), t) \) and \( g(x(t), t) \), it is of interest to solve for the incremental properties of the process defined by equation (2.5). The incremental moment functions of \( x(t) \) may be determined by taking the expectations of both sides of equation (2.5) then dividing by \( \Delta t \) and taking the limit as \( \Delta t \) approaches zero. Performing these operations gives

\[
E \{ x(t + \Delta t) - x(t) \} = E \{ f[x(t), t] \, dt \} + g[x(t), t] E \{ y(t + \Delta t) - y(t) \} \quad \ldots (2.7)
\]

\[
\lim_{\Delta t \to 0} E \left\{ \frac{x(t + \Delta t) - x(t)}{\Delta t} \mid x(t) = \xi \right\} = f[\xi, t] \quad \ldots (2.8)
\]

Equation (2.8) is evident in light of equation (2.4). Squaring \( dx(t) \) and performing the same operations as in equation (2.7) gives

\[
E \{ dx^2(t) \} = E \{(f[x(t), t]dt)^2 + 2(f[x(t), t]dt)(g[x(t), t]dy(t)) + g^2[x(t), t](dy(t))^2 \} \quad \ldots (2.9)
\]
\[ \lim_{\Delta t \to 0} \mathbb{E} \left\{ \frac{\left[ x(t + \Delta t) - x(t) \right]^2}{\Delta t} \ \bigg| \ x(t) = \tau \right\} \]

\[ = \lim_{\Delta t \to 0} \mathbb{E} \left\{ \frac{g^2[\tau, t] \Delta t}{\Delta t} \right\} = g^2[\tau, t] \quad \ldots \ (2.10) \]

The conditional first and second order incremental moment functions are defined by equations (2.8) and (2.10) respectively. The i-th order conditional incremental moment function of the \( x(t) \) process is defined by

\[ k_i(x, t) = \lim_{\Delta t \to 0} \mathbb{E} \left\{ \frac{[x(t + \Delta t) - x(t)]^i}{\Delta t} \ \bigg| \ x(t) \right\} \quad \ldots \ (2.11) \]

For the \( x(t) \) process defined by Ito's stochastic differential equation,

\[ k_i(x, t) = 0; \quad i \geq 3 \quad \ldots \ (2.12) \]

as is true for a continuous Markov process or diffusion process.

Kolmogorov's equations (2.13) and (2.14) give a relationship between the conditional probability density and incremental moment function of a continuous Markov Process.
\[
\frac{\partial P(x,t|x_0,t_0)}{\partial t_0} = -f(x_0,t_0) \frac{\partial P(x,t|x_0,t_0)}{\partial x_0} - g^2(x_0,t_0) \frac{\partial^2 P(x,t|x_0,t_0)}{\partial x_0^2} \tag{2.13}
\]

\[
\frac{\partial P(x,t|x_0,t_0)}{\partial t} = -\frac{\partial}{\partial x} \left[ f(x,t) P(x,t|x_0,t_0) \right] + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[ g^2(x,t) P(x,t|x_0,t_0) \right] \tag{2.14}
\]

The second Kolmogorov equation is also known as the Fokker-Planck equation after the physicists who first developed it. The conditional probability density function is shown in equation (2.15).

\[
P(x,t|x_0,t_0) = \frac{P(x,x_0,t,t_0)}{P(x_0,t_0)} \tag{2.15}
\]

A derivation of equations (2.13) and (2.14) is given by Gnedenko (1962).

In the stochastic differential equation formulated by Ito a conventional derivative is not defined. It has been shown by Wonham (1965) that Ito's stochastic differential equation cannot be treated as if it was an ordinary
differential equation. In an attempt to simplify the relationship between a stochastic differential equation and an ordinary differential equation Stratonovich (1966) has defined a new representation for a stochastic integral. The stochastic differential equation derived by Stratonovich is based on a symmetrical representation for a stochastic integral. This equation has the advantage that it can be manipulated, to a certain extent, as if it is an ordinary differential equation. An associated property of the Stratonovich stochastic differential equation is that it is similar in form to the ordinary differential equation that defines the corresponding near-Markov physical process. Miller (1969) has given a thorough discussion of Stratonovich's stochastic differential equation.

2.5 Ordinary and Stochastic Differential Equations

One might be tempted to formulate an ordinary differential equation by dividing through in equation (2.5) by $dt$ and replacing $dy(t)/dt$ by a realizable Gaussian band-limited noise source $z(t)$. This result is shown in equation (2.16).

$$dx(t) = f(x,t) + g(x,t) z(t) \quad \ldots \quad (2.16)$$

If in equation (2.16), $dy(t)/dt$ were differentiable, then
the autocorrelation of \( z(t) \) would be a delta function which has a flat power spectrum. However, as mentioned before, \( dy(t)/dt \) is not differentiable. Wong and Zakai (1965) have shown that as \( z(t) \) approaches white Gaussian noise, in general, the solution to equation (2.16) does not converge to that of the diffusion process defined by Ito's stochastic differential equation.

The relationship between the stochastic and ordinary differential equation is much more complex. Consider the first-order ordinary differential equation given in equation (2.17).

\[
\frac{dX(t)}{dt} = F(X) + G(X) z(t) \quad \ldots \quad (2.17)
\]

\( X(t) \) is a near-Markov Process and \( z(t) \) is a stationary Gaussian distributed noise source with zero mean value and having a very short correlation time. The correlation time of the noise and the intensity coefficient, which is equal to unity in equation (2.17), are defined by equations (2.19) and (2.20) respectively. Both equations involve the auto-correlation function of the noise. This is given in equation (2.18) for a stationary ergodic noise source.
\[
R_{zz}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{0}^{T} z(t + \tau)z(t)\,dt
\]

\[
= E \left\{ z(t + \tau)z(t) \right\} \quad \ldots \quad (2.18)
\]

\[
\tau_{\text{cor}}(z) = \frac{1}{\sigma_z^2} \int_{0}^{\infty} \left| R_{zz}(\tau) \right| \,d\tau; \quad \sigma_z^2 = R_{zz}(0) \quad (2.19)
\]

\[
C_i(z) = \int_{-\infty}^{\infty} R_{zz}(\tau)\,d\tau \quad \ldots \quad (2.20)
\]

Stratonovich (1963) has developed equations which relate the functions \( F(X) \) and \( G(X) \) appearing in the ordinary differential equation (2.17) to the functions \( K_1(X) \) and \( K_2(X) \) as given in equations (2.21) and (2.22).

\[
K_1(X) \approx F(X) + \frac{1}{2} \frac{\partial G(X)}{\partial X} G(X) \quad \ldots \quad (2.21)
\]

\[
K_2(X) \approx G^2(X) \quad \ldots \quad (2.22)
\]

These equations are only valid if \( C_i(z) \) is equal to unity. \( K_1(X) \) and \( K_2(X) \), computed from an ordinary differential equation, are functions analogous to the incremental moment functions of a diffusion process and are called the conditional incremental statistics.
The form of the stochastic differential equation, in the sense of Ito, that defines a diffusion process with corresponding incremental moment functions \( k_1(x) \) and \( k_2(x) \) is given in equation (2.25).

\[
k_1(x) = F(x) + \frac{1}{2} \frac{\partial G(x)}{\partial x} G(x)
\]

\[
k_2(x) = G^2(x)
\]

\[
dx(t) = \left[ F(x) + \frac{1}{2} \frac{\partial G(x)}{\partial x} G(x) \right] dt + G(x) dy(t). \quad (2.25)
\]

Stratonovich has thus been able to show the relationship between a stochastic and a first-order ordinary differential equation. He also gives the conditions under which the solution to equation (2.25) has statistical properties sufficiently close to the near-Markov Process to form a satisfactory model. A detailed treatise of this portion of Stratonovich's work is given by Miller (1969).

2.6 The Estimation of a Physical Model

The theory so far has dealt with finding the appropriate diffusion model for a near-Markov Process defined by a given ordinary differential equation. However, the problem is to formulate a model in terms of a first-order
ordinary differential equation based on estimates taken from a physical near-Markov Process.

If a stationary near-Markov Process is encountered that can be modelled by a first-order ordinary differential equation such as (2.26)

\[
\frac{dX(t)}{dt} = F(X) + G(X)z(t)
\]  

(2.26)

then the results of section 2.5 are useful. The functions \( F(X) \) and \( G(X) \) can be computed from a knowledge of the conditional incremental statistics \( K_1(X) \) and \( K_2(X) \). This can be shown by rearranging equations (2.21) and (2.22) as follows.

\[
F(X) = K_1(X) - \frac{1}{4} \frac{\partial K_2(X)}{\partial X} 
\]  

(2.27)

\[
G(X) = \sqrt{K_2(X)}
\]  

(2.28)

For these latter two equations to be valid, \( z(t) \) in equation (2.26) must have a small correlation time and a unit intensity coefficient.

In a realistic situation the functions \( K_1(X) \) and \( K_2(X) \) would have to be estimated from sampled-data information about the physical process. A digital computer would be required to compute the estimates from the
measured data. The values of \( F(x) \) and \( G(x) \) could be computed from these estimates when the statistical properties of \( z(t) \) are known.

Since a physical process is only at best near-Markov it is a problem to determine the optimum sampling interval. In the academic case where the near-Markov Process is defined by a known equation defining the transformation of a noise source the optimum sampling interval would be slightly greater than the correlation time of the noise source. In practice, however, any choice of sampling interval must be determined from measurements made on the physical process alone. Since a correlation time for the physical process itself can be computed from measured data the optimum sampling interval might be given as some fraction of this value. The problem is in essence to determine the smallest interval over which the physical process has Markovian-like properties.

Assuming that a model of the form given in equation (2.26) is estimated using measured data there does not seem to be any way in which the validity of this model can be verified directly. That is, one must actually compare the statistical properties of interest of both the process being modelled and the output defined by the model. Once the specific form
of a model of this nature has been estimated, the
Stratonovich convergence criterion can be evaluated for
it with the aid of a digital computer. It would thus be
possible to tell if the estimated model defined a near-
Markov Process. However, the statistical equivalence of
the process being modelled and the output from the model
cannot be established by this evaluation of the convergence
criterion.

The manner in which the estimates of the incremental
statistics of the near-Markov physical Process are used to
formulate the model depends upon how the model is to be
implemented. The method is different depending upon
whether analogue or digital oriented model is desired.
In the digital case the choice of the numerical integration
also influence how the estimates are used.

2.7 Summary

Ito has defined a stochastic differential equation whose
solution is a continuous Markov Process.

When a physically realizable process with near-Markov
properties is considered, one is concerned with the
infinitesimal properties of equations (2.29), (2.30), and
(2.31) rather than the incremental moment functions of
equations (2.8) and (2.10).
\[ E \{ x(t + \Delta t) - x(t) | x(t) \} = f(x(t), t) \Delta t + \eta(\Delta t) \quad \ldots (2.29) \]

\[ E \{ [x(t + \Delta t) - x(t)]^2 | x(t) \} = g^2(x(t), t) \Delta t + \eta(\Delta t) \quad \ldots (2.30) \]

\[ E \{ [x(t + \Delta t) - x(t)]^n | x(t) \} = \eta(\Delta t) \quad \ldots (2.31) \]

A suitable model for a near-Markov Physical Process must be based on an ordinary differential equation rather than a stochastic one.

As the identification technique in this thesis is closely associated with near-Markov Processes, a detailed treatise of this topic is given in this chapter.
CHAPTER III

LINEAR SYSTEM IDENTIFICATION USING
NEAR-MARKOV METHODS

3.1 Introduction

This chapter describes a computer-aided technique using near-Markov methods for the identification of multidimensional linear systems described by state equatons of the form

\[
\frac{d}{dt} [x(t)] = [A][x(t)] + [B][u(t)]
\]  \hspace{1cm} (3.1)

where \( [x(t)] \) is the pX1 state vector, \( u(t) \) is the qX1 input vector, and, \( [A] \) and \( [B] \) are pXp and pXq matrices respectively.

As has been pointed out earlier, no process in practice has strict Markov properties. Therefore its incremental properties do not strictly characterize the actual process. However, if the process is sufficiently near-Markov, its conditional incremental statistics help in its identification. In the identification technique described in this chapter, a near-Markov process is generated using equation (3.1) with wideband Gaussian noise \( (U) \). The conditional incremental statistics of the state variables are computed from their sampled-data information. The coefficient matrices \( [A] \) and
[B] in equation (3.1) are then identified from the conditional incremental statistics and the system vectors.

A mathematical derivation of the identification technique is given in sections 3.2 and 3.3. Various aspects associated with the identification procedure are discussed in section 3.4 and simulation studies on two second-order examples are treated in section 3.5(a) and 3.5(b).

3.2 Derivation of the Conditional Incremental Moment Functions for Multidimensional Linear Systems

Equation (3.1) can be represented by a system of equations as follows

\[
\begin{align*}
\dot{x}_1 &= \sum_{k=1}^{p} (a_{1k} x_k + b_{1k} u_k) = \lambda F_1[x, U(t)] \\
&\ldots \\
\dot{x}_p &= \sum_{k=1}^{p} (a_{pk} x_k + b_{pk} u_k) = \lambda F_p[x, U(t)]
\end{align*}
\]  

(3.2)

where, \( \lambda \) is a small parameter. In equation (3.2) if \( u_1(t) \ldots u_p(t) \) are independent wideband Gaussian noise sources with zero mean and unit intensity coefficients, the state variables \((x_1, x_2, \ldots, x_p)\) are near-Markov in nature.

Choosing the initial values

\[
x_\ell(t_0) = x_{\ell0} \quad [\ell = 1, \ldots, p]
\]  

(3.3)
we write the solution $x_{\ell}(1 \leq \ell \leq p)$ of the system (3.2) as an expansion

$$x_{\ell}(t) - x_{\ell}^{0} = \mathbf{I}_{\ell}(x_{0}) + \sum_{m=1}^{p} \lambda \int_{t_{0}}^{t} \int_{t_{0}}^{t_{1}} \mathbf{F}_{\ell}[x_{0}, u(t_{1})] \, dt_{1} \, dt_{2} \, \frac{\partial F}{\partial x_{m}}[x_{0}, u(t_{1})] \int_{t_{0}}^{t} \mathbf{F}_{m}[x_{0}, u(t_{2})] \, dt_{2} \, (3.4)$$

The increments $x_{1} - x_{10}, \ldots, x_{p} - x_{p0}$ have the joint characteristic function

$$E\left[\exp \left\{ j \sum_{\ell} u_{\ell}(x_{\ell} - x_{\ell}^{0}) \right\} \right] = 1 + \sum_{\ell=1}^{p} ju_{\ell}E[I_{\ell}(x_{0})] + \frac{1}{2} \sum_{\ell, m=1}^{p} ju_{\ell}u_{m}E[I_{\ell}(x_{0})I_{m}(x_{0})] + \cdots \cdots \cdots (3.5)$$

Taking the inverse fourier transform of equation (3.5), one
obtains the multidimensional probability density function

\[ p(x_1, \ldots, x_p, t \mid x_{10}, \ldots, x_{p0}, t_0) \]

\[(1 + L) \delta(x_1 - x_{10}) \cdots \delta(x_p - x_{p0}) \xi (3.6)\]

where

\[ L = \sum_{\ell=1}^{P} \left( -\frac{\partial}{\partial x_{\ell}} \right) E[I_{\ell}(x)] \]

\[ + \frac{1}{2} \sum_{\ell, m=1}^{P} \left( -\frac{\partial}{\partial x_{\ell}} \right) \left( -\frac{\partial}{\partial x_{m}} \right) E[I_{\ell}(x) I_{m}(x)] \]

\[ + \cdots \cdots \cdots (3.7) \]

It follows at once from equation (3.6) that the probability density \( p(x_1, \ldots, x_p, t \mid x_{10}, \ldots, x_{p0}, t_0) \) satisfies the equation

\[ \frac{\partial p}{\partial t} (x_1, \ldots, x_p, t \mid x_{10}, \ldots, x_{p0}, t_0) \]

\[ = \frac{\partial L}{\partial t} (1 + L)^{-1} p(x_1, \ldots, x_p, t \mid x_{10}, \ldots, x_{p0}, t_0) \quad (3.8) \]

We must now calculate the operator \( \dot{L}(1 + L)^{-1} \). Substituting equation (3.4) into equation (3.7), we obtain

\[ L = -\hbar \sum_{\ell=1}^{P} \frac{\partial}{\partial x_{\ell}} \left[ \int_{t_0}^{t} E \left[ F_{\ell}(x, U(t_1)) \right] dt_1 \right] \]

\[ + \hbar \int_{t_0}^{t} dt_1 \sum_{\ell=1}^{P} E \left[ \frac{\partial F_{\ell}(x, U(t_1))}{\partial x_m} F_m(x, U(t_2)) \right] dt_2 \]
\[ + \frac{\lambda^2}{2} \sum_{\ell, m=1}^P \int_0^t \int_0^t E \left[ F_{\ell}(X, U(t_1)) F_m(X, U(t_2)) \right] dt_1 dt_2 + o(\lambda^3) \]  

so that

\[ L = - \sum_{\ell=1}^P \frac{\partial}{\partial x_{\ell}} \left\{ F_{\ell}(X, U(t)) + \int_0^t \sum_{m=1}^P E \left[ \frac{\partial F_{\ell}}{\partial x_m}(X, U(t)) \right] F_m(X, U(t')) dt' \right\} \]

\[ + \lambda^2 \sum_{\ell, m=1}^P \frac{\partial^2}{\partial x_{\ell} \partial x_m} \int_0^t E \left[ F_{\ell}(X, U(t)) F_m(X, U(t')) \right] dt' + o(\lambda^3) \]  

... (3.10)

The operator \( L (\lambda + L)^{-1} \) differs from equation (3.10) only by having the correlation functions

\[ \text{Cov} \left[ \frac{\partial F_{\ell}}{\partial x_m}, F_{m'} \right], \text{Cov} \left[ F_{\ell}, F_{m'} \right] \]  
in place of the averages

\[ E \left[ \frac{\partial F_{\ell}}{\partial x_m} F_{m'} \right], E[F_{\ell} F_{m'}]. \]

As a result, equation (3.8) becomes
\[
\frac{\partial P}{\partial t} (x_1, \ldots, x_p, t \mid x_{10}, \ldots, x_{p0}, t_0) \\
= -\lambda \sum_{\ell=1}^{P} \frac{\partial}{\partial x_{\ell}} \left\{ \left( E[F_{\ell}] + \lambda \sum_{m=1}^{P} \int_{t_0}^{0} \text{Cov} \left[ \frac{\partial F_{\ell}}{\partial x_{m}}, F_{m\tau} \right] d\tau \right) \right\} \\
+ \lambda^2 \sum_{\ell, m=1}^{P} \frac{\partial^2}{\partial x_{\ell} \partial x_{m}} \left\{ \int_{t_0}^{0} \text{Cov} \left[ F_{\ell\tau}, F_{m\tau} \right] d\tau \right\} \\
\cdot P(x_1, \ldots, x_p, t \mid x_{10}, \ldots, x_{p0}, t_0) \right\}
\]

\[ \ldots \quad (3.11) \]

In the derivation of equation (3.11) the terms of order \( \lambda^3 \) and higher have been neglected. The conditions under which these terms can be justifiably neglected forms the basis of Stratonovich's convergence criterion. The condition that must be satisfied for equation (3.11) to be valid and the state variables \((x_1, \ldots, x_p)\) to be sufficiently near-Markov that a satisfactory diffusion model can be found is that the expression

\[ \tau_{\text{cor}}(U) \cdot \left| \frac{\partial F(X, U(t))}{\partial X} \right|_{\text{max}} \ll 1.0 \]

must be valid. The expression is Stratonovich's convergence criterion. If the above equation is valid, the appropriate
diffusion model for $X(t)$ can be found by comparing equation (3.11) with a Fokker-Planck equation.

The multidimensional Fokker-Planck equation is given by

$$
\frac{\partial P(x_1, \ldots, x_p, t \mid x_{10}, \ldots, x_{p0}, t_0)}{\partial t} = -\sum_{\ell=1}^{P} \frac{\partial}{\partial x_\ell} \left[ K_\ell(x_1, \ldots, x_p) \ P(x_1, \ldots, x_p, t \mid x_{10}, \ldots, x_{p0}, t_0) \right]
$$

$$
+ \frac{1}{2} \sum_{\ell, m=1}^{P} \frac{\partial^2}{\partial x_\ell \partial x_m} \left[ K_{\ell m}(x_1, \ldots, x_p) \ P(x_1, \ldots, x_p, t \mid x_{10}, \ldots, x_{p0}, t_0) \right] \quad \ldots (3.12)
$$

where the incremental moment functions $K_\ell(x_1, \ldots, x_p)$ and $K_{\ell m}(x_1, \ldots, x_p)$ are given by

$$
K_\ell(x_1, \ldots, x_p) = \lim_{\Delta t \to 0} E \left[ \frac{x_\ell(t + \Delta t) - x_\ell(t)}{\Delta t} \mid x_\ell(t) \right] \quad (3.13)
$$

$$
K_{\ell m}(x_1, \ldots, x_p) = \lim_{\Delta t \to 0} E \left[ \frac{\{x_\ell(t + \Delta t) - x_\ell(t)\} \{x_m(t + \Delta t) - x_m(t)\}}{\Delta t} \mid x_\ell(t), x_m(t) \right] \quad \ldots (3.14)
$$

Comparing equation (3.11) with equation (3.12), we find that
\[ K_p(x) = \lambda E[F_p] + \lambda^2 \sum_{m=1}^{P} \mathop{\int}_{0}^{t} \text{Cov} \left[ \frac{\partial F_p}{\partial x_m}, F_m \right] d\tau. \quad (3.15) \]

\[ K_{p,m}(x) = 2 \lambda^2 \mathop{\int}_{0}^{t} \text{Cov} \left[ F_p, F_{m,\tau} \right] d\tau \quad \ldots \quad (3.16) \]

Now as \( t \to 0 \to \infty \),

\[ K_p(x) \to \lambda E[F_p] + \lambda^2 \sum_{m=1}^{P} \mathop{\int}_{\infty}^{0} \text{Cov} \left[ \frac{\partial F_p}{\partial x_m}, F_m \right] d\tau \quad \ldots \quad (3.17) \]

\[ K_{p,m}(x) \to 2 \lambda^2 \mathop{\int}_{-\infty}^{0} \text{Cov} \left[ F_p, F_{m,\tau} \right] d\tau \quad \ldots \quad (3.18) \]

From equation (3.2), we have

\[ \lambda F_p = \sum_{k=1}^{P} \left( a_{pk} x_k + b_{pk} u_k \right) \quad \ldots \quad (3.19) \]

Taking expected values of both the sides of equation (3.19), we have

\[ \lambda E[F_p] = \sum_{k=1}^{P} E \left( a_{pk} x_k + b_{pk} u_k \right) = \sum_{k=1}^{P} a_{pk} x_k \quad (3.20) \]

Let us now calculate

\[ \text{Cov} \left[ \frac{\partial F_p}{\partial x_m}, F_{m,\tau} \right] \triangleq E \left[ \frac{\partial F_p}{\partial x_m} F_{m,\tau} \right] - E \left[ \frac{\partial F_p}{\partial x_m} \right] E \left[ F_{m,\tau} \right] \quad (3.21) \]
differentiating both the sides of equation (3.19) with respect to $x_m$ we obtain

$$\frac{\partial F}{\partial x_m} = \frac{1}{\lambda} \sum_{k=1}^{P} \left[ a_{\ell k} \frac{\partial x_k}{\partial x_m} + \frac{\partial b_{\ell k}}{\partial x_m} u_k \right]$$

$$= \frac{1}{\lambda} \sum_{k=1}^{P} a_{\ell k} \frac{\partial x_k}{\partial x_m} \quad (3.22)$$

Again

$$F_{m\tau} = \frac{1}{\lambda} \left[ \sum_{k=1}^{P} a_{mk} x_k + b_{mk} u_k(t+\tau) \right] \quad (3.23)$$

Using equations (3.21), (3.22) and (3.23) we have

$$\text{Cov} \left[ \frac{\partial E^2}{\partial x_m}, F_{m\tau} \right]$$

$$= \frac{1}{\lambda^2} E \left[ \left( \sum_{k=1}^{P} a_{\ell k} \frac{\partial x_k}{\partial x_m} \right) \left( \sum_{k=1}^{P} a_{mk} x_k + b_{mk} u_k(t+\tau) \right) \right]$$

$$= \frac{1}{\lambda^2} \left[ a_{\ell k} \sum_{k=1}^{P} \frac{\partial x_k}{\partial x_m} \right] E \left[ \sum_{k=1}^{P} \left( a_{mk} x_k + b_{mk} u_k(t+\tau) \right) \right]$$

\begin{align*}
\frac{1}{\lambda^2} \left[ \sum_{k=1}^{p} a_{jk} \frac{\partial x_k}{\partial x_m} a_{mk} x_k \right] \\
+ b_{mk} a_{jk} \frac{\partial x_k}{\partial x_m} E \left[ u_k(t+\tau) \right] \\
- \frac{1}{\lambda^2} \left[ \sum_{k=1}^{p} a_{jk} \frac{\partial x_k}{\partial x_m} \right] \left[ \sum_{k=1}^{p} a_{mk} x_k \right] \\
+ b_{mk} E \left\{ u_k(t+\tau) \right\} \right] \\
= \frac{1}{\lambda^2} \left[ \sum_{k=1}^{p} a_{jk} a_{mk} x_k \frac{\partial x_k}{\partial x_m} \right] - \frac{1}{\lambda^2} \left[ \sum_{k=1}^{p} a_{jk} a_{mk} x_k \frac{\partial x_k}{\partial x_m} \right] = 0 \\
(3.24)
\end{align*}
\[
\text{Cov}[F_{\ell}, F_{m, \tau}] = \frac{1}{\lambda^2} \mathbb{E} \left[ \sum_{k=1}^{P} (a_{\ell k} x_{k} + b_{\ell k} u_{k}(t)) \right] \cdot \mathbb{E} \left[ \sum_{k=1}^{P} (a_{mk} x_{k} + b_{mk} u_{k}(t + \tau)) \right] \\
= \frac{1}{\lambda^2} \left[ \sum_{k=1}^{P} (a_{\ell k} a_{mk} x_{k} + b_{\ell k} b_{mk} \mathbb{E} \{ u_{k}(t) \}) \right] \\
+ a_{\ell k} b_{mk} \mathbb{E} \{ u_{k}(t + \tau) \} \\
+ b_{\ell k} b_{mk} \mathbb{E} \{ u_{k}(t) u_{k}(t + \tau) \} \\
- \frac{1}{\lambda^2} \left[ \sum_{k=1}^{P} a_{\ell k} a_{mk} x_{k} \right] \\
= \frac{1}{\lambda^2} \left[ \sum_{k=1}^{P} (a_{\ell k} a_{mk} x_{k} + b_{\ell k} b_{mk} \mathbb{E} \{ u_{k}(t) u_{k}(t + \tau) \}) \right] \\
- a_{\ell k} a_{mk} x_{k} \right] \\
= \frac{1}{\lambda^2} \sum_{k=1}^{P} b_{\ell k} b_{mk} \mathbb{E} \left[ u_{k}(t) u_{k}(t + \tau) \right] \quad (3.25)
\]
From equations (3.17), (3.20) and (3.24),

\[
K_\ell(x) = \sum_{k=1}^{p} a_{\ell k} x_k
\]  
(3.26)

From equations (3.18) and (3.25),

\[
K_{\ell m}(x) = \int_{-\infty}^{\infty} \sum_{k=1}^{p} b_{\ell k} b_{m k} \mathbb{E}[u_k(t)u_k(t + \tau)] \, d\tau
\]

\[
= \sum_{k=1}^{p} b_{\ell k} b_{m k}
\]  
(3.27)

since

\[
\int_{-\infty}^{\infty} \mathbb{E}[u_k(t)u_k(t + \tau)] \, d\tau = 1
\]

for unit intensity coefficient.
3.3 The Relationships of the [A] and [B] Matrices to the Conditional Incremental Moment Functions and System Vectors.

For developing the relationships involving identification of $a_{jk}$ coefficients, equations (3.13) and (3.26) will be used. A second-order system will be treated first and the result will then be generalized.

For a second-order linear system, the state equations can be written as

$$\dot{x}_1 = a_{11}x_1 + a_{12}x_2 + b_{11}u_1 + b_{12}u_2$$  \hspace{1cm} (3.28)

$$\dot{x}_2 = a_{21}x_1 + a_{22}x_2 + b_{21}u_1 + b_{22}u_2$$  \hspace{1cm} (3.29)

Using equations (3.26), (3.28) and (3.29), the theoretical incremental moment functions can be written as

$$K_{1\text{ Theo}}(x_1, x_2) = a_{11}x_1 + a_{12}x_2$$  \hspace{1cm} (3.30)

$$K_{2\text{ Theo}}(x_1, x_2) = a_{21}x_1 + a_{22}x_2$$  \hspace{1cm} (3.31)

From equation (3.13), $K_1(x_1, x_2)$ and $K_2(x_1, x_2)$ can be estimated using the following relationships.
\[ K_{1\,\text{EST}}(x_1, x_2) = E \left[ \frac{x_1(t + \Delta t) - x_1(t)}{\Delta t} \bigg| x_1(t) \right] \] (3.32)

\[ K_{2\,\text{EST}}(x_1, x_2) = E \left[ \frac{x_2(t + \Delta t) - x_2(t)}{\Delta t} \bigg| x_2(t) \right] \] (3.33)

Let \( K_{1\,\text{EST}}(n), K_{1\,\text{THEO}}(n), K_{2\,\text{EST}}(n), K_{2\,\text{THEO}}(n), x_1(n) \) and \( x_2(n) \), \([n = 1, 2, \ldots, N]\) be the samples of \( K_{1\,\text{EST}}(x_1, x_2)\), \( K_{1\,\text{THEO}}(x_1, x_2), K_{2\,\text{EST}}(x_1, x_2), K_{2\,\text{THEO}}(x_1, x_2), x_1(t), \) and \( x_2(t) \) respectively, and \( N \) be the total number of samples. The coefficients \( a_{11}, a_{12}, a_{21} \) and \( a_{22} \) can be identified by forming the mean square errors between \( K_{1\,\text{THEO}}(n) \) and \( K_{1\,\text{EST}}(n) \), and, between \( K_{2\,\text{THEO}}(n) \) and \( K_{2\,\text{EST}}(n) \) and then applying least square technique.

Mean square error between \( K_{1\,\text{EST}}(n) \) and \( K_{1\,\text{THEO}}(n) \) with \( N \) samples is given by

\[ E_1 = \frac{1}{N} \sum_{n=1}^{N} \left[ K_{1\,\text{EST}}(n) - K_{1\,\text{THEO}}(n) \right]^2 \]

\[ = \frac{1}{N} \sum_{n=1}^{N} \left[ K_{1\,\text{EST}}(n) - a_{11}x_1(n) - a_{12}x_2(n) \right]^2 \] (3.34)

Differentiating equation (3.34) with respect to \( a_{11} \) and equating to zero, we have,
\[
\frac{\partial E_1}{\partial a_{11}} = \frac{1}{N} \sum_{n=1}^{N} 2 \left[ K_1^{\text{EST}(n)} - a_{11}x_1(n) - a_{12}x_2(n) \right] \left[ -x_1(n) \right] \\
= 0
\]

Therefore

\[
\sum_{n=1}^{N} \left[ - K_1^{\text{EST}(n)} x_1(n) + a_{11}x_1^2(n) + a_{12}x_1(n)x_2(n) \right] = 0
\]

\[
a_{11} \sum_{n=1}^{N} x_1^2(n) + a_{12} \sum_{n=1}^{N} x_1(n)x_2(n) = \sum_{n=1}^{N} K_1^{\text{EST}(n)}x_1(n)
\]

(3.35)

Similarly differentiating equation (3.34) with respect to \( a_{12} \) and equating to zero, we have,

\[
a_{11} \sum_{n=1}^{N} x_1(n)x_2(n) + a_{12} \sum_{n=1}^{N} x_2^2(n) = \sum_{n=1}^{N} K_1^{\text{EST}(n)}x_2(n)
\]

(3.36)

Again forming the mean square error between \( K_2^{\text{EST}(n)} \) and \( K_2^{\text{THEO}(n)} \) and then differentiating with respect to \( a_{21} \) and \( a_{22} \) respectively and then equating to zero, we have,
\[ a_{21} \sum_{n=1}^{N} x_1^2(n) + a_{22} \sum_{n=1}^{N} x_1(n)x_2(n) = \sum_{n=1}^{N} k_{2EST}(n)x_1(n) \]  
\[ (3.37) \]

\[ a_{21} \sum_{n=1}^{N} x_1(n)x_2(n) + a_{22} \sum_{n=1}^{N} x_2^2(n) = \sum_{n=1}^{N} k_{2EST}(n)x_2(n) \]  
\[ (3.38) \]

Writing equations (3.35), (3.36), (3.37) and (3.38) in matrix form, we have,

\[
\begin{bmatrix}
  a_{11} & a_{12} \\
  \sum_{n=1}^{N} x_1^2(n) & \sum_{n=1}^{N} x_1(n)x_2(n)
\end{bmatrix}
\begin{bmatrix}
  \sum_{n=1}^{N} k_{1EST}(n)x_1(n) \\
  \sum_{n=1}^{N} k_{1EST}(n)x_2(n)
\end{bmatrix}
\]

\[
\begin{bmatrix}
  a_{21} & a_{22} \\
  \sum_{n=1}^{N} x_1(n)x_2(n) & \sum_{n=1}^{N} x_2^2(n)
\end{bmatrix}
\begin{bmatrix}
  \sum_{n=1}^{N} k_{2EST}(n)x_1(n) \\
  \sum_{n=1}^{N} k_{2EST}(n)x_2(n)
\end{bmatrix}
\]

Therefore,

\[
\begin{bmatrix}
  a_{11} & a_{12} \\
  \sum_{n=1}^{N} k_{1EST}(n)x_1(n) & \sum_{n=1}^{N} k_{1EST}(n)x_2(n)
\end{bmatrix}
\begin{bmatrix}
  \sum_{n=1}^{N} x_1^2(n) \\
  \sum_{n=1}^{N} x_1(n)x_2(n)
\end{bmatrix}
\]

\[
\begin{bmatrix}
  a_{21} & a_{22} \\
  \sum_{n=1}^{N} k_{2EST}(n)x_1(n) & \sum_{n=1}^{N} k_{2EST}(n)x_2(n)
\end{bmatrix}
\begin{bmatrix}
  \sum_{n=1}^{N} x_1(n)x_2(n) \\
  \sum_{n=1}^{N} x_2^2(n)
\end{bmatrix}
\]

\[ (3.39) \]
In general, for pth order system, we have,

\[
\begin{bmatrix}
a_{11} & \cdots & a_{1p} \\
\vdots & \ddots & \vdots \\
a_{p1} & \cdots & a_{pp}
\end{bmatrix}
\begin{bmatrix}
\sum_{n=1}^{N} K_{1EST}(n)x_1(n) - \sum_{n=1}^{N} K_{1EST}(n)x_p(n) \\
\sum_{n=1}^{N} K_{pEST}(n)x_1(n) - \sum_{n=1}^{N} K_{pEST}(n)x_p(n)
\end{bmatrix}
= 
\begin{bmatrix}
\sum_{n=1}^{N} x_1^2(n) - \sum_{n=1}^{N} x_1(n)x_p(n) \\
\sum_{n=1}^{N} x_p(n)x_1(n) - \sum_{n=1}^{N} x_p^2(n)
\end{bmatrix}^{-1}
\]

Equation (3.40) can be used to identify [A] matrix.

Next the development of the relationships involving identification of [B] matrix will be considered.

For a second-order system equation (3.27) can be written as

\[
\begin{align*}
K_{11}(x_1, x_2) &= b_{11}^2 + b_{12}^2 \\
K_{12}(x_1, x_2) &= b_{11}b_{21} + b_{22}b_{12} \\
K_{21}(x_1, x_2) &= b_{11}b_{21} + b_{22}b_{22} \\
K_{22}(x_1, x_2) &= b_{21}^2 + b_{22}^2
\end{align*}
\]

(3.41)
The incremental moment functions $K_{11}(x_1, x_2), K_{12}(x_1, x_2), K_{21}(x_1, x_2)$ and $K_{22}(x_1, x_2)$ can be estimated from sampled-data information of the state variables $x_1$ and $x_2$ using equation (3.14). Equation (3.41) can be written in a matrix form as

$$[B][B]^T = [K]$$  \hspace{1cm} (3.42)

where

$$[B] = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

and $[B]^T$ is the transpose of $[B]$.

$$[K] = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}$$

Equation (3.42) is difficult to solve because it has four unknown parameters and three equations. But if some constraint is imposed on $[B]$, the solution becomes simpler. For example, assuming $[B]$ is a symmetric matrix and positive definite, the equation (3.42) becomes

$$[B]^2 = [K]$$

$$[B] = [K]^{1/2}$$ \hspace{1cm} (3.43)

$[K]^{1/2}$ can be evaluated from $[K]$ by reducing $[K]$ to
diagonal form using $[P][K][P]^{-1}$ where $[P]$ is an eigenvector matrix and then computing $[B]$ as

$$[B] = [P]^{-1} [K]^{1/2} [P]$$

Equation (3.44) can be used for identifying $[B]$, assuming $[B]$ to be symmetric and positive definite. But it is not worthwhile to identify $[B]$ with such constraint, because the nature of $[B]$ is unknown before identification. Therefore, some other method of identifying $[B]$, which is of general nature, is to be found. Efforts have been made to identify $[B]$ of general nature using geometrical methods.

For example, let

$$\bar{x}_1 = b_{11}$$
$$\bar{y}_1 = b_{12}$$
$$\bar{x}_2 = b_{21}$$
$$\bar{y}_2 = b_{22}$$

Therefore, the set of equations given by equation (3.41) becomes

$$K_{11} = \bar{x}_1^2 + \bar{y}_1^2$$
$$K_{12} = \bar{x}_1 \bar{x}_2 + \bar{y}_1 \bar{y}_2$$
$$K_{21} = \bar{x}_1 \bar{x}_2 + \bar{y}_1 \bar{y}_2$$
$$K_{22} = \bar{x}_2^2 + \bar{y}_2^2$$

(3.45)
Let \( \vec{D} = \vec{x}_1 \vec{a}_x + \vec{y}_1 \vec{a}_y \) \\
\( \vec{c} = \vec{x}_2 \vec{a}_x + \vec{y}_2 \vec{a}_y \)

Therefore,

\[
|D| = \sqrt{\vec{x}_1^2 + \vec{y}_1^2} = \sqrt{K_{11}} \tag{3.46}
\]

\[
|c| = \sqrt{\vec{x}_2^2 + \vec{y}_2^2} = \sqrt{K_{22}} \tag{3.47}
\]

Taking dot products between the vectors \( \vec{D} \) and \( \vec{c} \) we have,

\[
\vec{D} \cdot \vec{c} = (\vec{x}_1 \vec{a}_x + \vec{y}_1 \vec{a}_y) \cdot (\vec{x}_2 \vec{a}_x + \vec{y}_2 \vec{a}_y)
\]

\[
= \vec{x}_1 \vec{x}_2 + \vec{y}_1 \vec{y}_2 = K_{12} = K_{21} \tag{3.48}
\]

Therefore \( K_{12} = \vec{x}_1 \vec{x}_2 + \vec{y}_1 \vec{y}_2 \)

\[
= |D| |c| \cos \theta = \sqrt{K_{11}} \sqrt{K_{22}} \tag{3.49}
\]

Equations (3.46) and (3.47) are circles with radii \( K_{11} \) and \( K_{22} \) respectively, as shown in figure (3-1).
FIGURE 3-1: GEOMETRICAL REPRESENTATION OF EQUATIONS

(3.46) and (3.47)

From equation (3.49),

$$\cos \theta = \frac{K_{12}}{\sqrt{K_{11}} \sqrt{K_{22}}}$$  \hspace{1cm} (3.50)

Let \( \bar{x}_1 = \bar{x}_3 \)

\( \bar{y}_1 = \frac{\bar{x}_3}{\sqrt{K_{11} - \bar{x}_3^2}} \)

From figure (3-1)

$$\cos \theta_1 = \frac{\bar{x}_3}{\sqrt{K_{11}}} \quad \sin \theta_1 = \sqrt{\frac{K_{11} - \bar{x}_3^2}{K_{11}}}$$

\( \sin (\theta + \theta_1) = \frac{\bar{x}_2}{\sqrt{K_{11}}} \)
Therefore \[ a_2 = (\sin \theta) \sqrt{x_3} \pm \sqrt{K_{11} - \bar{x}_3^2} \cos \theta \] (3.51)

Again \[ \cos (\theta + \theta_1) = \frac{a_1}{\sqrt{K_{11}}} \] and
\[ \cos \theta \cos \theta_1 - \sin \theta \sin \theta_1 = \frac{a_1}{\sqrt{K_{11}}} \]

Therefore \[ a_1 = \bar{x}_3 \cos \theta \pm \sqrt{K_{11} - \bar{x}_3^2} \sin \theta \] (3.52)

where \[ \cos \theta = \frac{K_{12}}{\sqrt{K_{11}} \sqrt{K_{22}}} \] (3.53)

\[ \sin \theta = \pm \sqrt{1 - \cos^2 \theta} \] (3.54)

Again \[ \bar{x}_2 = \frac{\sqrt{K_{22}}}{\sqrt{K_{11}}} a_1 \]
\[ = \frac{\sqrt{K_{22}}}{\sqrt{K_{11}}} \left[ \bar{x}_3 \cos \theta \pm \sqrt{K_{11} - \bar{x}_3^2} \sin \theta \right] \] (3.55)

\[ \bar{y}_2 = \frac{\sqrt{K_{22}}}{\sqrt{K_{11}}} \left[ \bar{x}_3 \sin \theta \pm \sqrt{K_{11} - \bar{x}_3^2} \cos \theta \right] \] (3.56)

Therefore, \[ [B] = \begin{bmatrix} \sqrt{\frac{K_{22}}{K_{11}}} \left( \bar{x}_3 \cos \theta \pm \sqrt{K_{11} - \bar{x}_3^2} \sin \theta \right) \\ \sqrt{\frac{K_{22}}{K_{11}}} \left( \bar{x}_3 \sin \theta \pm \sqrt{K_{11} - \bar{x}_3^2} \cos \theta \right) \end{bmatrix} \] (3.57)
where $\cos \theta$ and $\sin \theta$ can be obtained using equations (3.53) and (3.54).

Equation (3.57) can be used for identifying $[B]$. But a problem arises because we have to have a guess on $b_{11}$ initially and also to choose the proper signed values of $b_{12}$ and $\sin \theta$. For high order systems, identification of $[B]$ using geometrical methods becomes cumbersome. Therefore, attempts have been made to develop some other technique for identifying $[B]$ matrix.

The state equations represented by equation (3.1) can be written as

$$
\dot{x}_k = \sum_{k=1}^{p} \left( a_{lk} x_k + b_{lk} u_k \right)
$$

\hspace{1cm} \text{[} l = 1, \ldots, p \text{]}

(3.58)

For a second-order system, equation (3.58) can be written as

$$
\dot{x}_1 = a_{11} x_1 + a_{12} x_2 + b_{11} u_1 + b_{12} u_2
$$

(3.59)

$$
\dot{x}_2 = a_{21} x_1 + a_{22} x_2 + b_{21} u_1 + b_{22} u_2
$$

(3.60)

Multiplying both the sides of equation (3.59) by $u_1$ and taking expected values of both the sides we have,
\[ E[\dot{x}_1 u_1] = a_{11} E[u_1] + a_{12} E[u_2] \]
\[ + b_{11} E[u_1^2] + b_{12} E[u_1 u_2] \]  
\((3.61)\)

Since \(u_1\) and \(u_2\) have zero mean values and they are independent, we have,
\[ E[u_1] = E[u_2] = 0 \]  
\((3.62)\)

\[ E[u_1 u_2] = E[u_1] E[u_2] = 0 \]  
\((3.63)\)

Applying the conditions of equations \((3.62)\) and \((3.63)\) to equation \((3.61)\) and after some manipulations, one obtains

\[ b_{11} E[u_1^2] = E[\dot{x}_1 u_1] \]

Therefore
\[ b_{11} = \frac{E[\dot{x}_1 u_1]}{E[u_1^2]} = \frac{E[\dot{x}_1 u_1]}{\sigma_1^2} \]  
\((3.64)\)

Where \(\sigma_1^2\) is the variance of \(u_1\). Again multiplying both the sides of equation \((3.59)\) by \(u_2\) and proceeding in the same manner, we have,

\[ b_{12} = \frac{E[\dot{x}_1 u_2]}{\sigma_2^2} \]  
\((3.65)\)

where \(\sigma_2^2\) is the variance of \(u_2\). Similarly multiplying both the sides of equation \((3.60)\) by \(u_1\) and \(u_2\) respectively and proceeding in the same way, one obtains,

\[ b_{21} = \frac{E[\dot{x}_2 u_1]}{\sigma_1^2} \]  
\((3.66)\)
\[ b_{22} = \frac{E[\dot{x}_2 u_2]}{\sigma^2} \quad (3.67) \]

In general, for \( p \)th order system,

\[ b_{\lambda k} = \frac{E[\dot{x}_\lambda u_k]}{\sigma^2_k} \quad (3.68) \]

where \( \lambda = 1, \ldots, p \)

\[ k = 1, \ldots, p \]

and \( \sigma_k^2 \) is the variance of \( u_k \).

Equation (3.69) can be used to identify \([B]\) matrix.

3.4 The Identification Procedure

The purpose of this section is to discuss the various aspects of the modelling procedure associated with the identification technique.

In order to identify the system, a model for \( X(t) \) is to be formulated first, based on the estimates of the conditional incremental statistics of the state variables. A procedure for computing the conditional incremental statistics is given in section 3.4(a).

The proper choice of the sampling interval is very important to evaluate the conditional incremental statistics of the state variables. This topic is discussed in section 3.4(b).

Noise source properties play an important role
in the simulation of near-Markov processes and therefore, this is treated in section 3.4(c).

3.4(a) Estimation of the Conditional Incremental Moment Functions

From equation (3.13)

$$\kappa_\ell (X) = \lim_{\Delta t \to 0} \mathbb{E} \left\{ \frac{x_\ell [(t + \Delta t)] - x_\ell (t)}{\Delta t} \right\} x_\ell (t) \quad (3.69)$$

The above relationship has been used to compute the estimate of the conditional first-order statistic, \( \kappa_\ell (X) \).

Two array of storage, each with, say 1000 locations, are allocated for computation of each state variable statistic; these arrays are all initially set to zeros. The range between the maximum and minimum values of the state variable of interest is then divided up into 1000 equal intervals. Each of the 1000 intervals is associated with one of the storage locations in each array. When the state variable is sampled, its value is within the range of one of the intervals. The storage location in the first array which corresponds to this interval has its value increased by one.

A second sample is taken \( \Delta t \) seconds later and the derivative of the state variable is calculated. This is
added to the contents of the storage location in the second array which is determined, as before, by the magnitude of the state variable sample. The second array is used to sum the values of the derivatives of the state variable. This procedure is continued until say, 1000 samples are taken.

The contents of the storage locations of the second array are then divided by the contents of the corresponding storage locations of the first array, except for, if the contents of the storage locations of either array are zero. The results obtained in the second array are the samples of the estimated conditional incremental moment function of the state variable.

3.4(b) Choice of the Sampling Interval

The proper choice of the sampling interval is very important to evaluate the conditional incremental statistics of the system. The process being only near-Markov, the choice of too small a sampling interval makes the resulting samples highly correlated. Since the near-Markov Process can be considered as a transformation of a realizable white noise with a short correlation time, the smallest sampling interval should not be smaller than the noise correlation time.
When a physically realizable process with near-Markov Properties is considered one is concerned with the following equations:

\[ E \left\{ x(t + \Delta t) - x(t) \mid x(t) \right\} = f \left\{ x(t), t \right\} \Delta t + \eta(\Delta t) \]  
(3.70)

\[ E \left\{ [x(t + \Delta t) - x(t)]^2 \mid x(t) \right\} = g^2 \left\{ x(t), t \right\} \Delta t + \eta(\Delta t) \]  
(3.71)

\[ E \left\{ [x(t + \Delta t) - x(t)]^n \mid x(t) \right\} = \eta(\Delta t) \]  
(3.72)

The above set of equations show that, the larger \( \Delta t \) is, the larger is the error. The Process correlation time sets the upper limit for the sampling interval. The sampling interval, therefore, lies between the noise and the process correlation times.

It would be advantageous if the optimum sampling interval could be given as a function of the correlation time of the physical process being modelled. This seems difficult to do, however, as the ratio of the correlation time of the physical process to that of the noise source is not known. Usually the best results would be obtained when the sampling interval is slightly larger than the correlation time of the noise source.
3.4(c) Noise Source Properties

If a physical model is to be simulated on a digital or analogue computer a suitable noise source must be available. In practice the correlation time of a noise source associated with the generation of a near-Markov Process will be quite short. In the case of digital computer simulations this requirement can be met by using a pseudo-random number generator.

A pseudo-random number generator of the congruential type described by Chambers (1967) is used in the digital simulations described in section 3.5. This algorithm produces Gaussian distributed numbers that are uncorrelated and have a specified mean and variance.

When an ordinary differential equation defining a physical process is solved on a digital computer one obtains an approximate solution at integral multiples of the differencing interval, $\Delta t$, used. The optimum size of this interval or increment depends on the type of equation being solved, the numerical integration formula used, and the accuracy desired. In this thesis the digital simulations are performed in such a manner that one pseudo-random number is required in the determination of each discrete value of
the physical process. As far as the integration formula is concerned the pseudo-random numbers are held constant over the integration interval. This means that the effective properties of the pseudo-random number sequence are the same as those of a piecewise constant noise source. The form of this type of noise source is a pulse train consisting of constant width pulses whose amplitudes are Gaussian distributed.

If \( x(t) \) is a piecewise constant noise source then

\[
z(t) = z(nH); \quad nH < t \leq (n + 1) H
\]

(3.73)

Where \( H \) is the differencing interval used in the digital simulations.

Since \( H \) is also the effective period over which the noise is held constant the following relationships can be written. If the mean value of the noise is zero and the variance is \( \sigma_z^2 \) then the intensity coefficient is

\[
I_c(z) = \frac{\sigma_z^2}{4H}
\]

(3.74)

and the correlation time is

\[
\tau_{Cor}(z) = \frac{H}{2}
\]

(3.75)

Thus when a piecewise constant noise source with
Figure 3-2: Gaussian distributed piecewise constant noise, its autocorrelation and power spectrum.
unit intensity coefficient is desired the variance, \( \sigma_z^2 \), is chosen as \( \sigma_z^2 = (H)^{-1} \). The autocorrelation function of a piecewise constant noise source is given by Korn (1966) as

\[
R_{zz}(\tau) = \sigma_z^2 \left(1 - \frac{|\tau|}{H}\right); \quad |\tau| < H
\]  

(3.76)

A plot of this triangular function along with a typical sample function of piecewise constant noise is shown in figure (3-2). If it is desirable to estimate this function using the uncorrelated pseudo-random numbers as data it is necessary to repeat the numbers an arbitrary number of times before performing the autocorrelation operation. The repetition of numbers gives the effect of the noise being held constant over a time interval \( H \).

3.5(a) An Example of System Identification

In order to verify the identification procedure for linear systems, which is the goal of this investigation, a second-order example is considered which is described by

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
-2 & -1 \\
2 & -5
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
1 & 0 \\
0 & -2
\end{bmatrix} \begin{bmatrix}
u_1 \\
u_2
\end{bmatrix}
\]  

(3.77)
Equation (3.77) is simulated on an IBM 360/65 computer using fourth-order Runge-Kutta method. For obtaining accurate sampled data information of the state variables $x_1$ and $x_2$, choice of the differencing interval ($H$) is very important. For this purpose, the following procedure is followed.

From equation (3.1),

$$[X(t)] = [A][X(t)] + [B][U(t)]$$

taking Laplace transform of both the sides of the above equation

$$S[X(s)] = [A][X(s)] + [B][U(s)] \quad (3.78)$$

$$(S[I] - [A])[X(s)] = [B][U(s)]$$

Therefore $[X(s)] = [S-I][A]^{-1}[B][U(s)] \quad (3.79)$

where $S$ is the Laplace variable, $[I]$ is the unit matrix.

$$[H(s)] = [S-I][A]^{-1}[B] \quad (3.80)$$

where $[H(s)]$ is the transfer function matrix.

In our case $[A] = \begin{bmatrix} -2 & -1 \\ 5 & 0 \end{bmatrix}$ \quad (3.81)

$[B] = \begin{bmatrix} -3 & -5 \\ 0 & 0 \end{bmatrix}$ \quad (3.82)
Substituting the \([A]\) and \([B]\) matrices from equations (3.81) and (3.82) into equation (3.80), we have

\[
[H(S)] = \begin{bmatrix}
S + 2 & +1 \\
-2 & S+5
\end{bmatrix}^{-1} \begin{bmatrix}
1 \\
0
\end{bmatrix}
\]

\[
= \frac{1}{s^2 + 7s + 12} \begin{bmatrix}
S+5 & -1 \\
2 & S+2
\end{bmatrix} \begin{bmatrix}
1 \\
0
\end{bmatrix}
\]

\[
= \frac{1}{s^2 + 7s + 12} \begin{bmatrix}
S+5 & 2 \\
2 & -2s-4
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\frac{S + 5}{s^2 + 7s + 12} & \frac{2}{s^2 + 7s + 12} \\
\frac{2}{s^2 + 7s + 12} & \frac{-2s - 4}{s^2 + 7s + 12}
\end{bmatrix}
\quad (3.83)
\]

Now \(X(S) = H(S) U(S)\) \quad (3.84)

If \(u_1\) and \(u_2\) are unit impulses, then \(u_1(S) = 1\) and \(u_2(S) = 1\).

Therefore
\[
\begin{bmatrix}
  x_1(s) \\
  x_2(s)
\end{bmatrix} =
\begin{bmatrix}
  \frac{s + 5}{s^2 + 7s + 12} & \frac{2}{s^2 + 7s + 12} \\
  \frac{-2}{s^2 + 7s + 12} & \frac{-2s - 4}{s^2 + 7s + 12}
\end{bmatrix}
\begin{bmatrix}
  1 \\
  1
\end{bmatrix}
\] (3.85)

From where,
\[x_1(s) = \frac{s + 7}{(s + 3)(s + 4)}\] (3.86)
\[x_2(s) = \frac{-2s - 2}{(s + 3)(s + 4)}\] (3.87)

Taking inverse Laplace transforms of equation (3.86) and (3.87) the theoretical values of \(x_1(t)\) and \(x_2(t)\) are given by
\[x_1^{\text{THEO}}(t) = 4e^{-3t} - 3e^{-4t}\] (3.88)
\[x_2^{\text{THEO}}(t) = 4e^{-3t} - 6e^{-4t}\] (3.89)

The state variables \(x_1\) and \(x_2\) are directly estimated from the equation (3.77) using fourth-order Runge-Kutta method and the sampled-data information of \(x_1^{\text{EST}}(n\Delta t)\) and \(x_2^{\text{EST}}(n\Delta t)\) are obtained. Then RMS error \((x_1)\) and RMS error \((x_2)\) are estimated from the following relationships.
\[
\text{RMS Error } (x_1) = \sqrt{\frac{\sum_{n=0}^{N-1} (x_{1\text{THEO}}(n\Delta t) - x_{1\text{EST}}(n\Delta t))^2}{N}}
\]  
(3.90)

\[
\text{RMS Error } (x_2) = \sqrt{\frac{\sum_{n=0}^{N-1} (x_{2\text{THEO}}(n\Delta t) - x_{2\text{EST}}(n\Delta t))^2}{N}}
\]  
(3.91)

RMS Error \( (x_1) \) and RMS Error \( (x_2) \) are calculated for different values of the differencing interval, \( H \). One thousand samples of \( u_1 \) and \( u_2 \) are taken. R.M.S. Error \( (x_1) \) and R.M.S. Error \( (x_2) \) are plotted as a function of \( H \) and are shown in figures (3-3) and (3-4) respectively. The R.M.S. Error \( (x_1) \) and R.M.S. error \( (x_2) \) are seen to decrease exponentially up to \( N = .03 \) seconds and after which computer underflow occurred. In our example as \( H \) varies from .001 seconds to .03 seconds, R.M.S. Error \( (x_1) \) decreases from .06 to .11 and R.M.S. Error \( (x_2) \) decreases from .556 to .102. The differencing interval, \( H \) is chosen to be .03 seconds in this example.

For identifying [A] and [B] matrices, \( u_1 \) and \( u_2 \) in equation (3.77) are Gaussian noises with zero mean values and a variance equal to \( [H]^{-1} \) or \( (.03)^{-1} \) in this example.
FIGURE 3-3    R.M.S. ERROR ($x_1$) VS. DIFFERENCING INTERVAL [H]
TOTAL NO. OF SAMPLES = 1000

Figure 3-4: R.M.S. Error ($x_2$) vs. Differencing Interval (H)
From equation (3.92) the conditional incremental moment functions of the state variables \( x_1 \) and \( x_2 \) are given by,

\[
K_1(x) = E \left\{ \frac{x_1[(n + 1)\Delta t] - x_1(n\Delta t)}{\Delta t} \mid x_1(n\Delta t) \right\}
\tag{3.92}
\]

\[
K_2(x) = E \left\{ \frac{x_2[(n + 1)\Delta t] - x_2(n\Delta t)}{\Delta t} \mid x_2(n\Delta t) \right\}
\tag{3.93}
\]

The conditional incremental moment functions \( \hat{K}_1^{\text{EST}}(X) \) and \( \hat{K}_2^{\text{EST}}(X) \) are estimated using equations (3.92) and (3.93). For this purpose, the procedure given in section 3.4 is followed.

Again from equations (3.30) and (3.31) the theoretical conditional incremental moment functions \( K_1^{\text{THEO}}(X) \) and \( K_2^{\text{THEO}}(X) \) of the state variables \( x_1 \) and \( x_2 \) are given by

\[
K_1^{\text{THEO}}(x) = a_{11}x_1 + a_{12}x_2
\tag{3.94}
\]

\[
K_2^{\text{THEO}}(x) = a_{21}x_1 + a_{22}x_2
\tag{3.95}
\]

The R.M.S. Error associated with the estimates of the conditional first-order incremental statistics of the \( x_1(n\Delta t) \) and \( x_2(n\Delta t) \) processes have been plotted as a
FIGURE 3-5  THE R.M.S. ERROR OF THE ESTIMATE OF THE INCREMENTAL STATISTIC OF $x_1(t)$ GIVEN AS A FUNCTION OF THE NUMBER OF SAMPLES USED IN MAKING THE ESTIMATES
Figure 3-6  The R.M.S. Error of the Estimate of the Incremental Statistic of $x_2(t)$ Given as a Function of the Number of Samples Used in Making the Estimates
function of the number of samples, $N$ used in making the estimates. The results are shown in figures (3-5) and (3-6) respectively. A curve which is proportional to $\frac{1}{\sqrt{N}}$ has been included in these figures for reference purposes. It can be seen from these figures that the R.M.S. errors associated with the estimates of these statistical properties is inversely proportional to the square root of the total number of samples used in making the estimate. It is also shown that the estimates of these incremental moment functions are unbiased.

The estimates of the incremental moment functions $K_1(X)$ and $K_2(X)$ and the corresponding sampled-data information of $x_1$ and $x_2$ are then applied to equation (3.39) for identifying the $[A]$ matrix.

The $[B]$ matrix is identified using equation (3.68), where sampled-data information of the input vector and the state variables are used.

A plot of the $a_{\ell k}$ coefficients vs. the total number of samples is shown in figure (3-7). The $a_{\ell k}$ coefficients converge to their true values as the total number of samples increases from 200 to 3800. The effect of the total number of samples on the identification of $b_{\ell k}$ coefficients
FIG. 7-7 THE EFFECT OF THE TOTAL NUMBER OF SAMPLES ON THE IDENTIFICATION OF $a_{jk}$ COEFFICIENTS

NO. OF INCREMENTS = 1000
DIFFERENCING INTERVAL ($H$) = 0.03

- $a_{11} = -2$, $a_{21} = 0.2$
- $a_{12} = -1$, $a_{22} = -5$
is shown in figure (3-8). The $b_{l,k}$ coefficients converge to their true values with only 200 samples and onward. A sampling interval of .03 seconds and 1000 increments are used.

R.M.S. error ($a_{l,k}$) and R.M.S. error ($b_{l,k}$) are computed using the following relationships:

\[
R.M.S. \text{ Error (} a_{l,k} \text{)} = \sqrt{\frac{1}{N_1} \sum_{l,k=1}^{p} \left( \frac{a_{l,k} - \hat{a}_{l,k}}{|a_{l,k,MAX}|} \right)^2}
\]

\[
(3.96)
\]

\[
R.M.S. \text{ Error (} b_{l,k} \text{)} = \sqrt{\frac{1}{N_1} \sum_{l,k=1}^{p} \left( \frac{b_{l,k} - \hat{b}_{l,k}}{|b_{l,k,MAX}|} \right)^2}
\]

\[
(3.97)
\]

where $l = 1, \ldots, p$

$k = 1, \ldots, p$

$N_1$ is the total number of coefficients.

$a_{l,k}$ and $\hat{a}_{l,k}$ are respectively the theoretical and identified values of $a_{l,k}$.

$b_{l,k}$ and $\hat{b}_{l,k}$ are respectively the theoretical and identified values of $b_{l,k}$.

$|a_{l,k,MAX}|$ and $|b_{l,k,MAX}|$ are the magnitudes of the
maximum values of $a_{\ell k}$ and $b_{\ell k}$ respectively.

A plot of the R.M.S. error ($a_{\ell k}$) vs. the total number of samples is given in figure (3-9). The total number of increments used in estimating the incremental moment functions is 1000. The R.M.S. error associated with this identification decreases from 20.1\% to 3.4\% with the total number of samples increasing from 200 to 3800. A study of figure (3-10) shows that the R.M.S. error associated with the estimation of $b_{\ell k}$ coefficients varies from 6\% to 10\% with the total number of samples increasing from 200 to 3800. The effects of the number of increments on the estimated $a_{\ell k}$ coefficients and R.M.S. error ($a_{\ell k}$) with 1000 samples are shown in figures (3-11) and (3-12) respectively. Figure (3-11) shows that as the number of increments increases from 25 to 925, the $a_{\ell k}$ coefficients converge to their true values. Figure (3-12) shows that the R.M.S. Error associated with the estimation of $a_{\ell k}$ coefficients decreases from 65.5\% to 5.3\% as the number of samples increases from 25 to 925.

Figure (3-13) shows that the R.M.S. Error associated with the estimation of $a_{\ell k}$ coefficients increases from 5.5\% to 45.7\% as the sampling interval varies from 0.03 seconds to 0.12 seconds. The effect of the choice of sampling interval
No. of increments = 1000
Differencing interval (H) = .03

The R.M.S. error associated with the identification of the $a_k$ coefficients given as a function of the total number of samples used.
FIG. 3-10 THE R.M.S. ERROR ASSOCIATED WITH THE IDENTIFICATION OF THE $b_{ik}$ COEFFICIENTS GIVEN AS A FUNCTION OF THE TOTAL NUMBER OF SAMPLES USED.

DIFFERENCING INTERVAL ($H$) = .2
TOTAL NO. OF SAMPLES = 1000
DIFFERENCING INTERVAL (H) = 0.5

$\alpha_{11} = -2$, $\alpha_{12} = -1$
$\alpha_{21} = 2$, $\alpha_{22} = -5$

PARAMETER ESTIMATE ($\hat{a}_{2k}$)

NO. OF INCREMENTS USED FOR ESTIMATING THE INCREMENTAL MOMENT FUNCTIONS

THE EFFECT OF THE NUMBER OF INCREMENTS ON THE IDENTIFICATION OF $a_{2k}$ COEFFICIENTS

FIG. 3-11
FIG. 3.12  THE R.M.S. ERROR ASSOCIATED WITH THE ESTIMATION OF $a_{\xi k}$ COEFFICIENTS PLOTTED AS A FUNCTION OF THE NUMBER OF INCREMENTS USED.
on R.M.S. Error ($b_k$) is shown in figure (3-14). As the
sampling interval varies from .03 seconds to .15 seconds, the
R.M.S. Error ($b_k$) increases from 7.5% to 45.2%.

One can define an exponentially correlated stochastic
process which is simultaneously stationary, Gaussian and
Markovian by the equation

$$\frac{dz_1(t)}{dt} = -uz_1(t) + \gamma_1(t), \quad z_1(t_0) = z_0$$

(3.98)

where $\gamma_1(t)$ is a Gaussian white noise process with

$$E\{\gamma_1(t)\} = 0, \quad R_{\gamma_1 \gamma_1}(\tau) = I_{0}(\gamma_1) \delta(\tau)$$

(3.99)

As $t - t_0$ increases the steady-state value of the auto-
correlation function of the $z_1(t)$ process is

$$R_{z_1 z_1}(\tau) = \frac{I_{c}(\gamma_1)}{2\alpha} e^{-\alpha|\tau|}$$

(3.100)

An exponentially correlated discrete physical process can be
obtained by using an Euler integration formula to obtain an
approximation to the process defined by equation (3.98).
The application of this formula leads to the difference
equation
\[ z_{1}(nH + 1)H = z_{1}(nH) + [-u_{2}(nH) + \eta_{1}(nH)]H \quad (3.101) \]

which defines a discrete exponentially correlated process.

The delta-correlated input appearing in equation (3.98) appears as the \( \eta_{1}(nH) \) term in equation (3.101). This term is realized by using uncorrelated pseudo-random numbers.

The effective properties of the \( z_{1}(nH) \) process are those of a piecewise constant exponentially correlated noise source.

The autocorrelation function associated with \( z_{1}(nH) \) has the form of a sequence of chords bounded below by the curve \( e^{-\alpha|\tau|} \) which they intersect at their end points. These points occur at multiples of the interval \( H \) over which the noise is assumed constant, that is, \( \tau = nH \).

From equation (3.100), the autocorrelation function of the \( z_{1}(t) \) process is

\[ R_{zz_{11}}(\tau) = \frac{I_{c}(\eta_{1})}{2\alpha} e^{-\alpha|\tau|} \quad (3.102) \]

and the variance of the \( z_{1}(t) \) process is

\[ \sigma_{z_{1}}^{2} = \frac{I_{c}(\eta_{1})}{2\alpha} \quad (3.103) \]

The intensity coefficient of \( z_{1}(t) \) is
\[ I_c(z_1) = 2 \int_0^\infty \Phi z_1z_1(\tau) \, d\tau \]

\[ = 2 \int_0^\infty \frac{I_c(\eta_1)}{2} e^{-\alpha \tau} \, d\tau \]

\[ = 2 \frac{I_c(\eta_1)}{2} \left( \frac{1}{\alpha} \right) \left[ e^{-\alpha \tau} \right]_0^\infty \]

\[ = - \frac{I_c(\eta_1)}{\alpha^2} [0 \rightarrow 1] \]

\[ = \frac{I_c(\eta_1)}{\alpha^2} \]  \hspace{1cm} (3.104)

From equation (3.104) if we want to keep the intensity coefficient of the \( z_1(t) \) process equal to that of \( \eta_1(t) \) then we need to multiply \( I_c(z_1) \) by \( \alpha^2 \), which is analogous to multiplying the \( z_1(t) \) process by \( \sqrt{2\alpha} \). Therefore multiplying the \( z_1(t) \) process by \( \sqrt{2\alpha} \), the intensity coefficient of the \( z_1(t) \) process can be kept constant.

The bandwidth and the correlation time of the \( z_1(t) \) process can be expressed as a function of \( \alpha \).

The power spectrum of the \( z_1(t) \) process is given by
\[ \Phi_{z_1 z_1}(w) = \frac{I_c(\gamma'_1)}{\alpha^2 + w^2} \]  \hspace{1cm} (3.105) \\

\[ \tau_{\text{Cor}}(z_1) = \frac{1}{\sigma_{z_1}^2} \int_0^\infty |R_{z_1 z_1}(\tau)| d\tau \]

\[ = \frac{1}{I_c(\gamma'_1)} \int_0^\infty \frac{1}{2\alpha} e^{-\alpha \tau} d\tau \]

\[ = \frac{1}{I_c(\gamma'_1)} \left[ \frac{e^{-d\tau}}{-\alpha} \right]_0^\infty \]

\[= \frac{1}{\alpha I_c(\gamma'_1)} \]  \hspace{1cm} (3.106) 

Now if \( I_c(\gamma'_1) = 1 \), then equation (3.106) becomes

\[ \tau_{\text{Cor}}(z_1) = \frac{1}{\alpha} \]  \hspace{1cm} (3.107)

Therefore, from equation (3.107) changing \( \alpha \), the correlation time of the \( z_1(t) \) process can be varied.

Again from equation (3.105),
\[
\Phi_{z_1 z_1}(w) = \frac{I_c(\eta_1)}{\alpha^2} \quad w = 0 \\
\Phi_{z_1 z_1}(w_b) = .707 \frac{I_c(\eta_1)}{\alpha^2}
\]

where \( w_b = 2\pi f_b \)

and \( f_b \) is the bandwidth of \( z_1(t) \)

Therefore, \[
\frac{1}{\alpha^2 + w_b^2} = .707 \frac{I_c(\eta_1)}{\alpha^2}
\]

with \( I_c(\eta_1) = 1 \), \[
\frac{1}{\alpha^2 + w_b^2} = \frac{.707}{\alpha^2}
\]

Hence \[
.707\alpha^2 + .707w_b^2 = \alpha^2 \]

\[
.293\alpha^2 = .707w_b^2 \\
w_b^2 = \frac{.293}{.707} \alpha^2 \\
w_b^2 = (.41443)\alpha^2 \\
w_b = .64376\alpha \\
f_b = \frac{.64376}{2\pi} \alpha
\]

Hence \( f_b(z_1) = (.10251)\alpha \)
Therefore from equations (3.107) and (3.110), varying 
\(a\), the correlation time and the bandwidth of \(z_1(t)\) can be 
varied. Equations (3.107) and (3.110) show that \(f_p(z_1)\) is 
inversely proportional to \(\tau_{Cor}(z_1)\). Also as mentioned 
before, the intensity coefficient of \(z_1(t)\) can be kept 
the same as that of \(\eta_1(t)\) by multiplying \(z_1(t)\) by \(\sqrt{2a}\). 
The concepts just described can be applied in studying the 
effects of variations of the ratios of correlation times 
or bandwidths of the inputs and the state variables have 
on the identification of \(a_{ik}\) coefficients. Details of 
the mathematical derivation are given in Appendix B.

In our example, exponentially correlated processes 
for both \(u_1\) and \(u_2\) are generated using equation (3.93). 
Experiments are performed for studying the effects of 
changing \(\frac{\tau_{Cor}(x_1)}{\tau_{Cor}(z_1)}\) and \(\frac{\tau_{Cor}(x_2)}{\tau_{Cor}(z_1)}\) 
identification of \(a_{ik}\) coefficients. A plot of the \(a_{ik}\) 
coefficients vs. \(\frac{\tau_{Cor}(x_1)}{\tau_{Cor}(z_1)}\) is shown in figure (3-15).

The figure shows that the \(a_{ik}\) parameters converge to 
their true values for approximately \(\frac{\tau_{Cor}(x_1)}{\tau_{Cor}(z_1)} \geq 14\).
FIGURE 3-15  THE EFFECT OF CHANGING $\frac{\tau \text{ Cor}(x_1)}{\tau \text{ Cor}(z_1)}$ ON THE IDENTIFICATION OF $a_{ik}$ COEFFICIENTS
NO. OF SAMPLES = 1000
NO. OF INCREMENTS = 1000
\( a_{11} = -2, a_{12} = -1 \)
\( a_{21} = 2, a_{22} = -5 \)

PARAMETER ESTIMATE \((a_{jk})\)

\( \frac{\tau \text{ Cor}(x_1)}{\tau \text{ Cor}(z_1)} \)

FIGURE 3-15: THE EFFECT OF CHANGING \( \frac{\tau \text{ Cor}(x_1)}{\tau \text{ Cor}(z_1)} \) ON THE IDENTIFICATION OF \( a_{jk} \) COEFFICIENTS
FIG. 3-16  THE R.M.S. ERROR ASSOCIATED WITH THE IDENTIFICATION OF $a_{ik}$ COEFFICIENTS PLOTTED AS A FUNCTION OF $\frac{\tau_{\text{Cor}(x_1)}}{\tau_{\text{Cor}(z_1)}}$.

NO. OF SAMPLES = 1000
NO. OF INCREMENTS = 1000
A study of figure (3-16) shows that the R.M.S. error associated with the estimation of \( a_{l_k} \) coefficients decreases from 56% to 6.4% with \( \frac{\tau_{\text{Cor}(x_1)}}{\tau_{\text{Cor}(z_1)}} \) varying from 1.38 to 17.38.

Similarly figures (3-17) and (3-18) respectively show that the \( a_{l_k} \) parameters converge to their true values for approximately \( \frac{\tau_{\text{Cor}(x_2)}}{\tau_{\text{Cor}(z_1)}} > 6.5 \) and R.M.S. error associated with the estimation of \( a_{l_k} \) parameters decreases from 56% to 6.4% with \( \frac{\tau_{\text{Cor}(x_2)}}{\tau_{\text{Cor}(z_1)}} \) varying from 1.12 to 9.12.

3.5.3 A Second Identification Example

In the example described in section 3.5.4, the linear system chosen for the demonstration of the identification technique was first simulated on an IBM 360/65 computer to obtain the sampled-data information of the state variables and the input vector. Then the samples of the state variables were used to compute the estimates of the conditional statistics. These estimates and the system vectors were then used to identify the system. In this example, an unknown system is considered whose sampled-data
FIGURE 3.17 THE EFFECT OF CHANGING $\frac{\tau_{\text{Cor}(x_2)}}{\tau_{\text{Cor}(z_1)}}$ ON THE IDENTIFICATION OF $a_{2k}$ COEFFICIENTS
NO. OF SAMPLES = 1000
NO. OF INCREMENTS = 1000

FIG. 3-18 THE R.M.S. ERROR ASSOCIATED WITH THE IDENTIFICATION OF $a_{jk}$ COEFFICIENTS GIVEN AS A FUNCTION OF $\frac{\tau \text{cor}(x_2)}{\tau \text{cor}(z_1)}$. 
information of the state variables and the input vectors were supplied on data cards. From the samples of the state variables, the estimates of the conditional incremental moment functions were computed using an IBM 360/65 computer and then the unknown system was identified using these estimates and the system vector.

Figure (3-19) shows that as the total number of samples vary from 25 to 1,425, the $a_{jk}$ coefficients converge to their true values. Figure (3-20) shows that the R.M.S. error associated with the estimation of $a_{jk}$ coefficients decreases from 38% to 5% with the total number of samples varying from 25 to 1,425.

A study of figure (3-21) shows that the $b_{jk}$ coefficients converge to their true values as the total number of samples vary from 25 to 1,425. Figure (3-22) shows that the R.M.S. error associated with the estimation of $b_{jk}$ coefficients decrease from 35% to 4% with the total number of samples varying from 25 to 1,425. In this example, 1000 increments and a sampling interval of .03 seconds are used.
NUMBER OF INCREMENTS = 1000

\[ a_{11} = -3, \ a_{12} = -2 \]

\[ a_{21} = 1, \ a_{22} = -4 \]

**FIG. 3-19** THE EFFECT OF THE TOTAL NUMBER OF SAMPLES ON THE IDENTIFICATION OF \( a_{jk} \) COEFFICIENTS
NO. OF INCREMENTS = 1000

FIG. 3-20 THE R.M.S. ERROR ASSOCIATED WITH THE IDENTIFICATION OF $a_k$ COEFFICIENTS PLOTTED AS A FUNCTION OF THE TOTAL NUMBER OF SAMPLES USED.
Fig. 3-21 The effect of the total number of samples on the identification of $b_{2k}$ coefficients.

$b_{11} = .5$, $b_{12} = .5$, $b_{21} = 1$, $b_{22} = -2$. 
Fig. 3-22 The R.M.S. error associated with the identification of the coefficients plotted as a function of the total number of samples used.
CHAPTER IV

DISCUSSION

4.1 Introduction

This chapter provides a discussion of the theoretical and experimental results presented in this thesis.

In section 4.2 a discussion of the theoretical developments which forms the basis of the identification technique is presented.

Section 4.3 deals with simulation studies. A discussion of the various aspects of simulation studies as a method of evaluating an identification technique is given.

Finally in section 4.4 some of the practical applications of the identification technique developed in this thesis are considered.

4.2 Discussion of Theoretical Developments

Equation (3.40) has been developed for the identification of the \([A]\) matrix. As mentioned before, this equation is derived using least square technique.

For the identification of the \([B]\) matrix, several approaches have been made. First of all, equation (3.27) is developed which can be used only if \([B]\) is symmetric and
positive definite. But before identification, nature of [B] is completely unknown. Therefore equation (3.27) cannot be used for identifying the [B] matrix which is of general nature. Next, efforts have been made to identify the [B] matrix using geometrical methods and as a result, equation (3.57) is developed for a second-order system. But a problem arises because we have to have a guess on $b_{11}$ initially and also to choose the proper signed values of $b_{12}$ and $s$ in $\theta$. For higher order systems, the identification of [B] matrix using geometrical approach becomes cumbersome. Finally, equation (3.68) has been developed for the identification of [B] matrix.

Hence, equations (3.40) and (3.68) form the basis of the identification technique developed in this thesis.

4.3 Simulation Studies

The identification technique described in Chapter III has been investigated by performing digital computer simulations on two second-order examples. In the first example, the system to be identified was simulated on an IBM 360/65 computer to obtain the sampled-data information of the state variances and system vectors. In the second example, this sampled-data information was supplied on
data cards. The main difference between the simulation studies of the first and second example is that in one case the sampled-data information of the state variables and the system vectors were generated from a known model and then the system was identified. In the other, the system was identified by reading in the data cards which were supplied and which characterized the input-output records associated with an unknown system.

Various aspects associated with the computer-aided identification technique are studied on the first example. For obtaining accurate sampled-data information of the state variables from the simulation of the known model, choice of the differencing interval is very important. For this purpose, a scheme is followed based on impulse response of the system as described in Chapter III. R.M.S. Error \( (x_1) \) and R.M.S. Error \( (x_2) \) are computed for different values of the differencing interval, \( H \). 1000 samples of \( u_1 \) and \( u_2 \) are taken. R.M.S. Error \( (x_1) \) and R.M.S. Error \( (x_2) \) are plotted as a function of \( H \) and are shown in figures (3-3) and (3-4) respectively. The R.M.S. Error \( (x_1) \) and R.M.S. Error \( (x_2) \) are seen to decrease exponentially up to \( H = .03 \) seconds and after which computer underflow occurred.
Initially as $H$ decreases, round-off error increases but truncation error decreases and R.M.S. Error ($x_1$) and R.M.S. Error ($x_2$) increase. Again in the fourth order Runge-Kutta method, truncation error varies directly as $H^5$. Therefore as $H$ increases, R.M.S. Error increases and there is a range of values of $H$ where the truncation and round-off errors neutralize each other. In our example as $H$ varies from $0.001$ seconds to $0.03$ seconds, R.M.S. Error ($x_1$) decreases from $0.6$ to $0.11$ and R.M.S. Error ($x_2$) decreases from $0.556$ to $0.102$. Therefore differencing interval, $H$ is chosen to be $0.03$ seconds for this example. However, the problem of the proper choice of the differencing interval does not arise in a real world situation.

For identifying the $[A]$ and $[B]$ matrices, $u_1$ and $u_2$ in equation (3.76) are Gaussian noise sources with zero mean value and a variance equal to $(H)^{-1}$ or $(0.03)^{-1}$ in this example.

To evaluate the conditional incremental statistics of the state variables $x_1$ and $x_2$, proper choice of the sampling interval is very important. As mentioned in Chapter III, the sampling interval must lie between the noise and process correlation times. The best results would be obtained when the sampling interval is slightly greater than the
correlation time of the noise source. In this example, the sampling interval is kept the same as the differencing interval, that is, .03 seconds. The effective auto-correlation function of the pseudo-random numbers as described in Chapter III is triangular in nature. From where, the correlation time of the noise source can be shown to be \( H/2 \). Therefore the correlation time of the noise source for this example is .015 seconds which is half of the sampling interval.

The conditional incremental moment functions of the state variables \( x_1 \) and \( x_2 \) are estimated using equations (3.91) and (3.92). The R.M.S. Errors associated with these estimates have been plotted as a function of the total number of samples, \( N \), used in making these estimates. The results are shown in figures (3-5) and (3-6) respectively. A curve which is proportional to \( 1/\sqrt{N} \) has been included in these figures for reference purposes. It can be seen from the figures that the R.M.S. Errors associated with the estimates of these statistical properties are inversely proportional to the square root of the total number of samples used in making the estimate. It is also shown that the estimates of these incremental moment functions are unbiased.
The estimates of the incremental moment functions $K_1(x)$ and $K_2(x)$ and the corresponding sampled-data information of $x_1$ and $x_2$ are then applied to equation (3.39) for identifying the $[A]$ matrix.

The $[B]$ matrix is identified using equation (3.68) where sampled-data information of the input vector and the state variables are used.

A plot of the $a_{\ell k}$ coefficients vs. the total number of samples is shown in figure (3-7). The convergence rate of the estimated parameters is seen to be rapid indicating the effectiveness of the technique in obtaining accurate identification results. The effect of the total number of samples on the identification of $b_{\ell k}$ coefficients is shown in figure (3-8). It can be seen that the identified parameters ($b_{\ell k}$) are not sensitive to the total number of samples since with 200 samples and onward, the $b_{\ell k}$ coefficients converge to their true values. A plot of the R.M.S. Error ($a_{\ell k}$) vs. the total number of samples is given in figure (3-9). The R.M.S. Error associated with this identification decreases from 20.1% to 3.4% with the total number of samples increasing from 200 to 3800. The number of increments used was 1000. A study of figure (3-10) shows that the
R.M.S. Error associated with the estimation of $b_{lk}$ coefficients varies from 6% to 10% with the total number of samples increasing from 200 to 3800.

The choice of the number of increments used in estimating the incremental moment functions is very important in identifying the $a_{lk}$ coefficients. As the number of increments increases, a better estimation of the incremental moment functions is expected. Therefore the identified parameters ($a_{lk}$) should converge to their true values with increasing number of increments. The effects of the number of increments on the estimation of $a_{lk}$ coefficients and R.M.S. Error ($a_{lk}$) with 1000 samples are shown in figures (3-11) and (3-12) respectively. Figure (3-11) shows that as the number of increments increases from 25 to 925, the $a_{lk}$ coefficients converge to their true values. Figure (3-12) shows that the R.M.S. Error associated with the estimation of $a_{lk}$ coefficients decreases from 65.5% to 5.3% as the number of increments increases from 25 to 925. Therefore the simulation results agree with the expected results.

As mentioned before, the best results in estimating the incremental moment functions are obtained if the sampling interval is slightly greater than the correlation
time of the noise source. In digital simulations using the IBM 360/65, the minimum sampling interval is the same as the differencing interval, that is, twice the noise correlation time. If the sampling interval is increased keeping the total number of samples, the number of increments and the differencing interval fixed, the error in estimating the incremental moment functions increases. Hence R.M.S. error associated with the estimation of $a_{kek}$ coefficients increases as shown in figure (3-13). R.M.S. Error ($a_{kek}$) increases from 5.5% to 45.7% as the sampling interval varies from .03 seconds to .12 seconds. The effect of the choice of the sampling interval on R.M.S. Error ($b_{kek}$) is shown in figure (3-14). As the sampling interval varies from .03 seconds to .15 seconds, the R.M.S. Error associated with the estimation of $b_{kek}$ coefficients increases from 7.5% to 45.2%. This can be reasoned in the following way. In identifying $b_{kek}$ coefficients using equation (3.68) the derivatives of the state variables are used. As the sampling interval increases, errors in estimating the derivatives of the state variables increase. Hence R.M.S. Error ($b_{kek}$) increases with increasing sampling interval.

The unbiased estimation procedure on the $[A]$ and $[B]$
matrices were performed for the first example. It was found that the bias error expressed in percent of the average value of the \([A]\) matrix was 6%. Correspondingly, a bias error of 10% was associated with the estimation of the \([B]\) matrix.

The main criteria about bandwidths is that the ratio of the noise bandwidth to system bandwidth will be as high as possible. This has been demonstrated and proved by Langseth and Lambert (1968) for a Markov process in general. They compared the performances of several high and low bandwidth Gaussian noise processes and came up with the above conclusion. Actually, experiments involving bandwidth and correlation time give identical results since bandwidth is inversely proportional to the correlation time. However, in the present case, experiments are performed for studying the effects of changing \(\frac{\tau_{Cor}(x_1)}{\tau_{Cor}(z_1)}\) and \(\frac{\tau_{Cor}(x_2)}{\tau_{Cor}(z_1)}\) have on the identification of \(a_{fk}\) coefficients.

A plot of the \(a_{fk}\) coefficients vs. \(\frac{\tau_{Cor}(x_1)}{\tau_{Cor}(z_1)}\) is shown in figure (3-15). The figure shows that the \(a_{fk}\) parameters converge to their true values for approximately
A study of figure (3-16) shows that the R.M.S. error associated with the estimation of $a_{lk}$ coefficients decreases from 56% to 6.4% with $r_{\text{Cor}(x_1)} \over r_{\text{Cor}(z_1)}$ varying from 1.38 to 17.38. Similarly, figures (3-17) and (3-18) respectively show that the $a_{lk}$ parameters converge to their true values for approximately $r_{\text{Cor}(x_2)} \over r_{\text{Cor}(z_1)} > 6.5$ and R.M.S. Error associated with the estimation of $a_{lk}$ parameters decreases from 56% to 6.4% with $r_{\text{Cor}(x_2)} \over r_{\text{Cor}(z_1)}$ varying from 1.12 to 9.12.

Next, a second identification example is chosen where an unknown system is considered whose sampled-data information of the state variables and the input vector were supplied on data cards. From the samples of the state variables, the estimates of the conditional incremental moment functions were computed using an IBM 360/65 computer and then the unknown system was identified using these estimates and the system vectors. Figure (3-19) shows that as the total number of samples vary from 25 to 1425, the $a_{lk}$ coefficients converge to their true values. The
execution time for computing one set of $a_{jk}$ coefficients with 1000 samples and 1000 increments was approximately 55 seconds. Figure (3-20) shows that the R.M.S. Error associated with the estimation of $a_{jk}$ coefficients decreases from 380% to 5% with the total number of samples increasing from 25 to 1425. A study of figure (3-21) shows that the $b_{jk}$ coefficients converge to their true values as the total number of samples vary from 25 to 1425. The execution time for estimating one set of $b_{jk}$ coefficients with 1000 samples was approximately 35 seconds. Figure (3-22) shows that the R.M.S. Error associated with the estimation of $b_{jk}$ coefficients decrease from 35% to 4% with the total number of samples increasing from 25 to 1425.

The main concepts of the identification technique have been adequately demonstrated by the simulation studies described in this thesis. A full study of the problem would require the identification of an actual physical system. However, the simulation studies do validate all the conceptually new material associated with the technique.

4.4 Practical Applications

The identification technique developed in this thesis has the potential to be implemented under the operating
conditions of a plant. For this purpose, the following procedure can be followed.

The sampled-data information of the state variables due to the operating input can be obtained. Then a wideband Gaussian noise can be added to the operating input and the state variable samples due to both will be obtained. Since the system is linear, the sampled-data information of the state variables of the system, due to wideband Gaussian noise only, can be estimated by subtracting the samples of the state variables due to the operating input plus the wideband Gaussian noise from the samples of the state variables due to the operating input only. The sampled-data information of the state variables thus computed and the wideband Gaussian noise can be recorded on a magnetic tape. Then in a computer facility, computing the conditional incremental statistics of the state variables and from the sampled-data information of the system vectors, the system can be identified.
CHAPTER V

CONCLUSIONS

This thesis is concerned with the development of a computer-aided identification technique for linear systems using near-Markov methods. The object is to develop an identification technique for linear systems involving Gaussian noise that does not require a cross correlation operation. These systems when excited by wideband Gaussian noise, generate state variables which are near-Markov in nature. It is possible to estimate the incremental statistics associated with the near-Markov processes from sampled-data information of the state variables. These estimates and the system vectors can be used to identify the system by the technique developed in this thesis.

As the theoretical basis of the identification technique, equations (3.40) and (3.68) have been developed for identifying the [A] and [B] matrices respectively.

The success of the identification scheme is demonstrated by means of digital computer simulations on two second-order examples. Various factors associated with the technique are also studied.
The important aspects associated with the implementation of the identification technique are the sampling interval, number of increments used and the correlation time.

To evaluate the conditional incremental statistics of the state variables, the proper choice of the sampling interval is very important. If a near-Markov physical process is considered to result from the transformation of a realizable noise source with a short correlation time, the minimum time interval over which the Markovian-like properties can be evaluated must be greater than this correlation time. In practice the sampling interval is determined from estimates of the statistical properties of the outputs. An upper-bound on the sampling interval is set by the process correlation time. The optimum sampling interval lies somewhere between the noise and process correlation times. The best results were obtained when the sampling interval was slightly greater than the correlation time of the noise source. In digital computer simulations, the minimum sampling interval is the same as the differencing interval, that is, twice the noise correlation time. In the first example, simulation results show that a sampling interval of .03 seconds gave best results.
The choice of the number of increments used in estimating the incremental moment functions is very important in identifying the [A] matrix. From the simulation results it can be seen that as the number of increments increases the identified parameters \(a_{E_k}\) converge to their true values. It can further be concluded from the experimental results that with 1000 increments, a typical R.M.S. Error associated with the estimation of \(a_{E_k}\) coefficients using this technique would be approximately 10%.

The main criteria about correlation time is that the ratio of the process correlation time to noise correlation time would be as high as possible in order to make the state variables sufficiently near-Markov. It can be concluded from the simulation results that for the first example, as long as the correlation times of \(x_1\) and \(x_2\) are approximately 14 times and 6.5 times respectively, of the noise correlation time, the state variables are sufficiently near-Markov. Hence accurate identification results were obtained. In general, a typical ratio of the correlation times of the state variables to the noise correlation times would be approximately 15.
The computer-aided identification technique developed in this thesis has several possible extensions. The output vector \( [Y] \) of a linear system can be described by the equation of the form \( [Y] = [C][X] + [D][U] \). In the technique developed in this thesis it is assumed that the sampled-data information of the state variables are available, that is, \([C]\) is a unit matrix and \([D]\) is a null matrix. Hence state variables are assumed to be the outputs of the system. However, in a practical situation, \([Y]\) would be observable and therefore a possible extension of this work would be to estimate the state vector \([X]\) from the available output vector \([Y]\) and then identify the system according to the technique developed in this thesis. The concepts described by Lüenberger (1966) may be useful for this purpose. A further extension of this work would be the identification of multivariable nonlinear systems. Both the storage requirements and the number of samples can be considerable, even in the two-dimensional case. It is of interest to determine if these extensions can be carried out in practice.
REFERENCES


Lee, Y. W. (1960), "Statistical Theory of Communications"
John Wiley and Sons, Inc.

M.I.T. - RLE Internal Report, April.

Luenberger, D.G. (1966), "Observers for Multivariable Systems",
IEEE Transactions on Auto. Control, Vol. AC-11,
pp. 190-197, April.

Mehra, R.K. (1971), "On line Identification of Linear Dynamic
Systems with Applications to Kalman Filtering",

Model Estimation", Ph.D. Thesis, University of
Waterloo, Waterloo, Ontario, Canada.

Papoulis, A. (1965), "Probability Random Variables and
Stochastic Processes", McGraw-Hill, New York,
pp. 349-351.

Pazdera, J.S. and Pottinger, H.J. (1969), "Linear System
Identification via Liapunov Design Techniques",

Sáridis, G.N. and Lobbia, R.N. (1972), "Parameter
Identification and Control of Linear Distrete-time
Systems", IEEE Trans. on Auto. Control, Vol. AC-17,
pp. 52-60.

Saridis, G.N. and Stein, G. (1968), "Stochastic Approximation
Algorithms for Linear Discrete-time System Identification",
IEEE Trans. Auto Control, Vol. AC-13, No. 5, October,
pp. 515-523.

Stratonovich, R.L. (1963), "Topics in the Theory of Random
Noise", Vol. 1, Gordon and Breach Science

Stratonovich, R.L. (1965), "A New Representation for
Stochastic Integrals and Equations", J. SIAM

APPENDIX A

Simulation Details

The computer-aided identification technique described in this thesis has been illustrated by means of simulation studies.

The identification procedure was demonstrated by means of two examples as described in Chapter III. In the first example, the system to be identified was simulated on an IBM 360/65 computer to obtain the sampled-data information of the state variables and the system vectors. In the second example, this sampled-data information was supplied on data cards. The main difference between the simulation studies of the first and second example is that in the first case the sampled-data information of the state variables and system vectors were generated from a known model and then the system was identified. In the second case, the system was identified using the input-output records associated with an unknown system.

Certain segments of the program associated with the second example in Chapter III will now be used to illustrate the computer-aided identification technique. The segments of the program chosen indicate the identification procedure adequately. The listings of these segments are given in the following
PROGRAM FOR IDENTIFYING THE [A] MATRIX FOR A SECOND-ORDER SYSTEM WITH THE INPUT - OUTPUT RECORDS SUPPLIED

DIMENSION DX(1010), X1MEAN(1010), X2MEAN(1010), X1MEAN1(1010)
DIMENSION X2MEAN1(1010), DX2(1010), SAMP1(1010), SAMP2(1010)
DIMENSION FUNCX1(1010), FUNCX2(1010), DX7(1010), DX77(1010)
DIMENSION DX8(1010), DX8B(1010), BK1(1010), BK2(1010)
DIMENSION AA1(6), AA2(6), SKUM(6), SSUM(6)
DIMENSION HI(25), AA3(25), AA4(25), AA5(25), AA6(25)

XMAX=1.0821
XMIN=-1.62363
XMAX=2.30263
XMIN=-2.09528
W=-15.45990
V=-10.06966
N=1000
H=.03
N+4=1000
X1=0.0
X2=0.0
A01=1000.0
N=2

A11=(X1MAX-X1MIN)/AB1
A12=(X2MAX-X2MIN)/AB1
D1 23 I=1, N44
DA1(I)=0.0
DA2(I)=0.0
XMEAN(I)=0.0
XMEAN1(I)=0.0
XMEAN2(I)=0.0
SAMP1(I)=0.0

23
SAMP2(I)=0.0
D2 24 I=1, 999
X1=X13
X2=X14
AABS(X1MIN)
BABS(X2MIN)
A11=X11+AA
A22=X2+BB
I1=A11/D1T1
JJ=A22/D2T2
II(I1-EQ.0) II=1
II(I1-EQ.0) JJ=1
R(A05, S5I1X13, X14, W1, V1)

56 INV.(45X, F10.5)
X1=X13
X2=X14
DA1(I)=DX1(I)+((X11-X1)/H)
XMEAN(I)=X1MEAN(I)+X1
XMEAN1(I)=X2MEAN1(I)+X2
DA2(I)=DX2(I)+((X22-X2)/H)
XMEAN(JJ)=X2MEAN1(JJ)+X2
XMEAN1(JJ)=X1MEAN1(JJ)+X1
SAMP1(I)=SAMP1(I)+1.0
SAMP2(JJ)=SAMP2(JJ)+1.0

24 CONTINUE
J1=0
J2=0
N=N44-1
D2 29 I=1, N55
IF(DX1(I).EQ.0.0) GO TO 30
J1=J1+1
FUNCXL(J1)=DX1(I)/SAMP1(I)
DA7(J1)=XMEAN(I)/SAMP1(I)
DA77(J1)=XMEAN2(I)/SAMP1(I)
30 IF(DX2(I).EQ.0.0) GO TO 29
J2=J2+1
FUNCX2(J2)=DX2(I)/SAMP2(I)
DA8(J2)=XMEAN(I)/SAMP2(I)
DA88(J2)=XMEAN2(I)/SAMP2(I)
29 CONTINUE
DO 71 I=1,4
SUM(I)=0.0
71 SUM(I)=0.0
DO 67 I=1,J1
SUM(I)=SUM(I)+(DX7(I)**2)
SUM(2)=SUM(2)+(DX7(I)**2)
SUM(3)=SUM(3)+(DX7(I)**2)
SUM(4)=SUM(4)+(DX7(I)**2)
SUM6=SUM6+SUM(I)
SUM6=SUM6+SUM(I)
DO 73 I=1,4
SUM(I)=SUM(I)+FUNCXL(I)*DX7(I)
73 SUM(I)=SUM(I)+FUNCXL(I)*DX7(I)
A11(I)=SUM(I)
A11(I)=SUM(I)
A12(I)=SUM(I)
A12(I)=SUM(I)
CALC SIMO(SUM,A11,A12,A13)
77 CALL SIMO(SUM,A11,A12,A13)
HALE(6,171,A11,A12)
171 FURMAT(2,5X,F10.5)
DO 75 I=1,J2
SUM(I)=SUM(I)+(DX8(I)**2)
SUM(2)=SUM(2)+(DX8(I)**2)
SUM(3)=SUM(3)+(DX8(I)**2)
SUM(4)=SUM(4)+(DX8(I)**2)
SUM6=SUM6+SUM(I)
SUM6=SUM6+SUM(I)
DO 76 I=1,4
SUM(I)=SUM(I)+FUNCX2(I)*DX8(I)
76 SUM(I)=SUM(I)+FUNCX2(I)*DX8(I)
A12(I)=SUM(I)
A12(I)=SUM(I)
A12(I)=SUM(I)
A12(I)=SUM(I)
CALC SIMO(SUM,A12,A13)
77 CALL SIMO(SUM,A12,A13)
HALE(6,172,A12,A12)
172 FURMAT(2,5X,F10.5)
STOP
END

NOTES: A11(1), A11(2), A12 and A12(2) represent a_{11}^{11}, a_{12}, a_{21} and a_{22} respectively.
PROGRAM FOR IDENTIFYING THE [B] MATRIX FOR A SECOND-ORDER SYSTEM WITH THE INPUT - OUTPUT RECORDS SUPPLIED.

N=25
XMAX=0.0
XMIN=0.0
XMAX=0.0
XMIN=0.0
H=0.03
SM=0.0
SM=0.0
SM=0.0
SM=0.0
SM=0.0
X=0.0
X=0.0
W=-15.45990
V=-10.06966
N=5-1
DO I=1,N
X=X13
X=X14
IF (X1.GT.X1MAX) X1MAX=X1
IF (X2.GT.X2MAX) X2MAX=X2
IF (X1.LT.X1MIN) X1MIN=X1
IF (X2.LT.X2MIN) X2MIN=X2
R=AD(5,56)X13,X14,W1,V1
56 FJ*MAT(5,56,F10.5)
X=X13
X=X14
SM=SM1+( (X1-X1)/H )*W
SM=SM4+( (X2-X2)/H )*V
SM=SM5+(W**2)
SM=SM6+(V**2)
SM=SM2+( (X1-X1)/H )*V
SM=SM3+( (X2-X2)/H )*W
W=W1
V=V1
CONTINUE
B11=SM1/SM5
B12=SM2/SM6
B21=SM3/SM5
B22=SM4/SM6
65 TE(5.571X1MAX,X1MIN,X2MAX,X2MIN
57 FJ*MAT(4,57,F10.5)
65 TE(16,58)B11,B12,B21,B22
58 FJ*MAT(5,3X,F10.5,3X,F10.5,3X,F10.5,3X,F10.5)
57 FJ*UP
END

NOTES: B11, B12, B21, and B22 represent b11', b12', b21' and b22' respectively.
APPENDIX B

BANDWIDTH CONSIDERATIONS

As mentioned in Chapter III, experiments are performed in studying the effects of the variations of the ratios of correlation times or bandwidths of the inputs and the state variables have on the identification of \( a_{lk} \) coefficients. Details of the mathematical derivations relating to these experiments will be given in the following.

From the transfer function matrix given by equation (3.83), the power spectrums of \( x_1 \) and \( x_2 \) can be shown to be

\[
\phi_{x_1x_1}(\omega) = \frac{29 + \omega^2}{(\alpha^2 + \omega^2)(9 + \omega^2)(16 + \omega^2)} \quad (B.1)
\]

\[
\phi_{x_2x_2}(\omega) = \frac{20 + 4\omega^2}{(\alpha^2 + \omega^2)(9 + \omega^2)(16 + \omega^2)} \quad (B.2)
\]

From equations (B.1) and (B.2)

\[
\phi_{x_1x_1}(0) = \frac{29}{144\alpha^2} \quad (B.3)
\]

\[
\phi_{x_2x_2}(0) = \frac{20}{144\alpha^2} \quad (B.4)
\]
Therefore the bandwidths of \( x_1 \) and \( x_2 \) that is \( f_b(x_1) \) and \( f_b(x_2) \) can be expressed in terms of \( a \) as shown in the following equations.

\[
\frac{(29 + w_b^2(x_1))}{(\sigma^2 + w_b^2(x_1)(9 + w_b^2(x_1))) (16 + w_b^2(x_1))} = \frac{-0.07 \times 29}{144 \sigma^2} \tag{B.5}
\]

\[
\frac{(20 + 4w_b^2(x_2))}{(\sigma^2 + w_b^2(x_2)(9 + w_b^2(x_2))) (16 + w_b^2(x_2))} = \frac{0.07 \times 20}{144} \tag{B.6}
\]

Hence using equations (3.110), (B.5), (B.6), and (3.39), experiments can be performed to study the effects of variation of \( \frac{f_b(x_1)}{f_b(z_1)} \) and \( \frac{f_b(x_2)}{f_b(z_1)} \) have on the identification of \( a_{lk} \) coefficients. In our case, experiments are performed for studying the effects of changing \( \frac{\tau_{Cor}(x_1)}{\tau_{Cor}(z_1)} \) and \( \frac{\tau_{Cor}(x_2)}{\tau_{Cor}(z_1)} \) have on the identification of \( a_{lk} \) coefficients. Actually, experiments involving correlation times and bandwidths give identical effects since bandwidth is inversely proportional to correlation time as mentioned before.

Let \( y_3 = w^2 \).

Therefore equations (B.1) and (B.2) become

\[
\Phi_{x_1x_1}(y_3) = \frac{29 + y_3}{(\sigma^2 + y_3)(9 + y_3)(16 + y_3)} \tag{B.7}
\]
\[ \Phi_{x_2 x_2}(y_3) = \frac{20 + 4y_3}{(\alpha^2 + y_3)(9 + y_3)(16 + y_3)} \]  

(B.8)

Let

\[ \frac{29 + y_3}{(y_3 + 9)(y_3 + 16)(y_3 + \alpha^2)} = \frac{A_1}{y_3 + 9} + \frac{B_1}{y_3 + 16} + \frac{C_1}{y_3 + \alpha^2} \]

(B.9)

29 + y_3 = A_1(y_3 + 16)(y_3 + \alpha^2) + B_1(y_3 + 9)(y_3 + \alpha^2) + C_1(y_3 + 9)(y_3 + 16)

In equation (B.9), equating the coefficients of \(y_3^2\),

\[ A_1 + B_1 + C_1 = 0 \]  

(B.10)

Equating coefficients of \(y_3\),

\[ A_1(16 + \alpha^2) + B_1(9 + \alpha^2) + 25C_1 = 1 \]  

(B.11)

Equating the constants,

\[ 16\alpha^2A_1 + 9\alpha^2B_1 + 144C_1 = 29 \]  

(B.12)

From (B.10), \[ C_1 = - (A_1 + B_1) \]  

(B.13)

Substituting the value of \(C_1\) in equation (B.11)

\[ A_1(16 + \alpha^2) + B_1(9 + \alpha^2) - 25 (A_1 + B_1) = 1 \]

\[ A_1 (\alpha^2 - 9) + B_1(\alpha^2 - 16) = 1 \]  

(B.14)
Substituting the value of $C_1$ from equation (B.13) into equation (B.12) we have,

$$16\alpha^2 A_1 + 9\alpha^2 B_1 - 144(A_1 + B_1) = 29$$

$$A_1(16\alpha^2 - 144) + B_1(9\alpha^2 - 144) = 29$$  \hspace{1cm} \text{(B.15)}

From equation (B.14)

$$A_1(\alpha^2 - 9) = 1 - B_1(\alpha^2 - 16)$$

therefore

$$A_1(16\alpha^2 - 144) = 16 - 16B_1(\alpha^2 - 16)$$  \hspace{1cm} \text{(B.16)}

Substituting the value of $A_1(16\alpha^2 - 144)$ from equation (B.16) into equation (B.15) we obtain,

$$16 - 16B_1(\alpha^2 - 16) + B_1(9\alpha^2 - 144) = 29$$

$$B_1(9\alpha^2 - 144 - 16\alpha^2 + 256) = 13$$

Hence, $B_1 = \frac{13}{112 - 7\alpha^2}$  \hspace{1cm} \text{(B.17)}

Substituting the value of $B_1$ from equation (B.17) into equation (B.14),
$$A_1(\alpha^2 - 9) + \frac{13(\alpha^2 - 16)}{112 - 7\alpha^2} = 1$$

$$A_1(\alpha^2 - 9) = 1 - \frac{13(\alpha^2 - 16)}{112 - 7\alpha^2}$$

$$= \frac{112 - 7\alpha^2 - 13\alpha^2 + 208}{112 - 7\alpha^2}$$

$$= \frac{320 - 20\alpha^2}{112 - 7\alpha^2}$$

Therefore

$$A_1 = \frac{320 - 20\alpha^2}{(112 - 7\alpha^2)(\alpha^2 - 9)} \quad (B.18)$$

Again from equation (B.13),

$$C_1 = -(A_1 + B_1)$$

$$= \frac{20\alpha^2 - 320}{(112 - 7\alpha^2)(\alpha^2 - 9)} - \frac{13}{112 - 7\alpha^2}$$

$$= \frac{20\alpha^2 - 320 - 13\alpha^2 + 117}{(112 - 7\alpha^2)(\alpha^2 - 9)}$$

$$= \frac{7\alpha^2 - 203}{(112 - 7\alpha^2)(\alpha^2 - 9)} \quad (B.19)$$

Therefore
\[ \Phi_{x_1 x_1}(w) = \frac{29 + w^2}{(w^2 + 9)(w^2 + 16)(w^2 + \alpha^2)} \]

\[ = \frac{A_1}{w^2 + 9} + \frac{B_1}{w^2 + 16} + \frac{C_1}{w^2 + \alpha^2} \quad \text{(B.20)} \]

where \[ A_1 = \frac{320 - 20\alpha^2}{(112 - 7\alpha^2)(\alpha^2 - 9)} \]

\[ B_1 = \frac{13}{(112 - 7\alpha^2)} \]

\[ C_1 = \frac{7\alpha^2 - 203}{(112 - 7\alpha^2)(\alpha^2 - 9)} \]

Again let

\[ \frac{20 + 4y_3}{(\alpha^2 + y_3)(9 + y_3)(16 + y_3)} = \frac{A_2}{y_3 + 9} + \frac{B_2}{y_3 + 16} + \frac{C_2}{y_3 + \alpha^2} \]

Therefore

\[ 20 + 4y_3 = A_2(y_3 + 16)(y_3 + \alpha^2) + B_2(y_3 + \alpha^2)(y_3 + 9) + C_2(y_3 + 9)(y_3 + 16) \quad \text{(B.21)} \]
Now in equation (B.20), equating coefficients of $y^2_3$,

$$A_2 + B_2 + C_2 = 0$$  \hspace{1cm} (B.22)

Equating coefficients of $y_3$

$$A_2(16 + a^2) + B_2(a^2 + 9) + 25C_2 = 4$$  \hspace{1cm} (B.23)

Equating constants

$$16a^2A_2 + 9a^2B_2 + 144C_2 = 20$$  \hspace{1cm} (B.24)

From equation (B.22)

$$C_2 = -(A_2 + B_2)$$  \hspace{1cm} (B.25)

Substituting the value of $C_2$ from equation (B.25) into (B.23) and (B.24) we obtain

$$A_2(16 + a^2) + B_2(a^2 + 9) - 25(A_2 + B_2) = 4$$

or

$$A_2(a^2 - 9) + B_2(a^2 - 16) = 4$$  \hspace{1cm} (B.26)

$$16a^2A_2 + 9a^2B_2 - 144(A_2 + B_2) = 20$$

$$A_2(16a^2 - 144) + B_2(9a^2 - 144) = 20$$  \hspace{1cm} (B.27)
From equation (B.26)

\[ 16A_2(a^2 - 9) = 64 - 16B_2(a^2 - 16) \]

or

\[ A_2(16a^2 - 144) = 64 - 16B_2(a^2 - 16) \]  \hspace{1cm} (B.28)

Substituting the value of \( A_2(16a^2 - 144) \) from equation (B.28)

into (B.27)

\[ 64 - 16B_2(a^2 - 16) + B_2(9a^2 - 144) = 20 \]

\[ B_2(9a^2 - 144 - 16a^2 + 256) = -44 \]

\[ \therefore B_2 = \frac{44}{9a^2 - 144 - 16a^2 + 256} \]  \hspace{1cm} (B.29)

Therefore from equations (B.26) and (B.29)

\[ A_2(a^2 - 9) + \frac{44(a^2 - 16)}{7a^2 - 112} = 4 \]

\[ \therefore A_2 = \frac{4 - 44(a^2 - 16)}{7a^2 - 112} \]

Therefore

\[ A_2 = \frac{28a^2 - 448 - 44a^2 + 704}{7a^2 - 112} \]

\[ \frac{256 - 16a^2}{(7a^2 - 112)(a^2 - 9)} \]  \hspace{1cm} (B.30)
Again

\[ c_2 = -a_2 - b_2 \]

\[ = \frac{16a^2 - 256}{(7a^2 - 112)(a^2 - 9)} - \frac{44}{7a^2 - 112} \]

\[ = \frac{16a^2 - 256 - 44a^2 + 396}{(7a^2 - 112)(a^2 - 9)} \]

\[ = \frac{140 - 28a^2}{(7a^2 - 112)(a^2 - 9)} \]

(B.31)

Therefore

\[ \Phi_{x_2x_2}(w) = \frac{20 \cdot w^2 + 4w^2}{(w^2 + 9)(w^2 + 16)(w^2 + a^2)} \]

\[ = \frac{a_2}{w^2 + 9} + \frac{b_2}{w^2 + 16} + \frac{c_2}{w^2 + a^2} \]

(B.32)

where

\[ a_2 = \frac{256 - 16a^2}{(7a^2 - 112)(a^2 - 9)} \]

\[ b_2 = \frac{44}{7a^2 - 112} \]

\[ c_2 = \frac{140 - 28a^2}{(7a^2 - 112)(a^2 - 9)} \]
From equations (3.20) and (B.32), taking inverse Fourier transform, the autocorrelation functions $R_{x_1 x_1}(\tau)$ and $R_{x_2 x_2}(\tau)$ are given by

$$R_{x_1 x_1}(\tau) = \frac{A_1}{6} e^{-3|\tau|} + \frac{B_1}{8} e^{-4|\tau|} + \frac{C_1}{2\alpha} e^{-\alpha|\tau|}$$  \hspace{1cm} (B.33)

$$R_{x_2 x_2}(\tau) = \frac{A_2}{6} e^{-3|\tau|} + \frac{B_2}{8} e^{-4|\tau|} + \frac{C_2}{2\alpha} e^{-\alpha|\tau|}$$  \hspace{1cm} (B.34)

The variances $\sigma_{x_1}^2$ and $\sigma_{x_2}^2$ respectively of $x_1(t)$ and $x_2(t)$ are

$$\sigma_{x_1}^2 = R_{x_1 x_1}(0) = \frac{A_1}{6} + \frac{B_1}{8} + \frac{C_1}{2\alpha}$$  \hspace{1cm} (B.35)

$$\sigma_{x_2}^2 = R_{x_2 x_2}(0) = \frac{A_2}{6} + \frac{B_2}{8} + \frac{C_2}{2\alpha}$$  \hspace{1cm} (B.36)

Again

$$\tau_{\text{Cor}}(x_1) = \frac{1}{\sigma_{x_1}^2} \int_{0}^{\infty} |R_{x_1 x_1}(\tau)| \, d\tau$$  \hspace{1cm} (B.37)

$$\tau_{\text{Cor}}(x_2) = \frac{1}{\sigma_{x_2}^2} \int_{0}^{\infty} |R_{x_2 x_2}(\tau)| \, d\tau$$  \hspace{1cm} (B.38)
From equations (B.33) and (B.37)

\[ \sigma_{x_1}^2 \tau_{\text{Cor}}(x_1) = \frac{A_1}{18} + \frac{B_1}{32} - \frac{C_1}{2a^2} \]

\[ = \frac{16A_1a^2 + 9a^2B_1 + 144C_1}{288a^2} \]

(B.39)

Again from equation (B.35)

\[ \sigma_{x_1}^2 = \frac{4A_1a + 3B_1a + 12C_1}{24a} \]

(B.40)

From equations (B.39) and (B.40) we have

\[ \tau_{\text{Cor}}(x_1) = \frac{(16A_1a^2 + 9a^2B_1 + 144C_1)}{288a^2} \]

(B.41)

From equations (3.106) and (B.41),
\[
\frac{\tau_{\text{Cor}(x_1)}}{\tau_{\text{Cor}(z_1)}} = \frac{\alpha(16A_1^2 + 9a^2B_1 + 144C_1)}{48A_1^2 + 36B_1^2 + 144C_1}
\]

(B.42)

\[
= \frac{16A_1^2 + 9a^2B_1 + 144C_1}{48A_1^2 + 36B_1^2 + 144C_1}
\]

Substituting the values of \(A_1\), \(B_1\) and \(C_1\) from equations (B.17), (B.18) and (B.19) into equation (B.42)

\[
\frac{\tau_{\text{Cor}(x_1)}}{\tau_{\text{Cor}(z_1)}} = \frac{16a^2 \frac{(320 - 20a^2)}{(112 - 7a^2)(a^2 - 9)} + 9a^2(13)}{144 \frac{(7a^2 - 203)}{(112 - 7a^2)(a^2 - 9)}}
\]

\[
= \frac{5120a^2 - 320a^4}{(112 - 7a^2)(a^2 - 9)} + \frac{117a^2}{(112 - 7a^2)} + \frac{1003a^2 - 29232}{(112 - 7a^2)(a^2 - 9)}
\]

\[
= \frac{15360a - 950a^3}{(112 - 7a^2)(a^2 - 9)} + \frac{468a}{(112 - 7a^2)} + \frac{1003a^2 - 29232}{(112 - 7a^2)(a^2 - 9)}
\]

\[
= \frac{5120a^2 - 320a^4 + 117a^4 - 1553a^2 + 1003a^2 - 29232}{(112 - 7a^2)(a^2 - 9)}
\]

\[
= \frac{15360a - 950a^3 + 468a^3 - 4212a + 1003a^2 - 29232}{(112 - 7a^2)(a^2 - 9)}
\]
Therefore,

\[
\tau_{\text{cor}(x_1)} = \frac{5120\alpha^2 - 320\alpha^4 + 117\alpha^4 - 1053\alpha^2 + 1008\alpha^2 - 29232}{15360\alpha - 960\alpha^3 + 468\alpha^3 - 4212\alpha + 1008\alpha^2 - 29232}
\]

\[
- 203\alpha^4 + 5075\alpha^2 - 29232
\]

\[
= \frac{-492\alpha^3 + 1008\alpha^2 + 1148\alpha - 29232}{(B.43)}
\]

Again from equations (B.34) and (B.38)

\[
\sigma_{x_2}^2 \tau_{\text{cor}(x_2)} = \frac{A_2}{18} + \frac{B_2}{32} + \frac{C_2}{2\alpha^2}
\]

\[
= \frac{16A_2\alpha^2 + 9B_2\alpha^2 + 1^{1/4}C_2}{288\alpha^2}
\]

(B.44)

From equation (B.35),

\[
\sigma_{x_2}^2 = \frac{4A_2\alpha + 3B_2\alpha + 12C_2}{24\alpha}
\]

(B.45)

From equations (B.44) and (B.45)
\[ \tau_{\text{Cor}}(x_2) = \frac{16A_2\alpha^2 + 9\alpha^2B_2 + 144C_2}{288\alpha^2} \times \frac{24\alpha}{4A_2\alpha + 3B_2\alpha + 12C_2} \]

\[ = \frac{16A_2\alpha^2 + 9\alpha^2B_2 + 144C_2}{48A_2\alpha^2 + 36B_2\alpha^2 + 144\alpha C_2} \]  

(B.46)

From equations (3.105) and (B.46)

\[ \frac{\tau_{\text{Cor}}(x_2)}{\tau_{\text{Cor}}(z_1)} = \frac{\alpha(16A_2\alpha^2 + 9\alpha^2B_2 + 144C_2)}{43A_2\alpha^2 + 36B_2\alpha^2 + 144\alpha C_2} \]

\[ = \frac{16A_2\alpha^2 + 9\alpha^2B_2 + 144C_2}{48A_2\alpha^2 + 36B_2\alpha + 144\alpha C_2} \]  

(B.47)

Substituting the values of \( A_2, B_2 \) and \( C_2 \) from equations (B.23), (B.29) and (B.30) into equation (B.47),
\[
\frac{\tau_{\text{Cor}(x_2)}}{\tau_{\text{Cor}(z_1)}} = \frac{16a^2(255 - 16a^2) + 9a^2(44) + 144(140 - 23a^2)}{(7a^2 - 112)(a^2 - 9)(7a^2 - 112)(a^2 - 9)}
\]

\[
= \frac{4096a^2 - 256a^4 + 396a^2 + 20160 - 4032a^2}{(7a^2 - 112)(a^2 - 9)(7a^2 - 112)(a^2 - 9)}
\]

\[
= \frac{12238a - 768a^3 + 1584a + 20160 - 4032a^2}{(7a^2 - 112)(a^2 - 9)}
\]

\[
= \frac{4096a^2 - 256a^4 + 396a^4 - 3564a^2 + 20160 - 4032a^2}{(7a^2 - 112)(a^2 - 9)}
\]

\[
= \frac{12238a - 768a^3 + 1584a^3 - 14256a + 20160 - 4032a^2}{(7a^2 - 112)(a^2 - 9)}
\]

Therefore

\[
\frac{\tau_{\text{Cor}(x_2)}}{\tau_{\text{Cor}(z_1)}} = \frac{4096a^2 - 256a^4 + 396a^4 - 3564a^2 + 20160 - 4032a^2}{12238a - 768a^3 + 1584a^3 - 14256a + 20160 - 4032a^2}
\]

\[
= \frac{140a^4 - 3500a^2 + 20160}{816a^3 - 4032a^2 - 1968a + 20160}
\]
Using equations (3.39), (B.43) and (B.48) the effects of
variation of $\frac{\tau_{\text{Cor}(x_1)}}{\tau_{\text{Cor}(z_1)}}$ and $\frac{\tau_{\text{Cor}(x_2)}}{\tau_{\text{Cor}(z_1)}}$
on the identification
of a $a_{kl}$ coefficients can be studied. For example, $\frac{\tau_{\text{Cor}(x_2)}}{\tau_{\text{Cor}(z_1)}}$
and $\frac{\tau_{\text{Cor}(x_2)}}{\tau_{\text{Cor}(z_1)}}$ can be varied by changing $\alpha$ and the
$a_{kl}$ coefficients can be identified accordingly.
VITA AUCTORIS

1949 Born in Bangladesh

1963 Completed highschool education in Sylhet, Bangladesh

1965 Completed college education in Dacca, Bangladesh

1969 B.Sc. in Electrical Engineering at E.P. University of Engineering and Technology, Dacca, Bangladesh

1972 M.A.Sc. in Electrical Engineering at the University of Windsor, Windsor, Ontario, Canada

1974 Candidate for Ph.D. in Electrical Engineering at the University of Windsor, Windsor, Ontario, Canada