CANADIAN THESES ON MICROFICHE

I.S.B.N.

THESES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL-339 (r. 82/08)

Canada

37

0-315-12352-4

National Library of Canada | Bibliothèque nationale du Canada

Canadian Theses Division | Division des thèses canadiennes

Ottawa, Canada
K1A 0N4

60263

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur

Chong Gun Kim

Date of Birth — Date de naissance

Jan. 2, 1948

Country of Birth — Lieu de naissance

Korea

Permanent Address — Résidence fixe

133, Michener Park, Edmonton, Alberta

Title of Thesis — Titre de la thèse

Identification of Linear stochastic systems

University — Université

Univ. of Alberta

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée

Ph. D.

Year this degree conferred — Année d'obtention de ce grade

1982

Name of Supervisor — Nom du directeur de thèse

Prof. Gourishankar

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

Date

Sept. 27, 1982

Signature

Chong Kim

NL-91 (4/77)

THE UNIVERSITY OF ALBERTA

IDENTIFICATION OF LINEAR STOCHASTIC SYSTEMS

by

© Chong G. Kim

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL ENGINEERING

EDMONTON, ALBERTA

FALL 1982

Microfilming

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR Chong G. Kim

TITLE OF THESIS IDENTIFICATION OF LINEAR STOCHASTIC SYSTEMS

DEGREE FOR WHICH THESIS WAS PRESENTED DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE GRANTED FALL 1982

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) Chong K. Kim

PERMANENT ADDRESS

133 Michener Park

Edmonton, Alberta

.....

DATED Sept. 1 19 82

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled IDENTIFICATION OF LINEAR STOCHASTIC SYSTEMS submitted by Chong G. Kim in partial fulfilment of the requirements for the degree of DOCTOR OF PHILOSOPHY.

V. Gomiskanker
.....

Supervisor

.....

.....

.....

.....

.....

External Examiner

Date..September 1 1982.....

ABSTRACT

In this thesis, a fast algorithm for the parameter identification of linear stochastic systems is derived and its unbiased convergence property is proved.

To lay the ground work for the above development, a new faster algorithm for linear deterministic system is derived by improving the previously published Kudva-Narendra algorithm. Geometric interpretation of the Kudva-Narendra algorithm helps one to understand how this algorithm works and consequently to improve the convergence rate, thus resulting in a rapidly converging algorithm with reduced computational requirements.

The algorithm for the stochastic system is obtained by modifying the Extended Kalman Filter algorithm using a canonical form of the innovations representation and using the concepts developed for the deterministic case. This development reveals that other existing algorithms which were developed using different approaches can be shown to be modifications of the EKF algorithm.

The effectiveness of this algorithm is compared with two other existing algorithms by means of simulation studies. These studies show the superiority of the new

algorithm with respect to convergence rate and computational requirements.

ACKNOWLEDGEMENTS

The author acknowledges with a deep gratitude the encouragement and direction from his supervisor Prof. V. G. Gourishankar throughout the research work.

The author is also very grateful to his family whose patience and sacrifice have made this work possible.

This work was supported by the National Research Council of Canada. Their financial support is acknowledged. The support of the Department of Electrical Engineering and its chairman Dr. C. R. James is also gratefully noted.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER I

GENERAL INTRODUCTION

1.1 Introduction

Any successful application of modern control theory largely depends on the information available with regard to the physical system of interest. This information is usually incorporated into a mathematical model that behaves identical to the physical system in terms of its input-output relations. The mathematical model can be characterized mainly by two aspects:

a. structure

b. parameters

Therefore, the collection of the information about the physical system is equivalent to the determination of the structure and parameters of the mathematical model on the basis of input and output information (Zadeh, 1962). Such a determination is often referred to as the identification problem and is the main subject pursued in this thesis.

It is extremely useful if it is possible to identify a black box only with input and output information. However, a basic principle of identification is that a completely unknown system cannot be identified. The first step, then, to be taken in the formulation of an identification problem is to characterize the system. This characterization, in general, involves the selection of an appropriate structure

for the system. The structure chosen should reflect factors such as, any à priori knowledge of the system and the ultimate purpose of the identification.

Due to the advent of the digital computer, it is now possible to employ sophisticated methods to control systems. Such methods require, however, that on-line estimation of the parameters be made even as the implementation of the control scheme is in progress. Recursive estimation is often mandatory in many modern adaptive control systems. Thus, considerable attention has recently been paid to recursive identification algorithms.

Of the many identification techniques which have been developed in recent years, the model reference technique possesses a strong intuitive appeal in that it fully utilizes the advantage of the recursive algorithm. The use of the term **model reference** in the context of identification perhaps needs some justification. As first conceived by Whittaker (1958) for aircraft control, the model reference technique adjusts the controller parameters in such a manner that the parameters of the composite system namely, the controller and the system to be controlled, approach the parameters of the reference model in some optimal sense. The philosophy of this technique can be used for the identification problem by adjusting the model parameters to follow the parameters of the unknown system, which is looked upon as the reference system. This adjustment is best done

in a recursive fashion. A natural way to come up with a suitable recursive algorithm would be to introduce a cost function in terms of the error between the model and the system, and to devise an adjustment mechanism for minimizing the cost function. This approach not only converts the problem of the model reference identification into an optimization problem but also provides a quantitative measure of the performance of the adjustment mechanism.

The recursive least-squares (LS) algorithm is perhaps the best known and widely used for linear systems (Astrom, et al., 1971). It is well known that this algorithm leads to a biased estimate when noise, if present, is correlated as is the case in the identification problem of linear stochastic systems. In order to avoid this serious drawback a number of modifications has been proposed (Isermann, et al., 1974; Saridis, 1974; Soderstrom, et al., 1978).

Among them the Extended Least Squares (ELS) algorithm (Panuska, 1968: 1969) is the easiest to implement and consequently has gained popularity in real-life applications (Ljung, et al., 1975). The basic idea of this approach is that noise is regarded as an additional input signal and the LS technique is applied to it. Since the noise is not available for measurement, it is replaced by previous residuals. Early simulation studies seemed to indicate that this method had good convergence property (Panuska, 1968; Kashyap, 1974). But a study carried out by Ljung, et. al

(1975; 1977), however, has revealed that, contrary to prior simulation results, the ELS method will not always converge to the true values of the parameters since convergence is dependent on a condition predetermined by the unknown process parameters. This nonconvergence, which apparently stems from the replacement of noise by previous residuals, can be corrected by the introduction of a prefilter involving the unknown parameters. Approximating the prefilter recursively with the updated estimates results in the recursive approximate maximum likelihood (RML) method (Soderstrom, 1973). This algorithm was developed by an approximation of the off-line maximum likelihood (ML) method (Astrom, et al., 1966) and it possesses a nice local convergence property. No global convergence results are, however, available at this time (Panuska, 1980b).

By the inclusion of the unknown parameters into its state vector a linear stochastic system becomes a nonlinear stochastic system. Hence the identification problem of a linear stochastic system is equivalent to the state estimation problem of a nonlinear stochastic system. Among the numerous approaches available for the nonlinear estimation problem, the Extended Kalman Filter (EKF) approach turns out to be suitable in many cases (Jazwinski, 1970). Since the EKF approach is based on linearization at each step of iteration around the previous estimates and the application of the Kalman filtering theory (Kalman, 1960),

the algorithm usually suffers from the lack of global stability and often involves excessive computation. This provides the motivation for modifying the EKF algorithm into purely a parameter identifier in order to reduce the computational requirements.

The main difficulty associated with the stochastic identification problem has been largely the lack of methods for the analysis of convergence properties. This motivated the extension of the stability theories, which had been developed mainly for the deterministic problem, to the stochastic problem (Kushner, 1967; Jumarie, 1979). Of a number of stability theories, the hyperstability theory (Popov, 1962; Zames, 1966) has been most extensively used in various applications to the deterministic problem in order to obtain adaptive stable systems (Landau, 1974). Thus, it is a rather natural choice to extend the applicable area of the hyperstability theorem to the stochastic identification problem. This was first successfully achieved by Landau (1976) with an introduction of a decreasing gain matrix to estimate the parameters of a linear system contaminated by a colored noise. This work was again extended by Ljung (1977a) to estimate the parameters of the noise as well as the system. The real importance of the Ljung's work, however, rests on finding the role of a positive real transfer function in the ELS method. This finding not only enables one to define the region of convergence but also

suggests a way leading to a convergent algorithm.

1.2 Scope of the Thesis

The main objective of this thesis is the discussion and the development of several model reference adaptive identification algorithms for linear stochastic systems.

In Chapter II, a representation of the linear stochastic system is fully discussed in order to facilitate the formulation of the identification problem treated in Chapter IV. Possible approaches to be taken, in general, to this problem are also discussed.

Chapter III presents a fast adaptive identification method for a linear deterministic system. The main reason for doing this is to develop certain ideas which will be used to solve the identification problem for the linear stochastic system. The identification problem for the stochastic case is, in general, a nonlinear problem and not amenable to direct solution. Therefore, a solution will be attempted by linearization. In this task, the concepts used in the development of the algorithm for a deterministic system will be useful.

In Section 3.2, a geometric interpretation of the Kudva-Narendra scheme (Kudva, et al., 1974) is given. This approach, which is very crucial to the development of the fast algorithm, provides greater insight into and consequently a clearer understanding of the K-N scheme. In

Section 3.3, the concept of orthogonal gradient is used in order to develop the algorithm for rapid convergence. In Section 3.4, the effectiveness of the algorithm proposed is demonstrated by the simulation studies

Chapter IV deals with the linear stochastic identification problem. Since this identification problem as formulated is inherently a nonlinear estimation problem, approximation techniques are used to obtain a filter that is physically realizable. Among the many techniques that can be used, the EKF technique is of particular interest because of its simplicity and thus popularity in real applications. Section 4.2 discusses the application of the EKF to the identification problem. In an attempt to reduce the computational requirements which arise in this approach, modifications of the EKF algorithm are made resulting in two simplified algorithms. These are given in Section 4.3 and 4.4. For obvious reasons these are labelled as EKF-M1 and EKF-M2, respectively. Though these two algorithms have similarities to some existing ones to the author's best knowledge, the exact algorithms reported here have not appeared in the literature.

Convergence analysis is probably one of the most important means to assess the usefulness of identification algorithms. However, since EKF-M1 turns out to be similar to the RML algorithm, no detailed discussion of its convergence property is included here. In Section 4.5,

attention is therefore focused on the convergence property of EKF-M2. In Section 4.6, a generalization of EKF-M2 leading to a third algorithm EKF-M3 is discussed. Comparison of the features of EKF-M3 with those of two other existing algorithms, namely, the RML method (Soderstrom, 1973) and the ELS method (Panuska, 1968; 1969).

The usefulness of EKF-M3 is evaluated by means of simulation studies. For purpose of comparison, simulation studies are also carried out with the RML and ELS algorithms. These resuls are reported in Chapter V.

Chapter VI provides a summary of the thesis, conclusions and suggestions for further research.

CHAPTER II

FORMULATION OF IDENTIFICATION PROBLEM

2.1 Introduction

An identification problem is characterized by three quantities (Zadeh, 1962): a class of model, a class of input signals and a criterion. The selection of the class of models generally leads to the determination of the structure that mathematically represents the system. Since input signals have a profound effect on the performance of the identification scheme some limitations are usually imposed on input signals. It has been shown (Aoki, et al., 1970; Yuan, et al., 1977) that the persistent excitation condition of input signals is sufficient in many cases to obtain consistent estimates or improve the performance of the identification scheme. However, the selection of input signals itself is the subject of an interesting but separate area of research, hence we will not pursue this matter any further.

The structure of models describing the system under consideration wields considerable influence on the complexity of identification schemes in various ways such as computational requirements, consistency and uniqueness of estimates, etc. Since different identification methods are necessary for different representations there are very few general rules available by which the optimality of the model

structure chosen can be determined. A convenient characterization of the model for the recursive model reference approach to the identification problem can be made through the parametric representation since this is relatively easy to implement on computers.

In the parametric representation of linear systems, the choice of model structure by which the system is represented is influenced by the number of unknown parameters to be identified and the applicability of the real-time optimization techniques which are readily available now. The number of parameters to be identified is directly related to the efficiency of the identification algorithm as well as the uniqueness of the solution because there exists a limit to number of the paramters which can be uniquely determined by the input-output relations. Since the identification problem is often formulated as an optimization problem by introducing a cost function, the applicability of the available optimization techniques should be taken into account in the choice of a model structure.

In this chapter we restrict our discussion to linear multiinput singleoutput (MISO) stochastic systems for the sake of simplicity. Since a linear multiinput multioutput (MIMO) system can be regarded as a group of MISO systems each of which has an input-output pair associated with it, the extension to MIMO systems can be readily carried out.

2.2 Representation of MISO Linear Stochastic Systems

Kalman (1963) has shown that only the completely controllable and completely observable part of a system can be characterized from input-output relations. Thus, the system that is of interest in this thesis is one which is completely controllable and completely observable. The phase variable representation (Chen, 1970) of such a system is:

$$x(k+1) = Ax(k) + bu(k) + \omega(k) \tag{2.1a}$$

$$y(k) = h'x(k) + \nu(k) \tag{2.1b}$$

where $x(k)$ is an nx1 state vector,

$u(k)$ is an input variable,

$\omega(k)$ is an nx1 system noise,

$y(k)$ is an output variable,

$\nu(k)$ is a measurement noise,

$$A = \left[\begin{array}{c|c} a & \begin{matrix} I \\ \hline \end{matrix} \\ \hline 0 & \end{array}\right]$$

$h = [1\ 0\ \cdots\ 0]'$,

$a = [a_1\ \cdots\ a_n]'$,

$b = [b_1\ \cdots\ b_n]'$,

I is the unity matrix of appropriate dimension.

and $'$ denotes tranpose of a matrix.

Here $\omega(k)$ and $\nu(k)$ are assumed to be $N(0,Q)$ and $N(0,r)$, respectively. By means of Kalman filtering theory (Kalman,

1960) the number of the unknown parameters which characterize noise can be reduced by an innovations representation (Kailath, 1968; Geesey, et al., 1969; Astrom, 1970) to a form which requires a total of only 3n+1 unknown parameters to completely characterize the n-th order system to be identified against [2n+1]+[n(n+1)/2] unknown parameters in equation (2.1), namely

$$x(k+1) = Ax(k) + bu(k) + dv(k) \qquad (2.2a)$$

$$y(k) = h'x(k) + v(k) \qquad (2.2b)$$

where $v(k)$ is $N(0,w)$,

$w = h'Ph + r$

P is the steady-state covariance matrix of the state error of the Kalman filter,

and $d = [d_1 \cdots d_n]'$ is the steady state Kalman filter gain.

Since the system matrix A in equation (2.2a) contains the unknown vector a, it is useful to replace the partially known matrix A with F, a known matrix but preserving the same structure of the matrix A for the identification problem.

Equation (2.2) can be thus rewritten as:

$$x(k+1) = Fx(k) + (a-f)x_1(k) + bu(k) + dv(k) \qquad (2.3a)$$

$$y(k) = h'x(k) + v(k) \qquad [2.2b]$$

where $x_1(k) = hx(k)$

$$F = \left[\begin{array}{c|c} f & \begin{matrix} I \\ \hline 0 \end{matrix} \end{array}\right]$$

and $f = [f_1 \cdots f_n]'$

Using equation (2.2b), equation (2.3a) becomes

$$x(k+1) = Fx(k) + (f-c)v(k) + (a-f)y(k) + bu(k) \qquad (2.3b)$$

$$y(k) = h'x(k) + v(k) \qquad [2.2b]$$

where $c = a - d$

Regarding the state vector $x(k)$ as the sum of the state vectors $x^{(1)}(k)$, $x^{(2)}(k)$ and $x^{(3)}(k)$ of the three systems driven by $v(k)$, $y(k)$ and $u(k)$, respectively, we obtain:

$$x^{(1)}(k+1) = Fx^{(1)}(k) + (f-c)v(k) \qquad (2.4a)$$

$$x^{(2)}(k+1) = Fx^{(2)}(k) + (a-f)y(k) \qquad (2.4b)$$

$$x^{(3)}(k+1) = Fx^{(3)}(k) + bu(k) \qquad (2.4c)$$

$$y(k) = h'x^{(1)}(k) + h'x^{(2)}(k) + h'x^{(3)}(k) + v(k) \qquad (2.4d)$$

Since equations (2.4) can be considered as a combination of three singleinput singleoutput (SISO) systems they are rewritten as:

$$z_1(k+1) = F'z_1(k) + hv(k) \qquad (2.5a)$$

$$z_2(k+1) = F'z_2(k) + hy(k) \qquad (2.5b)$$

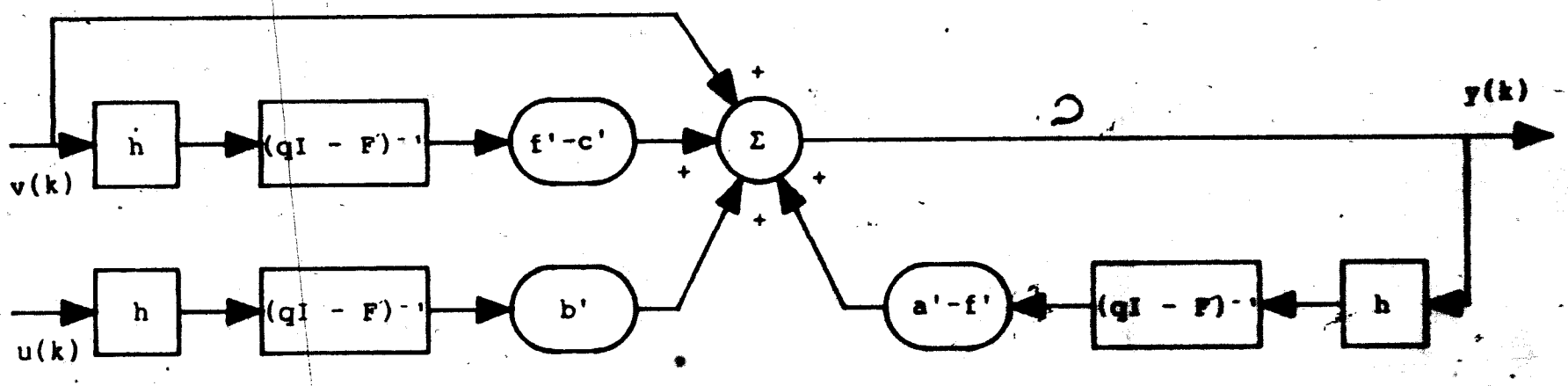
$$z_3(k+1) = F'z_3(k) + hu(k) \qquad (2.5c)$$

$$y(k) = (f-c)'z_1(k) + (a-f)'z_2(k) + b'z_3(k) + v(k) \qquad (2.5d)$$

which are input-output equivalent to equation (2.2). The block diagram representation of this system is provided in Fig. 2-1 where q denotes the shift operator i.e. $q^{-1}y(k) = y(k-1)$.

This representation has long been used in the adaptive observer problem for deterministic systems (Anderson, 1974; Kraft, 1976; Kreisselmeier, 1977) as is the case $v(k) = 0$. The apparent advantage of this representation for the deterministic system is the linearity in parameters (Astrom, et al., 1971) so that the usual least-squares method can be used for identification problems (Kudva, et al., 1974). However, for stochastic systems this advantage quickly disappears since equation (2.5d) is nonlinear in the parameters because $v(k)$, therefore $z_1(k)$, is not available for measurement. It should be noted that the above representation is equivalent to an ARMA-model if all the elements of f are zero.

Besides the determination of the structure of model an important requirement in the identification problem is the choice of an appropriate model order. It is well known that the parametric models result in estimates with large errors



2.1 The system representation

when the order of models does not match with that of the system to be identified. There are several methods available for choosing an appropriate model order (Anderson, 1962; van den Boom, et al., 1974; Unbehauen, et al., 1974; Soderstrom, 1977; Young, et al., 1980). Since this subject also stands alone as an interesting research project we leave it here without further discussion.

Since the omission of the control input u(k) does not affect the characteristics of equations (2.5) except to decrease the number of unknown parameters, we shall drop the control input u(k) in the sequel for the sake of simplicity, giving rise to a problem of time-series identification.

The formation of a state vector $\theta(k)$ by including all the unmeasurables in equations (2.5) as:

$$\theta(k)' = [\theta_1(k)', \theta_2(k)', \theta_3(k)']$$

$$\theta_1(k) = z_1(k)$$

$$\theta_2(k) = f\text{-}c$$

$$\theta_3(k) = a\text{-}f$$

lends itself to a nonlinear equation in the state space as:

$$\theta(k+1) = \Phi\theta(k) + \eta v(k) \qquad (2.6a)$$

$$y(k) = [0\ 0\ z_2(k)']\theta(k) + 1/2\theta(k)'S\theta(k) + v(k) \qquad (2.6b)$$

where

$$\Phi = \begin{bmatrix} F' & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \qquad \eta = \begin{bmatrix} h \\ 0 \\ 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 0 & I & 0 \\ I & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and 0's are understood as vectors or matrices of appropriate dimension.

2.3 Approaches to the Linear Identification Problem

Most of the approaches to the recursive identification problem of linear stochastic systems have been developed from statistical considerations and many of them are based on the modification of the LS method. Among them the best known are the RML method (Soderstrom, 1973), the recursive instrumental variable (IV) method (Young, 1974), the recursive prediction error (PE) method (Moore, et al., 1979; 1980) and the EKF algorithm (Kopp, et al., 1963; Cox, 1964). The main shortcoming of these methods, though some progress has been recently made (Ljung, 1977b), is the difficulty encountered when attempts to analyze the convergence properties of the corresponding algorithms are made.

Recursive identification problems could also be treated as a model tracking problem. It then becomes natural to introduce the model reference adaptive (MRA) method (Landau, 1974) as a design tool. In this approach

the recursive algorithm is designed to make the function of errors decrease asymptotically. For this task, it is useful to apply well-known stability theories, notably the Kalman-Yakubovich (K-Y) lemma (Hitz, et al., 1969) and hyperstability theorem (Landau, 1969; 1979). The advantage of this approach is that the overall stability of the resulting error system with respect to the convergence of the algorithm is assured.

Two configurations related to the MRA technique could be laid out depending on the error measurement methods (Lion, 1966). The equation error method† has been widely used for the problem of adaptive observers and schemes which work satisfactorily in real applications have been reported (Luder, et al., 1974). The popularity of this method for the particular problem is largely due to the ease with which the gradient of the quadratic cost function of the errors can be obtained and the simplicity of designing an recursive algorithm to decrease the cost function asymptotically. The results of this method, however, are known to be biased if correlated noise is present. For this reason, the applicability of this method is somewhat limited for stochastic problems. The output error method, on the other hand, has an appealing potential in that it has been shown

---

†The equation error method is often referred as the series-parallel approach and the output error method the parallel approach. These terms apparently have their origin in the configurations used for error measurement (Martin-Sanchez, 1976).

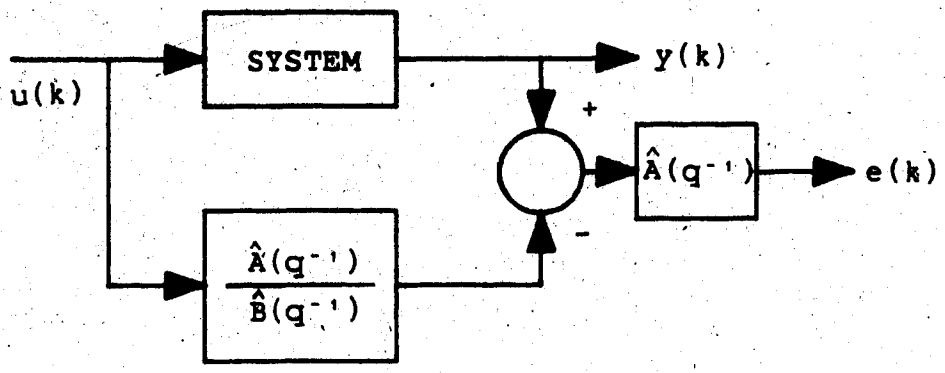
to yield an unbiased estimate even in the presence of correlated noise (Landau, 1976; 1978). Research has since concentrated on the possible extension of this method to the problem of the identification (Dugard, et al., 1980) as well as to adaptive control of stochastic systems. However, for stochastic systems it turns out that the conditions imposed for the overall stability are not always possible to be satisfied because of the involvement of the parameters that are to be identified (Ljung, 1977a; Egardt, 1980; Johnson, 1980).

Fig. 2-2 shows the block diagrams for two error measurement methods

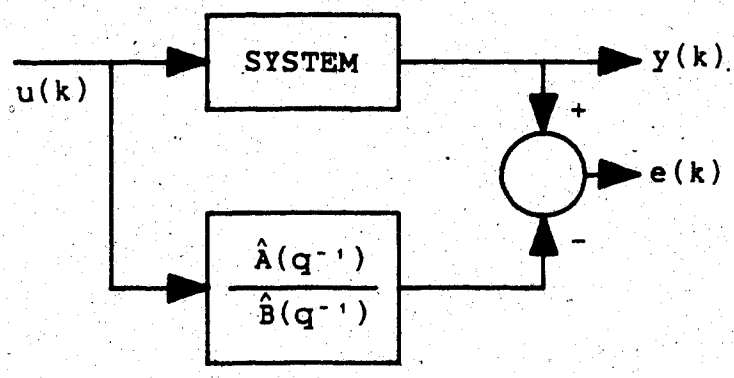
where $A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_n q^{-n}$ (2.7a)

$B(q^{-1}) = 1 + b_1 q^{-1} + \cdots + b_n q^{-n}$ (2.7b)

The inclusion of the unknown parameters to the augmented state vector transforms a parameter identification problem into a nonlinear estimation problem. Since the optimal solution to nonlinear filtering problems is not, in general, realizable in the finite dimensional systems (Bucy, et al., 1971), nonlinear filters are, in most cases, obtained by approximate design methods. While the EKF offers the best simplicity in implementation, thus, being most widely used among the numerous methods, the application of this approach often unnecessarily increases the computational requirements if only the parameters need to be known. It would appear to be worth while to attempt a



a) The equation error measurement



b) The output error measurement

2.2 Two error measurement methods

modification of the EKF in order to minimize the computational burden. It is noteworthy that the convergence properties of some modified algorithms of the EKF could be analyzed with the stability theories used in the case of the MRA method.

In the next chapter, we will first discuss the identification problem for the deterministic system and develop a fast algorithm through a clear insightful understanding of the existing algorithm.

CHAPTER III

A RAPID IDENTIFICATION SCHEME FOR DETERMINISTIC SYSTEMS

3.1 Introduction

Since the stochastic identification problem formulated in the previous chapter is intrinsically nonlinear and thus practically not amenable to an exact solution, an approximate solution will be sought through the linearization of the problem. While the setting in which the solutions sought for the deterministic identification problem can be considered to be too idealistic from the stochastic point of view, the technique developed in this chapter for the improvement of an existing deterministic identification algorithm will be useful to conceptually understand and thus solve the stochastic problem. The remaining portion of this chapter will be devoted to the application of MRA techniques to the deterministic identification problem.

As briefly mentioned in the previous chapter, the preservation of the linearity in the parameters, by which it is meant that a generalized error is linear in parameters (Astrom, et al., 1971), of the deterministic system provides a good opportunity for a wide range of applications of the MRA technique to various problems of the linear deterministic system. Furthermore, the linear system theories represent the most active research area and thus

are readily available for various applications including MRA schemes. Consequently, numerous successful MRA schemes for the identification problem as well as the control problem have appeared in the literature (Landau, 1974; Kreisselmeier, 1980).

With a few exceptions, most of the recent developments in the MRA schemes have adopted the stability-based approach in designing an adaptation mechanism. In this approach, stability theories are used to determine the sufficient conditions under which the adaptation mechanism makes the overall system asymptotically and globally stable. Thus, the design problem associated with the deterministic MRA method corresponds to a stability problem. The crucial point in the design procedure is therefore to choose an equation for updating the adjustable parameters in such a way that the overall nonlinear error system, which constitues the state error system and the adaptation mechanism, is asymptotically stable.

Kudva and Narendra (1974) employed the Lyapunov Direct Method, which is one of the most widely used stability-based design methods, to develop a simple adaptation mechanism for the identification of the discrete multivariable system. In the design procedure, it is assumed that all the state variables are accessible and that there is no noise involved. While this identification scheme (K-N scheme for short) possesses the overall stability, simulation studies

show that the convergence speed is often undesirably slow. Thus, there is need for increasing the speed of convergence.

A geometric interpretation of this adaptive scheme provides a greater insight into and consequently a clearer understanding of the scheme. Though the underlying concept is rather simple, this becomes very crucial in the development of an algorithm with rapid convergence property which is important for on-line applications. This approach also provides a general idea of how the gain of the improved algorithm is to be modified if noise is present. This, in turn, has an enormous implication in the development of an identification algorithm for the stochastic system developed in Chapter IV.

3.2 Geometric Interpretation of the K-N Scheme

Kudva and Narendra (1974) considered the problem of identifying the unknown but constant nxn and nxr, ($n \geq r$) matrices A and B of a noise free dynamic discrete system given by

$$x(k+1) = Ax(k) + Bu(k) \tag{3.1}$$

and have derived the following adaptive algorithm:

$$\hat{x}(k+1) = C\hat{x}(k) + [\hat{A}(k+1) - C]x(k) + \hat{B}(k+1)u(k) \tag{3.2}$$

$$[\hat{A}(k+1),\hat{B}(k+1)] = [\hat{A}(k),\hat{B}(k)] - \frac{\alpha\Sigma[e(k) - Ce(k-1)]\phi(k-1)'}{\lambda\phi(k-1)'\phi(k-1)} \qquad (3.3)$$

where C is an nxn stable matrix, i.e. all its eigenvalues lie inside the unit circle,

$$\phi(k)' = [x(k)',u(k)'],$$

λ is the largest eigenvalue of Σ

and $e(k) = \hat{x}(k) - x(k)$.

Kudva and Narendra have shown that estimate [A(k),B(k)] will converge asymptotically to [A,B] provided that

i) the gain matrix Σ is chosen to be symmetric and positive definite,

ii) $0 < \alpha < 2$,

iii) the system (3.1) is completely controllable

and iv) the input sequence u(k) is general enough (Kudva, et al., 1974).

It follows from equations (3.1-3) that

$$e(k+1) = Ce(k) + \tilde{P}(k+1)\phi(k) \qquad (3.4)$$

and $$\tilde{P}(k+1) = \tilde{P}(k) - \frac{\alpha\Sigma\tilde{P}(k)\phi(k-1)\phi(k-1)'}{\lambda\phi(k-1)'\phi(k-1)} \qquad (3.5)$$

where $\tilde{P}(k) = \hat{P}(k) - P$,

$\hat{P}(k) = [\hat{A}(k), \hat{B}(k)]$

and $P = [A, B]$.

Since the matrix C does not appear in equation (3.5), the convergence of this equation which is the main concern here, does not depend on whether $C = 0$ or $\neq 0$. In simulation studies, $C = 0$ is the most convenient choice.

In proving the stability of equation (3.5), for the sake of convenience, we institute the following change of variables:

$$\psi(k) = \phi(k-1)/||\phi(k-1)|| \tag{3.6}$$

where $||.||$ denotes the Euclidean norm.
Equation (3.5) reduces to:

$$\tilde{P}(k+1) = \tilde{P}(k) - \alpha\Sigma\tilde{P}(k)\psi(k)\psi(k)'/\lambda \tag{3.7}$$

Since P is a positive definite matrix, there exists a nonsingular matrix T such that

$$T\Sigma T^{-1}/\lambda = \Lambda \tag{3.8}$$

where Λ is a diagonal matrix whose diagonal elements λ_i lie in the range $0 < \lambda_i \leq 1$

Premultiplying equation (3.7) by T and using equation (3.8), we get:

$$\tilde{P}(k+1) = \tilde{P}(k) - a\Lambda\tilde{P}(k)\psi(k)\psi(k)' \qquad (3.9)$$

where $\tilde{P}(k) = T\tilde{P}(k)$ (3.10)

We can partition equation (3.10) row-wise as below:

$$\tilde{\rho}_i(k+1) = \tilde{\rho}_i(k) - a\lambda_i\psi(k)\psi(k)'\tilde{\rho}_i(k), \quad i\epsilon[1,n] \qquad (3.11)$$

where $\tilde{\rho}_i(k)'$ is the i-th row vector of the nx(n+r) matrix $\tilde{P}(k)$, λ_i is the i-th diagonal element of Λ

and $0 < a\lambda_i < 2$.

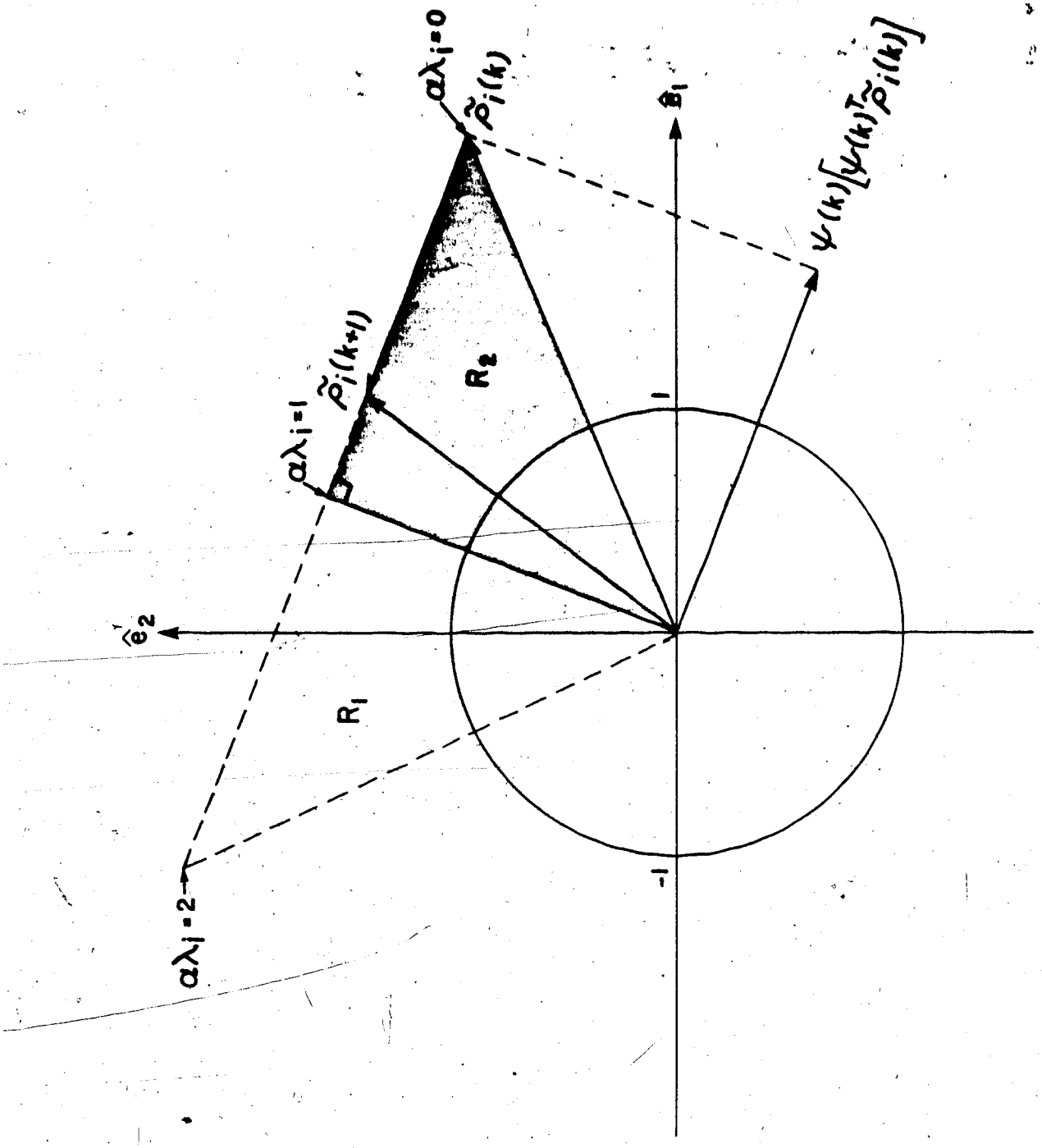
Since

$$\tilde{\rho}_i(k)'\psi(k) = ||\tilde{\rho}_i(k)||\cos\theta_i, \qquad (3.12)$$

the choice $a\lambda_i = 1$ makes $||\tilde{\rho}_i(k+1)||$ minimal for a given $\psi(k)$ where θ_i is the angle between $\tilde{\rho}_i(k)$ and $\psi(k)$.

If we choose $a = 1$ for convenience, λ_i must be 1 and $\Lambda = I$. Furthermore, if we choose $T = I$ for convenience then

$$a\Sigma/\lambda = I$$



3.1 The geometric effect of gain

Fig. 3-1 shows the geometric interpretation of equation (3.11). Now that we have obtained the optimal gain $\alpha\Sigma/\lambda = I$, we can go back to the original equation (3.7). With the choice of $\Sigma = I$ and $\alpha = 1$, equation (3.7) can be written as:

$$\tilde{p}_i(k+1) = \tilde{p}_i(k) - \psi(k)\psi(k)'\tilde{p}_i(k), \quad i\epsilon[1,n] \tag{3.13}$$

where $\tilde{p}_i(k)'$ is the i-th row vector of $\tilde{P}(k)$.

We can see that for a given $\psi(k)$ the best possible scalar gain at each step is $\psi(k)'\tilde{p}_i(k)$.

The global asymptotic stability of equation (3.13) is guaranteed by the fact that condition (iv) in the K-N scheme implies that any n+r consecutive vectors $\phi(k)$ span the n+r dimensional Euclidean space, i.e.

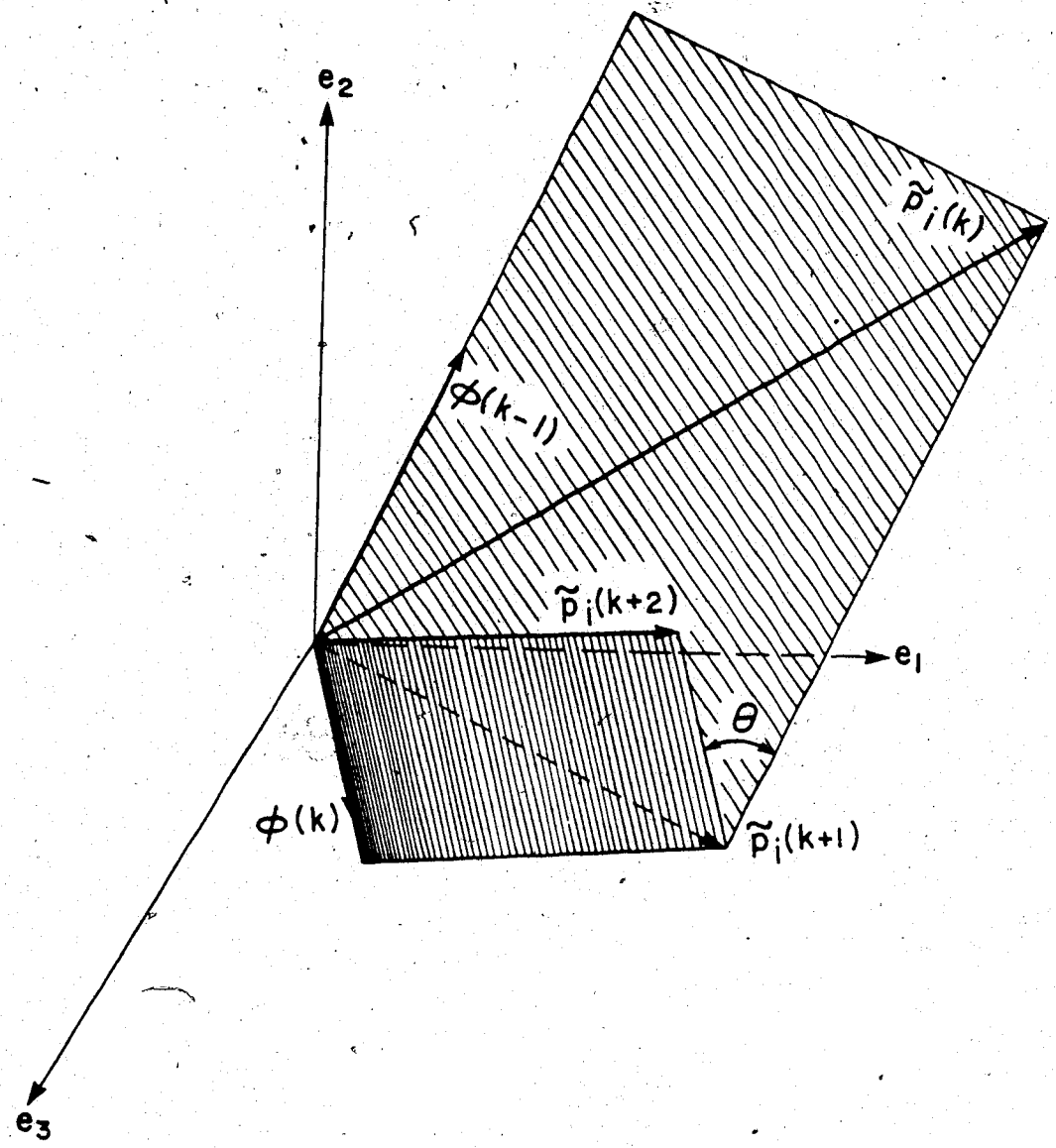
$[\phi(k),\phi(k+1),\ldots,\phi(k+n+r-1)]$ is nonsingular.

The implication of equation (3.13) is shown geometrically in Figure 3-2 for third order $\phi(k)$'s.

Keeping in mind the directions of $\phi(k-1)$ and $\psi(k)$ are the same, observation of Fiure 3-2 reveals the following properties of $\tilde{p}_i(k)$:

a) $\tilde{p}_i(k+1)$ is orthogonal to $\psi(k)$,

b) $\tilde{p}_i(k+1)$ lies in the plane made up of $\tilde{p}_i(k)$ and $\psi(k)$



3.2 The geometric effect of nonaligned vectors

c) if $\tilde{p}_i(k)$ is parallel to $\psi(k)$ then $\tilde{p}_i(k+1)$ is zero, and d) $||\tilde{p}_i(k+1)|| < ||\tilde{p}_i(k)||$.

3.3 A Rapidly Convergent Algorithm (Kim, et al., 1981)

While equation (3.13) maximizes the decrease of the parameter error $\tilde{p}_i(k)$ at each step and thus $\tilde{p}_i(k)$ is assured of convergence to the origin, the arbitrariness in the direction of $\psi(k)$ retards the speed of convergence $\tilde{p}_i(k)$ to the origin. In order to circumvent this difficulty, we introduce an orthogonal sequence of $\hat{\psi}(k)$ in the place of $\psi(k)$ to coordinate systematically the direction in which $\tilde{p}_i(k)$ decreases and adjust the scalar gain $\alpha(k)$ in such a way as to make $\tilde{p}_i(k+1)$ orthogonal to $\hat{\psi}(k)$ (This adjustable gain $\alpha(k)$ should not be confused with the fixed parameter α used in Section 3.2). Then it is not difficult to see from the properties a) and b) that this n+r vector can be made to be parallel to $\hat{\psi}(k)$ in n+r-1 steps. If this is done, then by property c), $\tilde{p}_i(k)$ will converge to the origin in one more step. Thus the total number of steps required to bring $\tilde{p}_i(k)$ to the origin will be utmost n+r.

Another way of explaining this is as follows. We choose $\psi(0)$ as the first basis vector of the n+r Euclidean space, $\hat{\psi}(0) = \psi(0)$. We select $\hat{\psi}(k)$ as the remaining basis vectors. We also choose $\alpha(k)$ to maximize $||\tilde{p}_i(k)|| - ||\tilde{p}_i(k+1)||$ at each step so that the number of steps required to bring $\tilde{p}_i(k)$ to the origin will be a minimum.

In the light of the above discussion, we can use the following algorithm in place of equation (3.13).

$$\tilde{p}_i(k+1) = \tilde{p}_i(k) - \alpha(k)\hat{\psi}(k)\phi(k-1)'\tilde{p}_i(k), \quad i\epsilon[1,n] \qquad (3.14)$$

The choice of $\alpha(k)$ to maximize $||\tilde{p}_i(k)|| - ||\tilde{p}_i(k+1)||$ is made as follows.

If an orthogonal sequence $\hat{\psi}(k)$ is generated from the sequence of $\phi(k)$ by the Gram-Schmidt process, the plane made up of $\hat{\psi}(k)$ and $\tilde{p}_i(k)$, and the plane made up of $\hat{\psi}(k)$ and $\phi(k-1)$ are orthogonal to each other (see Figure 3-3). $\tilde{p}_i(k)$ can be expressed as:

$$\tilde{p}_i(k) = \tilde{p}_{i\psi}(k) + \tilde{p}_{i\psi}^{\perp}(k)$$

where $\tilde{p}_{i\psi}(k)$ is a projection of $\tilde{p}_i(k)$ on $\hat{\psi}(k)$ and $\tilde{p}_{i\psi}^{\perp}(k)$ is a projection of $\tilde{p}_i(k)$ on a vector orthogonal to both $\hat{\psi}(k)$ and $\phi(k-1)$. Since $\phi(k-1)$ is orthogonal to $\tilde{p}_{i\psi}(k)$ and $\hat{\psi}(k)'\hat{\psi}(k) = 1$ because of orthonormality, we can write

$$\begin{aligned}\phi(k-1)'p_i(k) &= \phi(k-1)'\tilde{p}_{i\psi}(k) \\ &= \phi(k-1)'[\hat{\psi}(k)\hat{\psi}'(k)\tilde{p}_i(k)] \\ &= [\phi(k-1)'\hat{\psi}(k)][\hat{\psi}(k)'\tilde{p}_i(k)] \qquad (3.15)\end{aligned}$$

Using equation (3.15), we can write equation (3.14) as:

$$\tilde{p}_i(k+1) = \tilde{p}_i(k) - a(k)\psi(k)[\phi(k-1)'\hat{\psi}(k)]\hat{\psi}(k)'\tilde{p}_i(k) \quad (3.16)$$

for $i\epsilon[1,n]$

Following equation (3.13), it has been pointed out that $\psi(k)'\tilde{p}_i(k)$ is the best possible scalar gain for a given $\psi(k)$. By analogy, we see that the choice of

$$a(k) = 1/[\hat{\psi}(k)'\phi(k-1)] \quad (3.17)$$

in equation (3.16) will make $\hat{\psi}(k)'\tilde{p}_i(k)$ the best possible gain for a given $\hat{\psi}(k)$.

The resulting equation:

$$\tilde{p}_i(k+1) = \tilde{p}_i(k) - \frac{\hat{\psi}(k)\phi(k-1)'\tilde{p}_i(k)}{\phi(k-1)'\hat{\psi}(k)} \quad i\epsilon[1,n] \quad (3.18)$$

is basically the same as the equation (3.13) except that the direction of the sequence of $\tilde{p}_i(k+1) - \tilde{p}_i(k)$ is now orthogonal to that of all preceeding elements, which enables $\tilde{p}_i(k)$ to converge to the origin in utmost n+r steps.

The composite equation for all i, $i\epsilon[1,n]$ would be:

$$\tilde{P}(k+1) = \tilde{P}(k) - \frac{\tilde{P}(k)\phi(k-1)\hat{\psi}(k)'}{\hat{\psi}(k)'\phi(k-1)} \quad (3.19)$$

Thus the resulting algorithm is:

$$P(k+1) = P(k) - \frac{e(k)\hat{\psi}(k)'}{\hat{\psi}(k)'\phi(k-1)} \qquad (3.20)$$

Figure 3-3 shows the goemetric implication of equation (3.18).

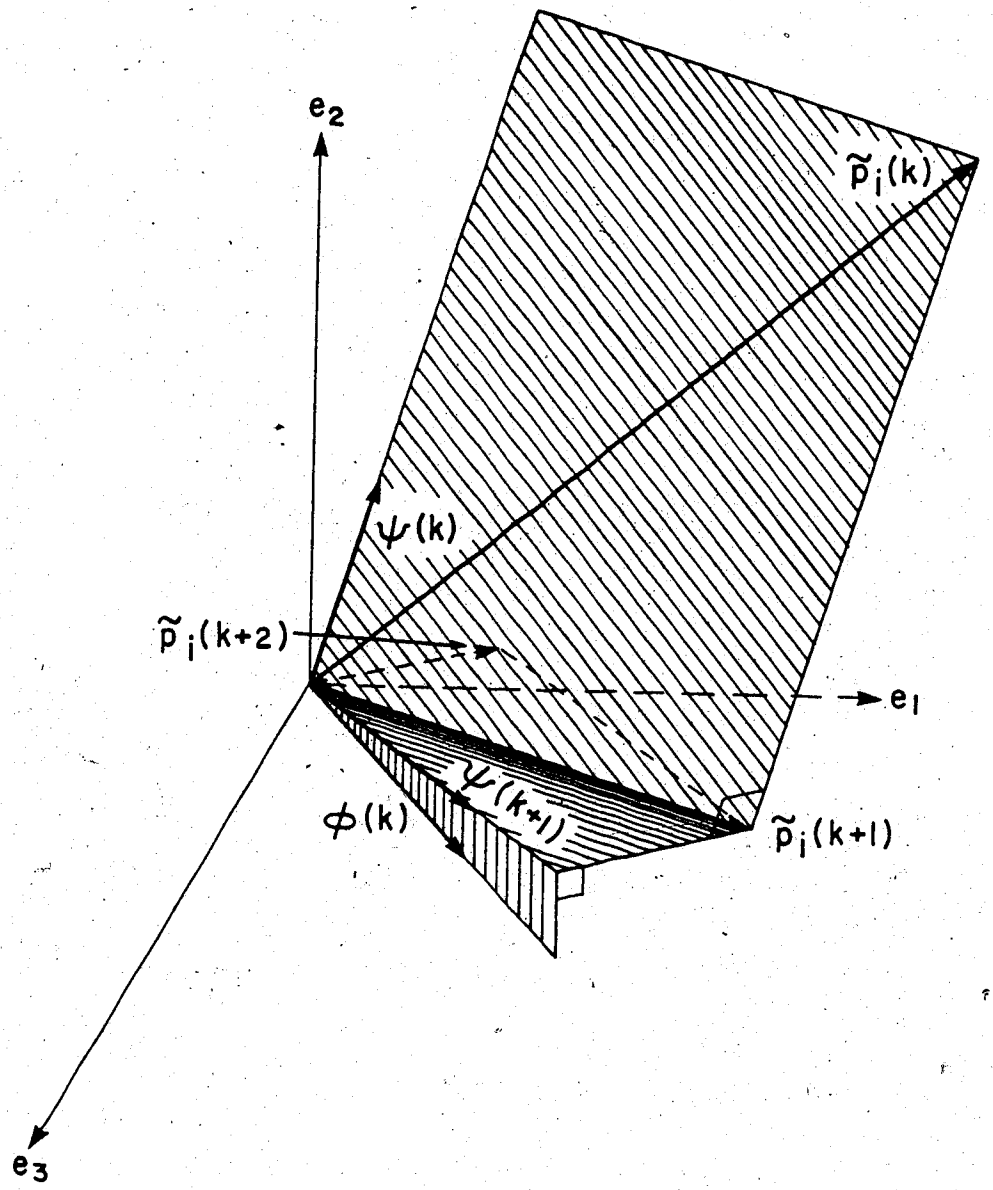
The algorithm proposed by Westphal (1978) yields the same results as the algorithm (3.20) on if p in the Westphal's scheme is set equal to n+r-1. However, the emphasis placed by Westphal on the utilization of à priori information about the system parameters to improve the the convergence unnecessarily obscures the real strength of the algorithm. As the simulation results show in the next section, à priori information is not at all necessary for ensuring rapid convergence of the algorithm.

This algorithm is also readily applicable to a deterministic MIMO system whose present output can be described as an inner product of the parameters and a sequence of the past outputs and inputs.

3.4 Simulations

Using the algorithm (3.20), a computer simulation is first carried out to identify the parameters of the dynamical discrete system:

$$x(k+1) = Ax(k) + Bu(k)$$



3.3 The geometric effect of orthonormal vectors

where A and B are the unknown but constant matrices. For the purpose of simulation, a fourth order discrete system with the following values for A and B matrices:

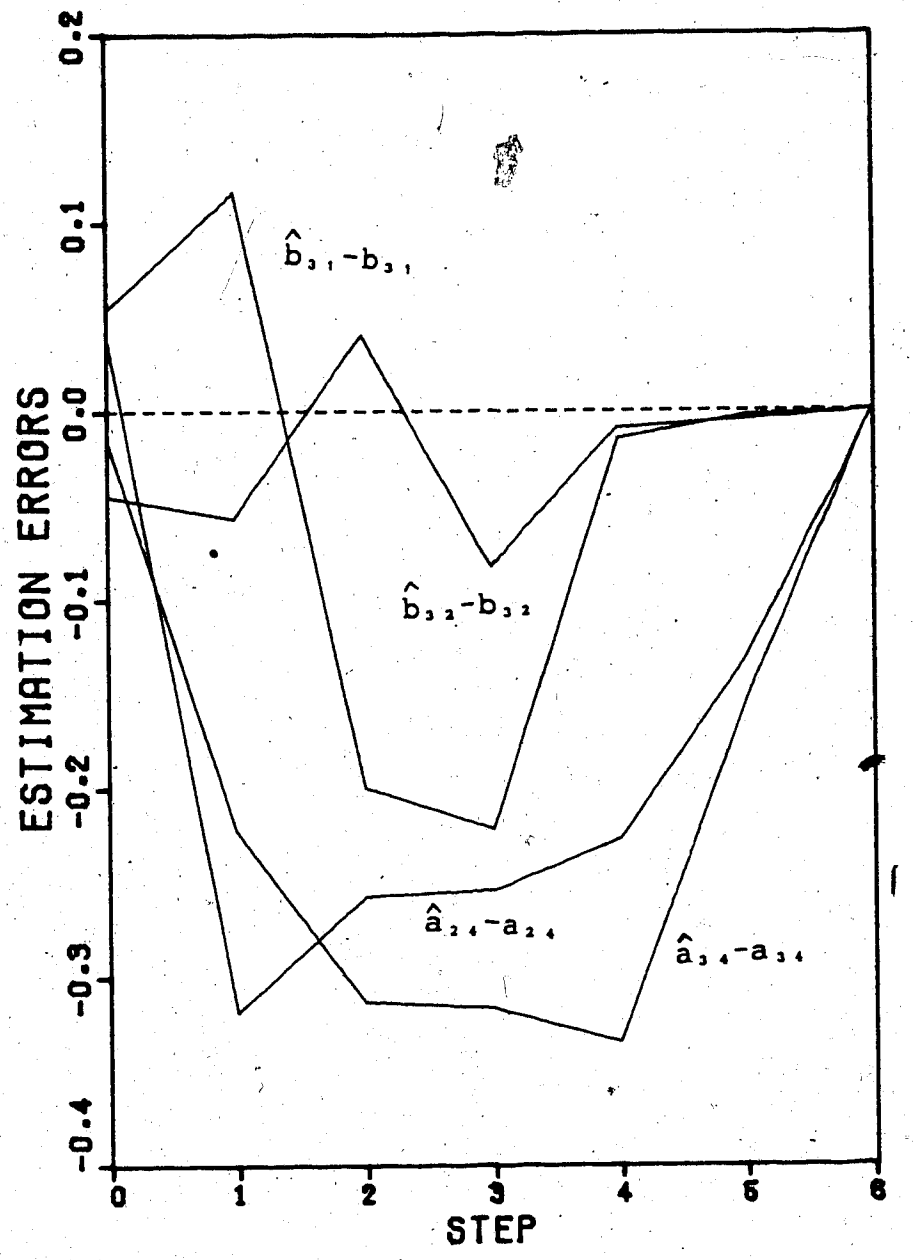
$$A = \begin{bmatrix} 9.9963E\text{-}1 & 2.6990E\text{-}4 & 1.6457E\text{-}4 & -4.5584E\text{-}3 \\ 4.7943E\text{-}4 & 9.8995E\text{-}1 & -1.7606E\text{-}4 & -4.0008E\text{-}2 \\ 9.9919E\text{-}4 & 3.6498E\text{-}3 & 9.9303E\text{-}1 & 1.4074E\text{-}2 \\ 5.0006E\text{-}6 & 1.8301E\text{-}5 & 9.9650E\text{-}3 & 1.0001E\text{-}0 \end{bmatrix}$$

$$B = \begin{bmatrix} 4.4212E\text{-}3 & 1.7543E\text{-}3 \\ 3.5272E\text{-}2 & -7.5542E\text{-}2 \\ -5.4940E\text{-}2 & 4.4605E\text{-}2 \\ -2.7513E\text{-}4 & 2.2351E\text{-}4 \end{bmatrix}$$

is chosen to demonstrate the effectiveness of the algorithm. This system is actually obtained by discretizing the continuous system in (Narendra, et al., 1973; Luders, et al, 1974) with sampling period 0.01s. Here the initial values of $\hat{A}(k)$ and $\hat{B}(k)$ are chosen to be all zero and the matrix C in equation (3.2) is conveniently set to be zero. A part of the results of the simulation is shown in Figure 3-4.

Another simulation is carried out in order to show the usefulness of the proposed algorithm to tackle a more realistic problem which was generated by assumption that only one state variable is measurable in the above example.

$$x(k+1) = Ax(k) + Bu(k)$$



3.4 The results of the first simulation

$$y(k) = [0\ 0\ 0\ 10]x(k)$$

This system is then equivalent to the system in equation (2.1) in terms of the input-output relation as:

$$x(k+1) = Ax(k) + b_1u_1(k) + b_2u_2(k)$$

$$y(k) = h'x(k)$$

where

$$A = \left[\begin{array}{c|c} a & \begin{matrix} I \\ \hline 0 \end{matrix} \end{array}\right]$$

$$a = [3.9827\ \ -5.9481\ \ 3.9480\ \ -9.8265]'$$

$$b_1 = [-2.751\ \ 2.738\ \ 2.734\ \ -2.721]' \times 10^{-3}$$

$$b_2 = [2.235\ \ -2.235\ \ -2.216\ \ 2.216]' \times 10^{-3}$$

$$h = [1\ 0\ 0\ 0]'$$

Following the procedure discussed in Section 2.2 to represent the above system in a more convenient form for the identification problem, we obtain:

$$z_1(k+1) = F_0'z_1(k) + hy(k)$$

$$z_2(k+1) = F_0'z_2(k) + hu_1(k)$$

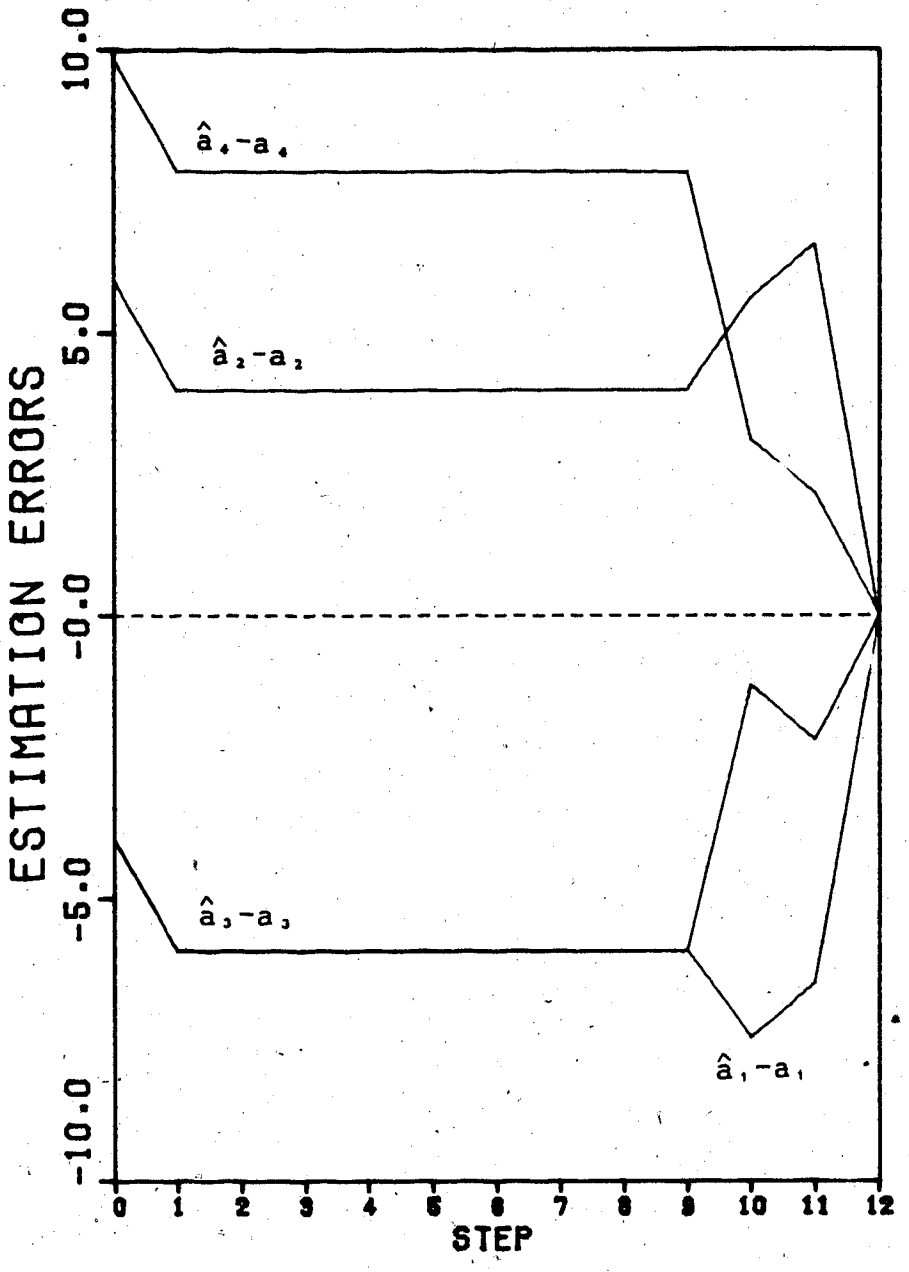
$$z_3(k+1) = F_0'z_3(k) + hu_2(k)$$

$$y(k) = a'z_1(k) + b_1'z_2(k) + b_2'z_3(k)$$

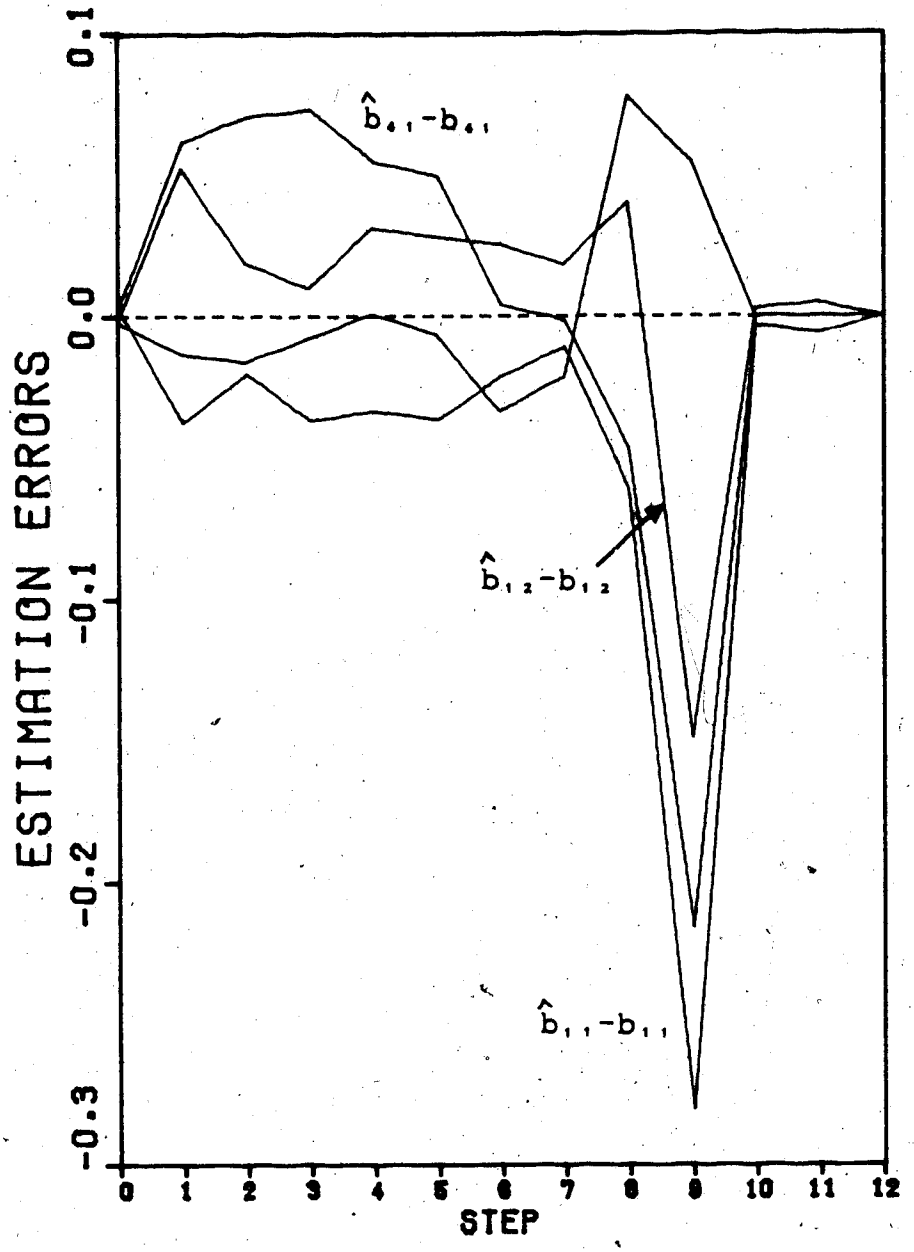
where F_0 is described in equation (2.3) with $f = 0$.

Then, we can readily apply the identification algorithm

proposed in Section 3.3. The simulation results are shown in Figure 3-5.



3.5 The estimates for vector a



3.6 The estimates for vector b

CHAPTER IV

ALGORITHMS FOR IDENTIFICATION OF STOCHASTIC SYSTEMS

4.1 Introduction

In this chapter we shall develop some new algorithms for the identification of the stochastic system described in Chapter II. Since the system is stochastic, it is also natural that we consider the use of stochastic estimation techniques in this development. It was pointed out in that chapter that the augmentation of the state vector of the stochastic system with the unknown parameters makes the error system nonlinear and consequently we have a nonlinear estimation problem of a stochastic system.

The main objective of stochastic estimation problems is to obtain the conditional mean values and covariance matrix of the unknown quantities. The mean is the estimate with the minimum variance for the given measured data and the covariance matrix represents the measure of the uncertainty in the estimates. It is well known that the conditional mean and covariance matrix cannot be, in general, characterized by a finite set of the moments of the conditional density function.

The involvement of solutions of functional integral difference equations (Bucy, et al., 1968) makes the computation of the exact conditional density function, in general, virtually impossible. Thus, in practice,

approximations of the conditional density function are used and this results in suboptimal filters. it is rather desirable that the conditional density function during the process of approximation is represented by a finite set of parameters so that the corresponding nonlinear filter is made up of finite number of equations of evolution for these parameters. Unfortunately this simplification cannot be achieved without the loss of sufficient statistics on the estimates.

It is known that in the linear estimation problem the conditional density function, being Gaussian, can completely determine the mean vector and covariance matrix. In this case, the mean vector is usually treated as the state of a linear filter. This filter, in turn, can be looked upon as a reference model in a model tracking context. The covariance matrix is used to compute the direction and magnitude of the gain vector. In order to take advantage of this useful feature of the linear estimation problem, the nonlinear estimation problem is linearized by taking only the linear part of the Taylor series representation of the nonlinear system. If this is done, then, the Kalman estimation theory can be readily applied to the approximated system because at each step only linear terms are involved. The resulting filter is called the Extended Kalman filter (EKF), which has been most widely used in real-life applications (Jazwinski, 1970).

It is known that the convergence of the estimates of the EKF algorithm to the true values is assured if the estimates stay within a small neighborhood of the true values. In the case of the MISO system described in Chapter II, the augmented state vector consists of the state vector and the parameter vector. If the parameter estimates stay within a small neighborhood of the true values, generation of good estimates of the state vector automatically follows in the innovations representation. There is no need to estimate the state vector separately. However, use of the standard EKF algorithm requires that the estimates of the entire augmented state vector be determined at each step. This unnecessarily increases the computational burden. One of objectives in this chapter is to develop certain modifications of the EKF algorithm to reduce this computational burden to facilitate on-line applications.

Stability is perhaps the most important consideration in assesing the performance of any techniques of recursive adaptive schemes. In the deterministic system, a number of methods are available to test or analyze the stability of recursive adaptive techniques, for example, Lyapunov method and the hyperstability criterion. Among these, the hyperstability criterion, which can be especially regarded as a specific formulation of the Lyapunov method, offers the most systematic approach to system design. The application of the hyperstability criterion to a stochastic

identification problem was first pioneered by Landau (1976) and became a popular tool to design a recursive adaptive scheme.

In Section 4.2 the application of the EKF algorithm to the identification problem is discussed. In the remaining sections of the chapter, several modified EKF algorithms for the stochastic identification problem are developed and certain similarities of these algorithms with earlier algorithms (Panuska, 1968: 1969; Soderstrom, 1973) are shown.

4.2 EKF Algorithm for Identification Problem

In this section, our aim is to derive EKF equations for the identification of the system described in Chapter II. The derivation follows the approach in (Anderson, et al., 1979).

$$\theta(k+1) = \Phi\theta(k) + \eta v(k) \qquad [2.6a]$$

$$y(k) = [0\ 0\ z_2(k)']\theta(k) + 1/2\theta(k)'S\theta(k) + v(k) \qquad [2.6b]$$

where

$$\Phi = \begin{bmatrix} F' & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \qquad \eta = \begin{bmatrix} h \\ 0 \\ 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 0 & I & 0 \\ I & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$F = \begin{bmatrix} f & \frac{I}{0} \end{bmatrix}$$

and 0's are understood as vectors or matrices of an appropriate dimension.

The nonlinear function, which appears only in the observation system (2.6b), can be expanded in Taylor series as:

$$y(k) = y(k) + \psi(k)'\tilde{\theta}(k) + 1/2\tilde{\theta}(k)'S\tilde{\theta}(k) + v(k) \qquad (4.1)$$

where $\hat{y}(k) = \hat{\theta}_1(k)'\hat{\theta}_2(k) + z_2(k)'\hat{\theta}_3(k)$ (4.2)

$$\psi(k)' = \left.\frac{\partial y(k)}{\partial \theta(k)}\right|_{\theta = \hat{\theta}(k)} = [\hat{\theta}_2(k)', \hat{\theta}_1(k)', z_2(k)'] \qquad (4.3)$$

$$\tilde{\theta}(k) = \theta(k) - \hat{\theta}(k) \qquad (4.4)$$

and $\tilde{\theta}(k) = [\tilde{\theta}_1(k)', \tilde{\theta}_2(k)', \tilde{\theta}_3(k)']'$ is an estimate of $\theta(k)$ based on the observations up to k-1 step.

Neglecting the second order term in equation (4.1), we obtain a linearized approximate version of the system (2.6) as:

$$\theta(k+1) = \Phi\theta(k) + \eta v(k) \tag{4.5a}$$

$$y(k) = \phi(k)'\tilde{\theta}(k) + v(k) + \hat{y}(k) \tag{4.5b}$$

Then, the application of the Kalman estimation method to equation (4.5) yields the following algorithm:

$$\hat{\theta}(k+1) = \Phi\hat{\theta}(k) + g(k)[y(k) - \hat{y}(k)] \tag{4.6a}$$

$$\hat{y}(k) = \hat{\theta}_1(k)'\hat{\theta}_2(k) + z_2(k)'\hat{\theta}_3(k) \tag{4.6b}$$

$$g(k) = [\Phi P(k)\psi(k) + \eta w]m(k) \tag{4.6c}$$

$$P(k+1) = \Phi P(k)\Phi' - g(k)m(k)^{-1}g(k)' + \eta w\eta' \tag{4.6d}$$

$$P(0) = P(0)' > 0$$

$$m(k) = [\psi(k)'P(k)\psi(k) + w]^{-1} \tag{4.6e}$$

where w is the covariance of $v(k)$ and $P(k)$ is the approximate covariance matrix of EKF.

$g(k)$ and $P(k)$ can be also, after a matrix manipulation (see APPENDIX 4A), expressed as:

$$g(k) = \Phi_n(k)P(k)\psi(k)m(k) + \eta \tag{4.7}$$

$$P(k+1) = \Phi_n(k)[P(k) - P(k)\psi(k)m(k)\psi(k)'P(k)]\Phi_n(k)' \tag{4.8}$$

where $\Phi_n(k) = \Phi - \eta\psi(k)'$ (4.9)

The rearrangement of $g(k)$ and $P(k)$ makes it clear that $P(k)$ goes to zero as $k \to \infty$ if $\Phi_n(k)$ is stable for all k. In this case $g(k)$ approaches η.

The error state $\theta(k)$ satisfies the difference equation:

$$\begin{aligned}
\tilde{\theta}(k+1) &= \Phi\tilde{\theta}(k) - g(k)[y(k) - \hat{y}(k)] + \eta v(k) \\
&= \Phi\tilde{\theta}(k) - \Phi_n(k)P(k)\psi(k)m(k)[\psi(k)'\tilde{\theta}(k) \\
&\quad + 1/2\tilde{\theta}(k)'S\tilde{\theta}(k) + v(k)] \\
&\quad - \eta[\psi(k)'\tilde{\theta}(k) - 1/2\tilde{\theta}(k)'S\tilde{\theta}(k)] \\
&= \Phi_n(k)[I - P(k)\psi(k)m(k)\psi(k)']\tilde{\theta}(k) \\
&\quad - 1/2[\Phi_n(k)P(k)\psi(k)m(k) + \eta]\tilde{\theta}(k)'S\tilde{\theta}(k) \\
&\quad - \Phi_n(k)P(k)\psi(k)m(k)v(k) \qquad (4.10)
\end{aligned}$$

As clearly seen in equation (4.10), the effect of the nonlinear term $\tilde{\theta}(k)'S\tilde{\theta}(k)$ is obvious. The convergence of the error state $\tilde{\theta}(k)$ to the origin could be thus assured, only in a neighborhood of the origin of the error state space where the effect of thenonlinearity is negligible.

Another advantage of employing the innovations representation, besides the reduction in number of parameters to characterize noise, is that the solution to a

Riccati equation in order to obtain the steady state Kalman filter gain can be avoided (Astrom, et al., 1971). This implies that the noise, thus the exact state vector, in the innovations representation could be easily generated as the error between the outputs of the system and the filter if the system parameters are known. In the identification problem under consideration, the estimate $\hat{\theta}_1(k)$ can be assumed to be equal to the state vector $\theta_1(k)$ (thus $z_1(k)$) in the neighborhood of the origin of the parameter error space, where the EKF is assured of convergence to a true values of the parameters.

4.3 First Modification of the EKF Algorithm (EKF-M1)

The approximate covariance matrix P(k) can be partitioned as:

$$P(k) = \begin{bmatrix} P_{11}(k) & P_{12}(k) \\ P_{12}(k)' & P_{22}(k) \end{bmatrix} \begin{matrix} \}n \\ \}2n \end{matrix}$$

Then, we obtain from equation (4.6d):

$$\begin{aligned} P_{11}(k+1) = {} & F_n(k)'P_{11}(k)F_n(k) - h\phi(k)'P_{12}(k)'F_n(k) \\ & - F_n(k)'P_{12}(k)\phi(k)h' + h\phi(k)'P_{22}(k)\phi(k)h' \\ & - [F_n(k)'P_{11}(k)\hat{\theta}_2(k) - h\phi(k)'P_{12}(k)'\hat{\theta}_2(k) \\ & - F_n(k)'P_{12}(k)\phi(k) - h\phi(k)'P_{22}(k)\phi(k)] \end{aligned}$$

$$m(k)[F_n(k)'P_{11}(k)\hat{\theta}_2(k) - h\phi(k)'P_{12}(k)\hat{\theta}_2(k)$$

$$- F_n(k)'P_{12}(k)\phi(k) - h\phi(k)'P_{22}(k)\phi(k)]' \tag{4.11}$$

$$P_{12}(k+1) = [F_n(k)'P_{12}(k) - h\phi(k)'P_{22}(k)]$$

$$- [F_n(k)'P_{11}(k)\hat{\theta}_2(k) - h\phi(k)'P_{12}(k)'\hat{\theta}_2(k)$$

$$- F_n(k)'P_{12}(k)\phi(k) - h\phi(k)'P_{22}(k)\phi(k)]$$

$$m(k)[P_{12}(k)'\hat{\theta}_2(k) + P_{22}(k)\phi(k)]' \tag{4.12}$$

$$P_{22}(k+1) = P_{22}(k) - [P_{12}(k)'\hat{\theta}_2(k) + P_{22}(k)\phi(k)]$$

$$m(k)[P_{12}(k)'\hat{\theta}_2(k) + P_{22}(k)\phi(k)]' \tag{4.13}$$

where $F_n(k) = F - \hat{\theta}_2(k)h'$ (4.14)

and $\phi(k) = [\hat{\theta}_1(k)', z_2(k)']'$ (4.15)

Suppose we assume that $\hat{\theta}_1(k)$ is nearly equal to $\theta_1(k)$, then $P_{11}(k)$ which is an approximate measurement of the uncertainity in $\hat{\theta}_1(k)$, can be set equal to 0, thus reducing the computational requirements. A convenient choice f = 0 considerably simplifies $P_{12}(k)$ and $P_{22}(k)$ for computational purpose as:

$$P_{12}(k+1) = [F_o' - h\hat{\theta}_2(k)']P_{12}(k) - h\phi(k)'P_{22}(k) \tag{4.16}$$

$$P_{22}(k+1) = P_{22}(k) - [P_{12}(k)'\hat{\theta}_2(k) + P_{22}(k)\phi(k)]$$

$$m_1(k)[P_{12}(k)'\hat{\theta}_2(k) + P_{22}(k)\phi(k)]' \quad (4.17)$$

where $F_o = F\Big|_{f = 0}$ (4.18)

and $m_1(k) = m(k)\Big|_{P_{11}(k) = 0}$ for all k

$$= [\phi(k)'P_{22}(k)\phi(k) + \phi(k)'P_{12}(k)'\hat{\theta}_2(k) + \hat{\theta}_2(k)'P_{12}(k)\phi(k) + w]^{-1} \quad (4.19)$$

It should be mentioned that a certain numerical problem associated with equations (4.16-17) will arise such that $P_{22}(k)$, which approximately represents the uncertainity in the estimates $\hat{\theta}_2(k)$ and $\hat{\theta}_3(k)$, could be nonpositive definite. In order to get around this difficulty, we simply add a positive scalar quantity μ to $m_1(k)^{-1}$ so that $P_{22}(k)$ remains positive definite for all k. Then this algorithm will be identical with the algorithm proposed by Ljung (1979). If this is done, define

$$U(k) \triangleq P_{12}(k)P_{22}(k)^{-1} \quad (4.20)$$

Thus, equations (4.16-17) can be rewritten as:

$$U(k+1)P_{22}(k+1)P_{22}(k)^{-1} = [F_o' - h\hat{\theta}_2(k)']U(k) - h\phi(k)' \quad (4.21)$$

$$P_{22}(k+1) = P_{22}(k) - P_{22}(k)[U(k)'\hat{\theta}_2(k) + \phi(k)]$$

$$m_1(k)[U(k)'\hat{\theta}_2(k) + \phi(k)]'P_{22}(k) \qquad (4.22)$$

Since in a small neighborhood of the true values of the parameter in $[\theta_2', \theta_3']'$ vector space, it can be assumed that

$$P_{22}(k+1)P_{22}(k)^{-1} \cong 1 \qquad (4.23)$$

Hence we can write $U(k)$ and $P_{22}(k)$ as:

$$U(k+1) = [F_0' - h\hat{\theta}_2(k)']U(k) - h\phi(k)' \qquad (4.24)$$

$$P_{22}(k+1) = P_{22}(k) - P_{22}(k)U(k+1)'hm_1(k)h'U(k+1)P_{22}(k) \qquad (4.25)$$

If we choose a positive scalar quantity μ such that

$$\mu = \hat{\theta}_2(k)'U(k)P_{22}(k)U(k)'\hat{\theta}_2(k) \qquad (4.26)$$

then we obtain

$$m_1(k) = [h'U(k+1)P_{22}(k)U(k+1)'h + w]^{-1} \qquad (4.27)$$

so that $P_{22}(k)$ is assured to be positive definite for all k. Hence the modified EKF becomes:

$$\hat{p}(k+1) = \hat{p}(k) + g_1(k)[y(k) - \hat{y}(k)] \tag{4.28a}$$

$$\hat{\theta}_1(k+1) = F_o'\hat{\theta}_1(k) + g_2[y(k) - \hat{y}(k)] \tag{4.28b}$$

$$\hat{y}(k) = \phi(k)'\hat{p}(k) \tag{4.28c}$$

$$g_1(k) = -P_{22}(k)U(k+1)'hm_1(k) \tag{4.28d}$$

$$g_2(k) = [F_o'U(k)P_{22}(k)\phi(k) + h]m_1(k) \tag{4.28e}$$

$$\hat{p}(k) = [\hat{\theta}_2(k)', \hat{\theta}_3(k)']' \tag{4.28f}$$

$P_{22}(k)$ and $U(k)$ are described in equations (4.24-25).

The salient feature of this algorithm, which we call the first modified EKF (EKF-M1) in the sequel, is the retention of $P_{12}(k)$ so as to preserve the local convergence property which is considered to be a very useful feature of EKF. However, since this modification of EKF is based on the assumption that the nonlinearity effect is negligible in a neighborhood of the origin of the parameter error space, it, in general, suffers from the lack of the global convergence property. However, under certain conditions, such global convergence is possible. The convergence of P(k), as shown in equation (4.8), completely depends on the stability of $\Phi_n(k)$, which, in turn, is determined by $\hat{\theta}_2(k)$. It is therefore necessary to ensure that $\hat{\theta}_2(k)$ lies in the region for which $\Phi_n(k)$ is stable. Such confinement of $\hat{\theta}_2(k)$ will

restrict the extent of observation nonlinearity. A recent study (Ljung, 1979) has shown that the modified EKF algorithm (4.28) when complemented by a facility to confine $\hat{\theta}_2(k)$ to the region where $F_o' - h\hat{\theta}_2(k)'$ is stable, will enable the parameter estimates to converge to the true values with probability 1 (w.p.1).

It will now be shown that the approximate RML algorithm (Soderstrom, 1973) can be derived from EKF-M1 as follows. Since $\hat{\theta}_1(k)$ is approximately equal to $\theta_1(k)$ in a neighborhood of the parameter error space, it follows that $\epsilon(k) = y(k) - \hat{y}(k)$ is equal to the noise term $v(k)$. This, in turn, suggests that $\hat{\theta}_1(k)$ can be obtained from equation (2.5a) using $\epsilon(k)$ instead of $v(k)$. If this is done, the algorithm turns out to be the same as the approximate RML algorithm with $w = 1$. That the approximate RML method possesses the local convergence property (Ljung, 1975) can be seen by noting that the gain $g_2(k)$ in equation (4.28e) is almost equal to h in a neighborhood of the origin of the parameter error space. However, the global convergence remains in question (Holst, 1977).

4.4 Second Modification of the EKF Algorithm (EKF-M2)

$P_{12}(k)$ represents the correlation between two error vectors, $\tilde{\theta}_1(k)$ and $\tilde{p}(k)$ where

$$\tilde{\theta}_1(k) = \theta_1(k) - \hat{\theta}_1(k) \tag{4.29}$$

$$\tilde{p}(k) = p(k) - \hat{p}(k) \tag{4.30}$$

$$p(k) = [\theta_2(k)' \ \theta_3(k)']' \tag{4.31}$$

The assumption that $\tilde{\theta}_1(k) = 0$, however, reduces the importance of $P_{12}(k)$, thus providing grounds for further simplification by setting $P_{12}(k) = 0$ (thus $U(k)$) in equation (4.28).

This results in the following algorithm:

$$\hat{p}(k+1) = \hat{p}(k) + g_3(k)[y(k) - \hat{y}(k)] \tag{4.32a}$$

$$\hat{\theta}_1(k+1) = F_0'\hat{\theta}_1(k) + hm_2(k)[y(k) - \hat{y}(k)] \tag{4.32b}$$

$$\hat{y}(k) = \phi(k)'\hat{p}(k) \tag{4.32c}$$

$$g_3(k) = P_{22}(k)\phi(k)m_2(k) \tag{4.32d}$$

$$P_{22}(k+1) = P_{22}(k) - P_{22}(k)\phi(k)m_2(k)\phi(k)'P_{22}(k) \tag{4.32e}$$

$$m_2(k) = [\phi(k)'P_{22}(k)\phi(k) + w]^{-1} \tag{4.32f}$$

In the sequel we shall refer to this algorithm as the second modified EKF algorithm (EKF-M2).

Following the same reasoning as before namely $\epsilon(k)$ could be assumed to be equal to $v(k)$, we use equation (2.5a) with $\epsilon(k)$ in the place of $v(k)$ to obtain $\hat{\theta}_1(k)$. Then the resulting algorithm with $w = 1$ in equation (4.32f) turns out to be the ELS method derived by Panuska (1968; 1969). The main drawback of the ELS method as well as the second

modified EKF algorithm lies in the requirement of additional restrictions on the system in order to ensure the convergence of the estimates to the true values.

4.5 The Convergence Property of the EKF-M2 Algorithm

In this section, the convergence property of the second modified EKF algorithm will be discussed using the Kalman-Yakubovich theorem which up to now has been used to study the stability of deterministic problems. Recall the brief discussion in section 2.3 that the identification problem can be formulated as a model tracking problem. In order to analyze the convergence property of EKF-M2 in the model tracking context, we shall examine the stability of the error model. We obtain the error states $\tilde{\theta}_1(k)$ and $\tilde{p}(k)$ satisfying the following equations with assumption $w = 1$ as:

$$\tilde{\theta}_1(k+1) = F_o'\tilde{\theta}_1(k) - hm_3(k)\epsilon(k) + hv(k) \tag{4.33a}$$

$$\tilde{p}(k+1) = \tilde{p}(k) - g_3(k)\epsilon(k) \tag{4.33b}$$

$$\begin{aligned}\epsilon(k) &= y(k) - \hat{y}(k) \\ &= [\theta_1(k)', z_2(k)']p(k) - \phi(k)'\hat{p}(k) + v(k) \\ &= \theta_2'\tilde{\theta}_1(k) + \phi(k)'\tilde{p}(k) + v(k)\end{aligned} \tag{4.33c}$$

where $m_3(k) = m_2(k)\Big|_{w = 1} = [1 + \phi(k)'P_{22}(k)\phi(k)]^{-1}$

Let $\epsilon_1(k) = m_3(k)\epsilon(k)$, then

$$\epsilon_1(k) = \theta_2'\tilde{\theta}_1(k) + \phi(k)'\tilde{p}(k) - \phi(k)'P_{22}(k)\phi(k)\epsilon_1(k) + v(k)$$

$$= \theta_2'\tilde{\theta}_1(k) + \gamma(k) + v(k) \qquad (4.34)$$

where $\gamma(k) = \phi(k)'\tilde{p}(k) - \phi(k)'P_{22}(k)\phi(k)\epsilon_1(k)$ (4.35)

Substituting equations (4.32d) and (4.34) into (4.33b), we obtain the following feedback system:

$$\tilde{\theta}_1(k+1) = [F_0' - h\tilde{\theta}_2']\theta_1(k) - h\gamma(k) \qquad (4.36a)$$

$$\epsilon_1(k) = \theta_2'\tilde{\theta}_1(k) + \gamma(k) + v(k) \qquad (4.36b)$$

$$\tilde{p}(k+1) = \tilde{p}(k) - P_{22}(k)\phi(k)\epsilon_1(k) \qquad (4.36c)$$

$$\gamma(k) = \phi(k)'\tilde{p}(k) - \phi(k)'P_{22}(k)\phi(k)\epsilon_1(k) \qquad (4.36d)$$

$$P_{22}(k+1) = P_{22}(k) - P_{22}(k)\phi(k)m_3(k)\phi(k)'P_{22}(k) \qquad (4.36e)$$

$$P_{22}(0) = P_{22}(0)' > 0$$

Since $\gamma(k)$ can be expressed as $\phi(k)'\tilde{p}(k+1)$, we get

$$\epsilon_1(k) = \theta_2'\tilde{\theta}_1(k) + \phi(k)'\tilde{p}(k+1) + v(k) \qquad (4.37)$$

Comparing the above equation with (4.33c) we can easily notice that $\epsilon_1(k)$ is the error after a correction is made on $\hat{p}(k)$. In the sequel $\epsilon_1(k)$ is referred to as the a posteriori error while $\epsilon(k)$ is the à priori error. The

condition for the convergence of the EKF-M2 algorithm to the true values will now be stated as below.

Theorem 1: Suppose that $F_o - h\theta_2'$ in equation (4.36a) has all its eigenvalues inside the unit circle, then the errors $\tilde{\theta}_1(k)$ and $\tilde{p}(k)$ satisfying equation (4.36) vanish as $k \to \infty$

$$\text{if} \qquad \frac{1}{1 - c_1 q^{-1} - \cdots - c_n q^{-n}} - \frac{1}{2} \qquad (4.38)$$

is strictly positive real (s.p.r.)

where $c' = [c_1, \cdots, c_n]$

and q is the shift operator.

The proof of Theorem 1 consists of two steps. First, we will prove using the Kalman-Yakubovich theorem that $\tilde{\theta}_1(k)$ and $\tilde{p}(k)$ converge to the origin without $v(k)$ in equation (4.36b). Then we prove that $v(k)$ actually has no effect on the convergence of $\tilde{\theta}_1(k)$ and $\tilde{p}(k)$ to the origin.

We need the following lemma.

Lemma: Given the following set of time-varying systems:

$$\tilde{\theta}_1(k+1) = F'\tilde{\theta}_1(k) - h\gamma(k) \qquad (4.39a)$$

$$e_1(k) = \theta_2'\tilde{\theta}_1(k) + \gamma(k) \qquad (4.39b)$$

$$\tilde{p}(k+1) = \tilde{p}(k) - P_{22}(k)\phi(k)e_1(k) \qquad (4.39c)$$

$$\gamma(k) = \phi(k)'\tilde{p}(k) - \phi(k)'P_{22}(k)\phi(k)e_1(k) \qquad (4.39d)$$

$$P_{22}(k+1) = P_{22}(k) - P_{22}(k)\phi(k)\dot{m}_3(k)\phi(k)'P_{22}(k) \qquad (4.39e)$$

$$P_{22}(0) = P_{22}(0)' > 0$$

the errors $\tilde{\theta}_1(k)$ and $\tilde{p}(k)$ converge to the origin as $k \to \infty$

if $$H(q^{-1}) = 1/2 - \theta_2'(qI - F_c')^{-1}h \qquad (4.40)$$

is s.p.r.

where $F_c = F_o - \theta_2 h'$ is assumed stable

proof: See APPENDIX 4B.

The above lemma is the discrete version of the Kudva-Narendra theorem (Lin, et al., 1978) with a modification that the gain matrix $P_{22}(k)$ decreases in magnitude as k increases. The decrease of $P_{22}(k)$ is necessary to account for the rejection of the noise in equation (4.36b) by gradually decreasing the weight on the noisy observations. It, however, requires a more stringent restriction on the system than is neccesary with a fixed $P_{22}(k)$ (Landau, 1979).

Resuming the proof of Theorem 1, note that the real positivity condition (4.38) is equivalent to $H(q^{-1})$ in (4.40) in that θ_2 is identical to -c with a choice f = 0 in

F matrix.

From equations (4.34-35) we have:

$$\epsilon_1(k) = [\theta_2'\tilde{\theta}_1(k) + \phi(k)'\tilde{p}(k) + v(k)]m_3(k) \qquad (4.41)$$

$$\gamma(k) = \{\phi(k)'\tilde{p}(k) - \phi(k)'P_{22}(k)\phi(k)[\theta_2'\tilde{\theta}_1(k) + v(k)]\}$$
$$m_3(k) \qquad (4.42)$$

Substituting (4.41-42) into (4.36), we have:

$$\chi(k+1) = \Xi(k)\chi(k) + \beta(k)v(k) \qquad (4.43)$$

where

$$\Xi(k) = \begin{bmatrix} \Xi_{11}(k) & \Xi_{12}(k) \\ \Xi_{21}(k) & \Xi_{22}(k) \end{bmatrix} \qquad (4.44)$$

$$\Xi_{11}(k) = F_c' + h\phi(k)'P_{22}(k)\phi(k)\theta_2'm_3(k) \qquad (4.45a)$$

$$\Xi_{12}(k) = -h\phi(k)'m_3(k) \qquad (4.45b)$$

$$\Xi_{21}(k) = -P_{22}(k)\phi(k)\theta_2'm_3(k) \qquad (4.45c)$$

$$\Xi_{22}(k) = I - P_{22}(k)\phi(k)\phi(k)'m_3(k) \qquad (4.45d)$$

$$\beta(k) = \begin{bmatrix} h\phi(k)'P_{22}(k)\phi(k)m_3 \\ -P_{22}(k)\phi(k)m_3(k) \end{bmatrix} \qquad (4.46)$$

$$\chi(k) = [\tilde{\theta}_1(k)', \tilde{p}(k)']' \qquad (4.47)$$

Hence, $\chi(k)$ can be expressed as:

$$\chi(k) = \left\{ \prod_{i=0}^{k-1} \Xi(i)\chi(0) + \sum_{j=0}^{k-2} \prod_{i=j+1}^{k-1} \Xi(i)\beta(j)v(j) \right.$$

$$\left. + \beta(k-1)v(k-1) \right. \tag{4.48}$$

Since $\Xi(k)$ is stable from Lemma and $v(k)$ is a white noise,

$$\lim_{k\to\infty} \prod_{i=0}^{k-1} \Xi(i) = 0 \tag{4.49}$$

$$\lim_{k\to\infty} E\{ \prod_{i=j+1}^{k-1} \Xi(i)\beta(j)v(j)\} = 0 \qquad \text{for } 0\le j\le k-2 \tag{4.50}$$

$$\lim_{k\to\infty} E\{\beta(k-1)v(k-1)\} = 0 \tag{4.51}$$

where $E(\cdot)$ denotes expectation.

Thus (Mendel, 1973),

$$\lim_{k\to\infty} E\{\chi(k)\} = 0 \tag{4.52}$$

we have proved Theorem 1 which, in turn, establishes the convergence of the EKF-M2 algorithm to the true values.

The difference between the ELS algorithm and EKF-M2 algorithm (4.32) is that in ELS à priori error $\epsilon(k)$ is used for the generation of $\hat{\theta}_1(k)$ while the EKF-M2 algorithm (4.32) uses à postériori error $\epsilon_1(k)$. Simulation studies show that this enables the EKF-M2 algorithm to converge more

rapidly than the ELS algorithm.

A limitation of the EKF-M2 algorithm, however, lies in the fact that the positive realness requirement (4.40) for unbiased convergence depends on the unknown parameter vector θ_2. A method which helps to overcome this limitation to some extent is discussed in the next section.

4.6 Third modification of the EKF (EKF-M3)

In the modifications of the EKF algorithm discussed so far, we have set the vector f in the matrix F equal to 0. In this section, the implication of using a general f, not neccesarily 0, will be discussed. The vector f will be chosen such that a stable polynominal of the difference between f and certain system parameter satisfies a positive realness condition.

If a general vector f is introduced in equation (4.32), the following algorithm is obtained:

$$z_2(k+1) = F'z_2(k) + hy(k) \tag{4.53a}$$

$$\hat{\theta}_1(k+1) = F'\hat{\theta}_1(k) + hm_3(k)[y(k) - \hat{y}(k)] \tag{4.53b}$$

$$\hat{p}(k+1) = \hat{p}(k) + P_{22}(k)\phi(k)m_3(k)[y(k) - \hat{y}(k)] \tag{4.53c}$$

$$\hat{y}(k) = \phi(k)'\hat{p}(k) \tag{4.53d}$$

$$\phi(k) = [\hat{\theta}_1(k)', z_2(k)']' \tag{4.53e}$$

$$P_{22}(k+1) = P_{22}(k) - P_{22}(k)\phi(k)m_3(k)\phi(k)'P_{22}(k) \qquad (4.53f)$$

This algorithm is referred to in the sequel as EKF-M3.

Convergence of the EKF-M3 algorithm

Recall from equation (2.5) that the system can be described as:

$$\theta_1(k+1) = F'\theta_1(k) + hv(k) \qquad (4.54a)$$

$$z_2(k+1) = F'z_1(k) + hy(k) \qquad (4.54b)$$

$$y(k) = [\theta_1(k)', z_2(k)']p + v(k) \qquad (4.54c)$$

where $p = [\theta_2', \theta_3']'$ (4.55)

$\theta_2 = f-c$

$\theta_3 = a-f$

We obtain the following set of error equations:

$$\tilde{\theta}_1(k+1) = F'\tilde{\theta}_1(k) - hm_3(k)\epsilon(k) + hv(k) \qquad (4.56a)$$

$$\tilde{p}(k+1) = \tilde{p}(k) - P_{22}(k)\phi(k)m_3(k)\epsilon(k) \qquad (4.56b)$$

$$\epsilon(k) = y(k) - \hat{y}(k)$$

$$= \theta_2'\tilde{\theta}_1(k)' + \phi(k)'\tilde{p}(k) + v(k) \qquad (4.56c)$$

The introduction of

$$\epsilon_1(k) = m_3(k)\epsilon(k) \tag{4.57}$$

$$\gamma(k) = \phi(k)'\tilde{p}(k) - \phi(k)P_{22}(k)\phi(k)\epsilon_1(k) \tag{4.58}$$

makes it clear that (4.56) constitues a block feedback system described in the following equations as:

$$\tilde{\theta}_1(k+1) = [F' - h\theta_2']\tilde{\theta}_1(k) - h\gamma(k) \tag{4.59a}$$

$$\epsilon_1(k) = \theta_2'\tilde{\theta}_1(k) + \gamma(k) + v(k) \tag{4.59b}$$

$$\tilde{p}(k+1) = \tilde{p}(k) - P_{22}(k)\phi(k)\epsilon_1(k) \tag{4.59c}$$

$$\gamma(k) = \phi(k)'\tilde{p}(k) - \phi(k)'P_{22}(k)\phi(k)\epsilon_1(k) \tag{4.59d}$$

$$P_{22}(k+1) = P_{22}(k) - P_{22}(k)\phi(k)m_3(k)\phi(k)'P_{22}(k) \tag{4.59e}$$

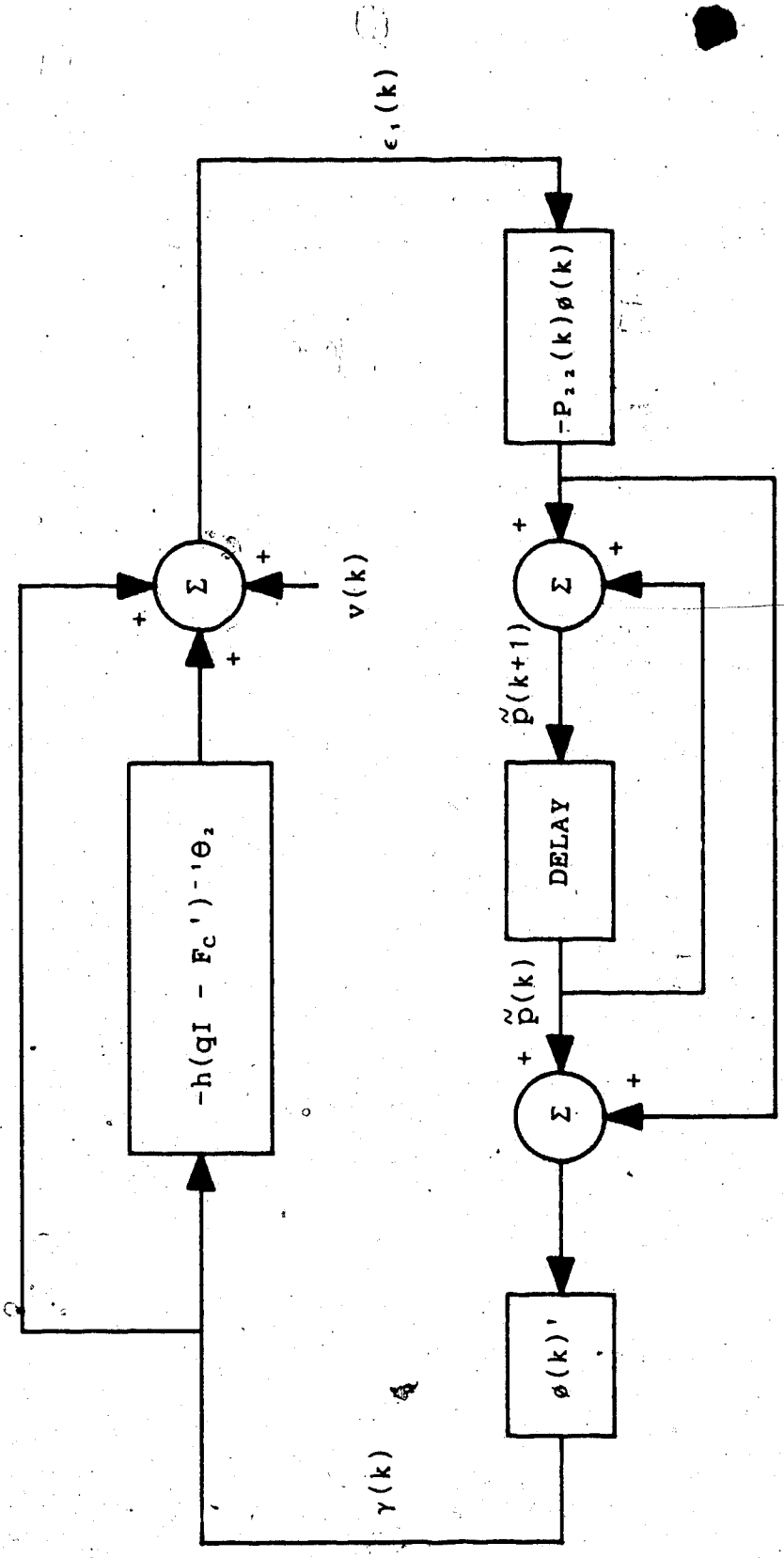
$$P_{22}(0) = P_{22}(0)' > 0$$

Fig. 4-1 shows the block diagram of the above feedback system.

A theorem similar to Theorem 1 can be stated as follows

Theorem 2: Suppose that $F' - h\theta_2'$ in equation (4.59a) has all its eigenvalues inside the unit circle, then the errors $\tilde{\theta}_1(k)$ and $\tilde{p}(k)$ satisfying equation (4.59) vanish as $k \to \infty$

if $$\frac{1 - f_1q^{-1} - \cdots - f_nq^{-n}}{1 - c_1q^{-1} - \cdots - c_nq^{-n}} - \frac{1}{2} \tag{4.60}$$



4.1 The error feedback system

is s.p.r.

proof: The proof is entirely analoguous to that of Theorem 1.

In recent years, the use of the approach utilizing the supermartingale theory to the convergence proof of recursive identification algorithms has appeared in the literature (Solo, 1979; Goodwin, et. al, 1979, Landau 1981: 1982). In this approach, it is necessary to find what is known as a corresponding supermartingale for the error feedback system. Since a stochastic Lyapunov function in the error feedback system should be positive and decrease in magnitude with time, it can be looked upon as a supermartingale. It is shown below that the use of this idea in the convergence proof for the EKF-M3 algorithm leads to a stronger result, namely, martingale convergence compared to the mean-square convergence obtained in the earlier proof.

Proof of Martingale Convergence

This proof should be read in conjunction with APPENDIX 4D where the definition and lemmas are given.

Since, as stated previously, $\tilde{\theta}_1(k)$ becomes zero as $\hat{p}(k)$ converges to p as a consequence of the use of innovations representation, we only need to prove that $\hat{p}(k)$ actually converges to p if the positive realness condition (4.60) holds.

Now, let s(k) be defined as:

$$s(k) \triangleq \tilde{p}(k)' P_{22}(k)^{-1} \tilde{p}(k) \tag{4.61}$$

where $\tilde{p}(k) = \hat{p}(k) - p$ as defined in (4.30)

Then

$$\begin{aligned} s(k+1) &= \tilde{p}(k+1)' P_{22}(k+1)^{-1} \tilde{p}(k+1) \\ &= \tilde{p}(k+1)' [P_{22}(k)^{-1} + \phi(k)\phi(k)'] \tilde{p}(k+1) \\ &= \gamma(k)^2 + [\tilde{p}(k) - P_{22}(k)\phi(k)e_1(k)]' P_{22}(k)^{-1} \\ &\quad [\tilde{p}(k) - P_{22}(k)\phi(k)e_1(k)] \\ &= \gamma(k)^2 + s(k) - 2e_1(k)\phi(k)'\tilde{p}(k) \\ &\quad + e_1(k)^2 \phi(k)' P_{22}(k)\phi(k) \\ &= \gamma(k)^2 + s(k) - 2e_1(k)\gamma(k) \\ &\quad - e_1(k)^2 \phi(k)' P_{22}(k)\phi(k) \end{aligned} \tag{4.62}$$

Taking conditional expectations in (4.62), we obtain

$$E\{s(k+1) + 2e_1(k)\gamma(k) \mid \mathcal{F}(k-1)\} \leq s(k) + \sigma(k)^2 \tag{4.63}$$

where $\sigma(k)^2 = E\{\gamma(k)^2 \mid \mathcal{F}(k-1)\}$ (4.64)

Since $e_1(k)$ and $\gamma(k)$ are the output and input signals of the strictly positive real system, the inequality

$$\sum_{i=1}^{k} e_1(i)\gamma(i) \geq 0 \quad \text{for all } k \tag{4.65}$$

always holds (Landau, 1969).

Defining a new martingale r(k)

$$r(k) = [s(k) + \sum_{i=1}^{k} e_1(i)\gamma(i)]/k-1 \qquad (4.66)$$

we have

$$E\{r(k+1)\lfloor\mathcal{F}(k-1)\} \leq r(k) - r(k)/k + \sigma(k)^2/k \qquad (4.67)$$

Since $\gamma(k)$ is the signal of the stable feedback system,

we have $\sum_{k=1}^{\infty} \sigma(k)^2 < \infty$

It follows from Lemma 5 in APPENDIX 4D that $r(k)$ converges to a finite random variable r with probabilty 1 (w.p.1) and

$$\sum_{k=1}^{\infty} r(k)/k < \infty \qquad \text{w.p.1} \qquad (4.68)$$

which makes $r = 0$ w.p.1.

Since $s(k)$ and $\sum_{i=1}^{k} e_1(i)\gamma(i)$ are positive, $s(k)/k$ and $e_1(i)\gamma(i)/k$ should converge to zero w.p.1.

Setting $T(k) = kP(k)$, we obtain

$$T(k+1) = T(k) - \frac{T(k)\phi(k)\phi(k)'T(k)}{k + \phi(k)'T(k)\phi(k)} + R(k) \qquad (4.69)$$

where

$$R(k) = P(k) - \frac{P(k)\phi(k)\phi(k)'P(k)}{k + \phi(k)'P(k)\phi(k)}$$

Since $y(k)$ is the output of the assumed stable system and $\epsilon(k)$ is the signal of the stable feedback system, $y(k)$ and $\epsilon(k)$ are bounded, so is $\phi(k)$. Hence $T(k)$ is upper bounded. This fact and $s(k)/k \to 0$ lead to the conclusion that

$$\lim_{k\to\infty} \tilde{p}(k) = 0 \quad \text{w.p.1.}$$

Since (4.60) is equivalent to (Siljak, 1968; 1970)

$$\frac{1}{1 - (c_1-f_1)q^{-1} - \cdots - (c_n- f_n)q^{-n}} - \frac{1}{2} \tag{4.70}$$

if the polynomials

$$F^*(q^{-1}) = 1 - f_1q^{-1} - \cdots - f_nq^{-n} \tag{4.71}$$

$$C^*(q^{-1}) = 1 - c_1q^{-1} - \cdots - c_nq^{-n} \tag{4.72}$$

are stable, the real strength of Theorem 2 is that the estimates $\hat{\theta}_1(k)$ and $\hat{p}(k)$ are assured of convergence to the true values if a vector f is chosen such that f-c i.e. θ_2, satisfies the condition (4.70) regardless of the vector c unlike in the case of the EKF-M2 algorithm (4.32). Theorem 2 defines the extent of the sensitivity of the algorithm (4.53) to the à priori knowledge of the parameters. In fact, the algorithm (4.32) could be considered as a special case of the algorithm (4.53) with f = 0.

APPENDIX 4A

$$g(k) = [\Phi P(k)\psi(k) + \eta w][\psi(k)'P(k)\psi(k) + w]^{-1}$$

$$= \{[\Phi - \eta\psi(k)']P(k)\psi(k) + \eta[\psi(k)'P(k)\psi(k) + w]\}$$
$$[\psi(k)'P(k)\psi(k) + w]^{-1}$$

$$= [\Phi - \eta\psi(k)']P(k)\psi(k)[\psi(k)'P(k)\psi(k) + w]^{-1} + \eta$$

$$= \Phi_n(k)P(k)\psi(k)m(k) + \eta$$

$$\Phi P(k)\Phi' - [\Phi P(k)\psi(k) + \eta w]m(k)[\psi(k)'P(k)\Phi' + w\eta'] + \eta w\eta'$$

$$= \Phi P(k)\Phi' - \Phi P(k)\psi(k)m(k)\psi(k)'P(k)\Phi' \qquad (A1.1)$$

$$- \Phi P(k)\psi(k)[\psi(k)'P(k)\psi(k) + w]^{-1}w\eta' \qquad (A1.2)$$

$$- w[\psi(k)'P(k)\psi(k) + w]^{-1}\psi(k)'P(k)\Phi' \qquad (A1.3)$$

$$- \eta w[\psi(k)'P(k)\psi(k) + w]^{-1}w\eta' + \eta w\eta' \qquad (A1.4)$$

$$= \Phi P(k)\Phi' - \Phi P(k)\psi(k)m(k)\psi(k)'P(k)\Phi' \qquad (A2.1)$$

$$+ \Phi P(k)\psi(k)[\psi(k)'P(k)\psi(k) + w]^{-1}$$
$$[-\psi(k)'P(k)\psi(k) - w + \psi(k)'P(k)\psi(k)]\eta' \qquad (A2.2)$$

$$+ \eta[-\psi(k)'P(k)\psi(k) - w + \psi(k)'P(k)\psi(k)]$$
$$[\psi(k)'P(k)\psi(k) + w]^{-1}\psi(k)'P(k)\Phi' \qquad (A2.3)$$

$$+ \eta\{[\psi(k)'P(k)\psi(k) + w - w][\psi(k)'P(k)\psi(k) + w]^{-1}w\}\eta' \qquad (A2.4)$$

$$= \Phi P(k)\Phi' - \Phi P(k)\psi(k)m(k)\psi(k)'P(k)\Phi' \qquad (A3.1)$$

$$+ \Phi P(k)\psi(k)[\psi(k)'P(k)\psi(k) + w]$$
$$[-\psi(k)'P(k)\psi(k) - w + \psi(k)'P(k)\psi(k)]\eta' \qquad (A3.2)$$

$$+ \eta[-\psi(k)'P(k)\psi(k) - w + \psi(k)'P(k)\psi(k)]$$
$$[\psi(k)'P(k)\psi(k) + w]^{-1}\psi(k)'P(k)\Phi' \qquad (A3.3)$$

$$+ \eta\psi(k)'P(k)\psi(k)[\psi(k)'P(k)\psi(k) + w]^{-1}w\eta' \qquad (A3.4)$$

$$= \Phi P(k)\Phi' - \Phi P(k)\psi(k)m(k)\psi(k)'P(k)\Phi' \qquad (A4.1)$$

$$+ \Phi P(k)\psi(k)[\psi(k)'P(k)\psi(k) + w]^{-1}$$
$$[-\psi(k)'P(k)\psi - w + \psi(k)'P(k)\psi(k)]\eta' \qquad (A4.2)$$

$$+ \eta[-\psi(k)'P(k)\psi(k) - w + \psi(k)'P(k)\psi(k)]$$
$$[\psi(k)'P(k)\psi(k) + w]^{-1}\psi(k)'P(k)\Phi' \qquad (A4.3)$$

$$+ \eta\psi(k)'P(k)\psi(k)[\psi(k)'P(k)\psi(k) + w]^{-1}$$
$$[\psi(k)'P(k)\psi(k) + w - \psi(k)'P(k)\psi(k)]\eta \qquad (A4.4)$$

$$= \Phi P(k)\Phi' - \Phi P(k)\psi(k)m(k)\psi(k)'P(k)\Phi' \qquad (A5.1)$$

$$+ \Phi P(k)\psi(k)\{-I + [\psi(k)'P(k)\psi(k) + w]^{-1}$$
$$\psi(k)'P(k)\psi(k)\}\eta' \qquad (A5.2)$$

$$+ \eta[-I + \psi(k)'P(k)\psi(k)[\psi(k)'P(k)\psi(k) + w]^{-1}\}$$
$$\psi(k)'P(k)\Phi(k) \qquad (A5.3)$$

$$+ \eta\psi(k)'P(k)\psi(k)\{I - [\psi(k)'P(k)\psi(k)+ w]^{-1}$$
$$\psi(k)'P(k)\psi(k)\}\eta' \qquad \text{(A5.4)}$$

$$= [\Phi - \eta\psi(k)']\{P(k) - P(k)\psi(k)m(k)\psi(k)'P(k)\}$$
$$[\Phi - \eta\psi(k)']' \qquad \text{(A6)}$$

$$= \Phi_n(k)\{P(k) - P(k)\psi(k)m(k)\psi(k)'P(k)\}\Phi_n(k)' \qquad \text{(A7)}$$

APPENDIX 4B

Proof of Lemma 1

The feedback system (4.39) can be described as a time-varying autonomous system:

$$\chi(k+1) = \Xi(k)\chi(k) \tag{B1}$$

where

$$\Xi(k) = \begin{bmatrix} \Xi_{11}(k) & \Xi_{12}(k) \\ \Xi_{21}(k) & \Xi_{22}(k) \end{bmatrix} \tag{B2}$$

$$\Xi_{11}(k) = F_c' + h\phi(k)'P_{22}(k)\phi(k)\theta_2'm_3(k) \tag{B3}$$

$$\Xi_{12}(k) = -h\phi(k)'m_3(k) \tag{B4}$$

$$\Xi_{21}(k) = -P_{22}(k)\phi(k)\theta_2'm_3(k) \tag{B5}$$

$$\begin{aligned}\Xi_{22}(k) &= I - P_{22}(k)\phi(k)\phi(k)'m_3(k) \\ &= [I + hP_{22}\phi(k)\phi(k)']^{-1}\end{aligned} \tag{B6}$$

$$\chi(k) = [\tilde{\theta}_1(k)', \tilde{p}(k)']' \tag{B7}$$

In equation (B6) we have used the matrix inversion lemma:

$$P_{22}(k+1)^{-1} = P_{22}(k)^{-1} + \phi(k)\phi(k)' \tag{B8}$$

Now, consider the following quadratic function as a candidate for the Lyapunov function:

$$V(k) = \chi(k)'T(k)\chi(k) \tag{B9}$$

where

$$T_{1,1}(k) = \Sigma \tag{B10}$$

$$T_{1,2}(k) = T_{2,1}(k) = 0 \tag{B11}$$

$$T_{2,2}(k) = P_{2,2}(k)^{-1} \tag{B12}$$

We obtain

$$\nabla V(k) = V(k+1) - V(k) = \chi(k)'\Omega(k)\chi(k) \tag{B13}$$

where $$\Omega(k) = \Xi(k)'T(k+1)\Xi(k) - T(k) \tag{B14}$$

A tedious but straightforward matrix manipulation of $\Omega(k)$ yields

$$\begin{aligned}\Omega_{1,1}(k) = {} & F_c\Sigma F_c' - \Sigma \\ & + \phi(k)'P_{2,2}(k)\phi(k)[\theta_2 h\Sigma F_c' + F_c\Sigma h\theta_2]m_3 \\ & + [\phi(k)'P_{2,2}(k)\phi(k)]^2\theta_2 h'\Sigma h\theta_2 m_3(k)^2 \\ & + \phi(k)'P_{2,2}\phi(k)\theta_2\theta_2'm_3(k)\end{aligned} \tag{B15}$$

$$\begin{aligned}\Omega_{1,2}(k) = {} & F_c\Sigma h\phi(k)'m_3(k) - \theta_2\phi(k)'m_3(k) \\ & - \phi(k)'P_{2,2}\phi(k)\theta_2 h'\Sigma h\phi(k)'m_3(k)\end{aligned} \tag{B16}$$

$$\Omega_{2,2}(k) = \phi(k)h'\Sigma\phi(k)'m_3(k)^2 - \phi(k)\phi(k)'m_3(k) \tag{B17}$$

where

$$\Omega(k) = \begin{bmatrix} \Omega_{1,1}(k) & \Omega_{1,2}(k) \\ \Omega_{1,2}(k)' & \Omega_{2,2}(k) \end{bmatrix} \tag{B18}$$

Since $H(q^{-1}) = 1/2 - \theta_2'(qI - F_c')^{-1}h$ is s.p.r., there exist a real vector ξ, a real scalar τ and a real positive definite matrix Σ such that (Hitz, et al., 1969):

$$F_c \Sigma F_c' - \Sigma = -\xi\xi' \tag{B19}$$

$$F_c \Sigma h = -\theta_2 + \tau\xi \tag{B20}$$

$$h'\Sigma h = 1 - \tau^2 \tag{B21}$$

Using these relationships on equation (B18), we have:

$$\Omega_{11}(k) = -\zeta(k)\zeta(k)' - \phi(k)'P_{22}(k)\phi(k)\theta_2\theta_2'm_3(k)^2 \tag{B22}$$

$$\Omega_{12}(k) = -\tau\xi\phi(k)'m_3(k) - \phi(k)'P_{22}(k)\phi(k)\theta_2\phi(k)'(1 - \tau^2)m_3(k)^2 \tag{B23}$$

$$\Omega_{22}(k) = -\phi(k)\phi(k)'[\tau^2 + \phi(k)'P_{22}(k)\phi(k)]m_3(k) \tag{B24}$$

where

$$\zeta(k) = \xi - \phi(k)'P_{22}(k)\phi(k)\theta_2\tau m_3(k)^2 \tag{B25}$$

Since

$$\Omega_{11}(k) - \Omega_{12}(k)\Omega_{22}(k)^{+}\Omega_{12}(k)'$$

$$= -\frac{\phi(k)'P_{22}(k)\phi(k)}{\tau^2 + \phi(k)'P_{22}(k)\phi(k))}\{[\zeta + \tau\theta_2 m_3][\zeta + \tau\theta_2 m_3]' \tag{B26}$$

is semi-negative definite

and $$\Omega_{22}\Omega_{22}^{+}\Omega_{12}' = \Omega_{12}' \tag{B27}$$

by Lemma 2 in APPENDIX C, the matrix $\Omega(k)$ is semi-negative definite where "$^+$" denotes pseudoinverse. It immediately follows that the matrix $\Xi(k)$ is stable, so is the system (4.39).

APPENDIX 4C

Lemma 2: A given symmetric matrix Ω

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12}' & \Omega_{22} \end{bmatrix} \tag{C1}$$

is semi-negative definite

if $\Omega_{22} \leq 0$

$\Omega_{11} - \Omega_{12}\Omega_{22}^{*}\Omega_{12}' \leq 0$

and $\Omega_{22}\Omega_{22}^{*}\Omega_{12}' = \Omega_{12}'$

where "*" denotes pseudoinverse.

Proof:

Consider the quadratic form:

$$V = (x',y')\Omega(x',y')'$$
$$= x'\Omega_{11}x + 2x'\Omega_{12}y + y'\Omega_{22}y \tag{C5}$$

where x and y are any non-zero vectors of appropriate dimensions. V can be expressed as:

$$V = x'\Omega_{11}x + 2x'\Omega_{12}(\Omega_{22}^{*}\Omega_{22})y + y'\Omega_{22}y$$
$$+ x'\Omega_{12}\Omega_{22}^{*}\Omega_{22}\Omega_{22}^{*}\Omega_{12}'x$$
$$- x'\Omega_{12}\Omega_{22}^{*}\Omega_{22}\Omega_{22}^{*}\Omega_{12}'x$$

$$= x'(\Omega_{11} - \Omega_{12}\Omega_{22}^{*}\Omega_{12}')x$$
$$+ (y + \Omega_{22}^{*}\Omega_{12}'x)'\Omega_{22}(y + \Omega_{22}^{*}\Omega_{12}'x) \tag{C6}$$

V is non-positive. Hence, Ω is semi-negative definite.

APPENDIX 4D

In this appendix, we introduce an important type of stochastic process called supermartingale and discuss those of its properties pertinent to the proof presented in Section 4.6.

Definition (Kushner, 1971): Let $\{x(k)\}$ be a sequence of random variables measurable with respect to (w.r.t) a random variable $z(k)$ and $\{\mathcal{F}(k)\}$ a sequence of an ever-increasing σ-algebras generated by $\{z(1), \cdots, z(k)\}$. The sequence of random variable $x(k)$ is then called a supermartingale if

$$E\{x(k) | \mathcal{F}(k-1)\} \leq x(k-1)$$

with probability 1 (w.p.1) for each k.

One of most successful applications of the supermartingale theory is in the proof of the convergence of recursive algorithms for stochastic systems. The convergence proof using the supermartingale theory usually results in stronger convergence properties compared to that in the mean squares sense. The application of the supermatingale theory for the convergence proof largely hinges on the following lemmas.

Lemma 3 (Kushner 1971): If $\{x(k)\}$ is a nonnegative supermartingale sequence and $E\{x(k)\} < \infty$ for all k, then $E\{x(k) | \mathcal{F}(k-2)\} \leq x(k-2)$

Proof: since $\mathcal{F}(k-2) \subset \mathcal{F}(k-1)$

$$E\{x(k)|\mathcal{F}(k-2)\} = E\{E\{x(k)|\mathcal{F}(k-1)\}|\mathcal{F}(k-2)\}$$

$$\leq E\{x(k-1)|\mathcal{F}(k-1)\} \leq x(k-2)$$

The important implication of the above lemma is that $E\{x(k)\}$ converges to a finite random variable. The proof of this implication could be found in (Neuve, 1975).

Lemma 4 (Neuve, 1975): Let $\{a(k)\}$ and $\{x(k)\}$ be sequences of nonnegative random variables measurable w.r.t $\mathcal{F}(k)$ and

$$E\{x(k)|\mathcal{F}(k-1)\} \leq x(k-1) + a(k)$$

then $x(k) \to x < \infty$ w.p.1 exists

if $\sum a(k) < \infty$

Proof: Let $r(k) = x(k) - \sum_{i=1}^{k} a(i)$

then $E\{r(k)|\mathcal{F}(k-1)\} \leq x(k-1) - \sum_{i=1}^{k} a(i) - a(k)$

$$= r(k-1) - a(k) \leq r(k)$$

thus r(k) is a supermartingale and this implies, according to Lemma 3 and its consequence, that r(k) converges to r w.p.1

Since $\sum a(i) < \infty$, $x(k) = r(k) + \sum_{i=1}^{k} a(i)$ converges to a finite random variable $x < \infty$.

Lemma 5 (Solo, 1979): Let $\{\beta(k)\}$ and $\{s(k)\}$ be sequences of nonnegative random variables measurable w.r.t $\mathcal{F}(k)$ and let

$$E\{s(k)|\mathcal{F}(k-1)\} \leq s(k-1) + \alpha(k) - \beta(k)$$

and $\sum_{k=1}^{\infty} \alpha(k) < \infty$ w.p.1

then $s(k) \to s$ w.p.1 and $\sum_{k=1}^{\infty} \beta(k) < \infty$ w.p.1

where s is a finite random variable.

Proof: Define $t(k) = s(k) + \sum_{i=1}^{k} \beta(i)$

then $E\{t(k)|\mathcal{F}(k-1)\} \leq t(k-1) + \alpha(k)$

It follows from Lemma 4 that

$$s(k) \to s < \infty \text{ w.p.1} \quad \text{and} \quad \sum_{k=1}^{\infty} \beta(k) < \infty \text{ w.p.1.}$$

CHAPTER V

SIMULATIONS

5.1 Introduction

In the previous chapter, three recursive identification algorithms for the linear stochastic system were derived. The algorithms were obtained by modifying the standard EKF algorithm for identification in order to reduce the computational burden and/or to avoid biasedness of the estimates. Of the three algorithms, it was shown that the third algorithm (EKF-M3) is the most general and hence most useful. In this chapter, the effectiveness of this algorithm as compared with two other algorithms previously reported in the literature will also be discussed with the help of simulation studies.

Since the ultimate objective of developing these algorithms is to apply them to real-life systems, it would seem appropriate that the comparison of various algorithms should be carried out with respect to real-life systems. However, this has certain disadvantages. In applying an algorithm to a particular system, especially if very little is known about the system, special approaches or tricks have to be used. These tricks may vary from algorithm to algorithm and system to system. Consequently a proper comparison of the different algorithms may become difficult. On the other hand, computer simulations provide a great deal

of insight into the relative strengths and weaknesses of the different algorithms tested for the following reasons.

i) complete control can be exercised on the conditions associated with the system in order to achieve clear comparison

and ii) it is possible to eliminate or reduce dispensable elements which may otherwise obscure the comparison of results.

The comparison of algorithms will therefore be carried out by means of computer simulations.

Many different recursive algorithms have been proposed (Saridis, 1974; Isermann, et al., 1974; Soderstrom, et al., 1978; Dugard, et al., 1980). Listed below are the most well-known algorithms which have been published:

RLS - the recursive least-squares algorithm

RIV - the recursive instrumental variable algorithm

RGLS- the generalized least-squares algorithm

ELS - the extended least-squares algorithm

RML - the recursive approximate maximum likelihood algorithm.

Considering that both RLS and RIV algorithms are formulated to estimate only the process parameters, not the noise parameters which are of main interest here, these two algorithms will not be used in the simulation studies. Although the RGLS method is intended to estimate the noise

parameters as well as the process parameters, the model structure representing noise differs from that used in the EKF-M3 and the last two algorithms listed. Consequently, the RGLS method will be of little value in the comparison studies in terms of assessement for possible real applications; thus this algorithm is also excluded. The ELS and RML algorithms will be used in simulations along with the EKF-M3 algorithm developed in this thesis. It is noted that the features of the ELS and RML algorithms were dicussed in the previous chaper.

Two examples will be used for the simulation studies. In Section 5.2, results based on a well known example which first appeared in (Ljung, et al., 1975) are discussed. These studies will compare the convergence properties of the algorithms. In Section 5.3, general comparison of the algorithms is reported based on a fourth order example used extensively by Lee (1964) and Saridis (1974).

5.2 First Simulation Study

A second order system is chosen:

$$x(k+1) = \begin{bmatrix} a_1 & 1 \\ a_2 & 0 \end{bmatrix} x(k) + \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} v(k)$$

$$y(k) = [1 \quad 0]x(k) + v(k)$$

with the following parameter values

$$a_1 = -0.9$$

$$a_2 = -0.95$$

$$d_1 = 0.6$$

$$d_2 = -0.2$$

thus, $c_1 = a_1 - d_1 = -1.5$

$$c_2 = a_2 - d_2 = -0.75$$

and $v(k) \sim N(0,1)$

This example is the same as that used by Ljung, et al. (1975) as a counterexample to demonstrate a biased convergence of the ELS algorithm. Since the EKF-M3 algorithm as well as the RML algorithms yield unbiased estimates, the ELS algorithm was modified also to yield an unbiased estimate so that a true comparison of rates of convergence of the three algorithms could be made. The modified ELS algorithm is designated as GELS. The details of the modifications are given in Appendix 5A.

Several values of initial state were used in the simulation studies. While the GELS and the EKF-M3 algorithm converged every time with a proper choice of vector f, the convergence of the RML algorithm was possible only by the use of a projection facility such that estimate c(k) stays in a region where

$$\frac{1}{1 - \hat{c}_1(k)q^{-1} - \hat{c}_2(k)q^{-2}}$$

is always stable.

Finally for the EKF-M3 and GELS algorithms, a vector f was selected as:

$$f = [-1.0 \ -0.35]'$$

Figures 5-1 and 2 show the results as well as the measurement y(k).

5.3 Second Simulation Study

For the second comparison study, the following fourth order system is chosen:

$$x(k+1) = Ax(k) + \omega(k)$$

$$y(k) = h'x(k) + \nu(k)$$

where

$$A = \left[\begin{array}{c|c} a & \begin{matrix} I \\ \hline 0 \end{matrix} \end{array}\right]$$

$$h = [1 \ 0 \ 0 \ 0]'$$

with the Gaussian noises

$$\omega(k) \sim N(0.0, Q)$$

$$\nu(k) \sim N(0.0, r)$$

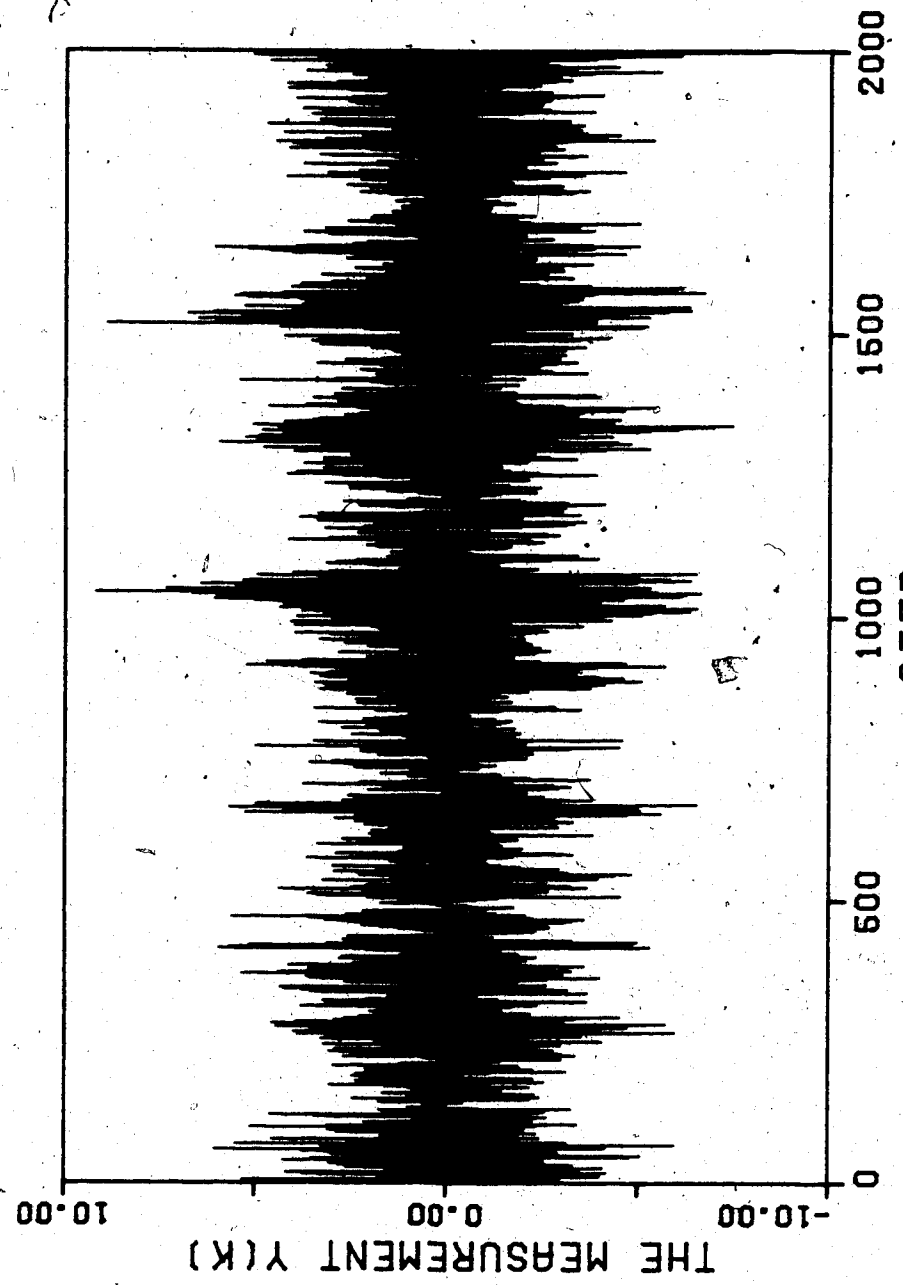
For purpose of simulation, the parameters to be identified are set as:

$$a = [\ 1.00 \ -0.18 \ \ 0.78 \ -0.66]'$$

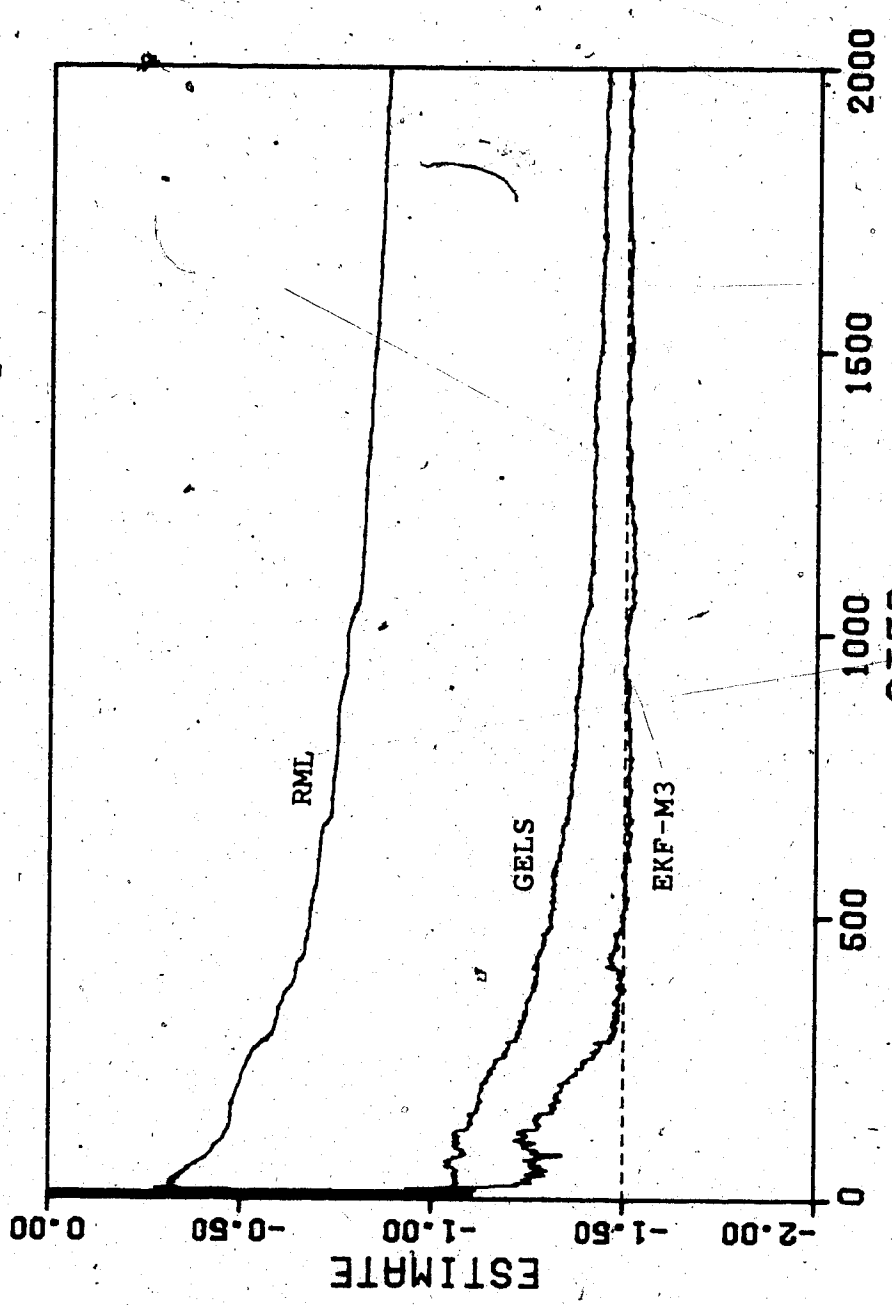
$$Q = \begin{bmatrix} 1.00 & 0.00 & 1.18 & -1.60 \\ 0.00 & 0.00 & 0.00 & 0.00 \\ 1.18 & 0.00 & 1.39 & -1.89 \\ -1.60 & 0.00 & -1.89 & 2.56 \end{bmatrix}$$

$$r = 0.25$$

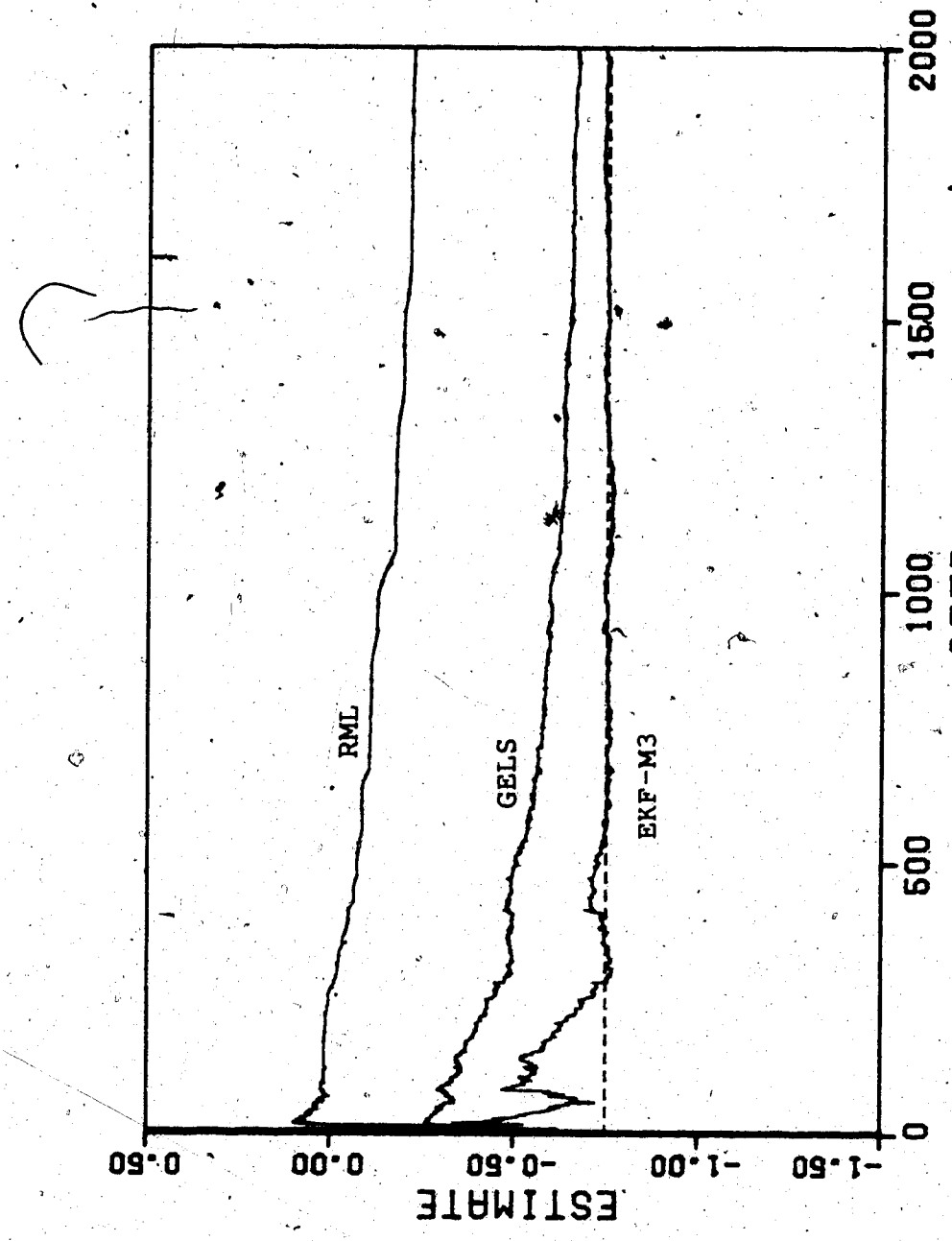
This system is equivalent in terms of input-output relation to:



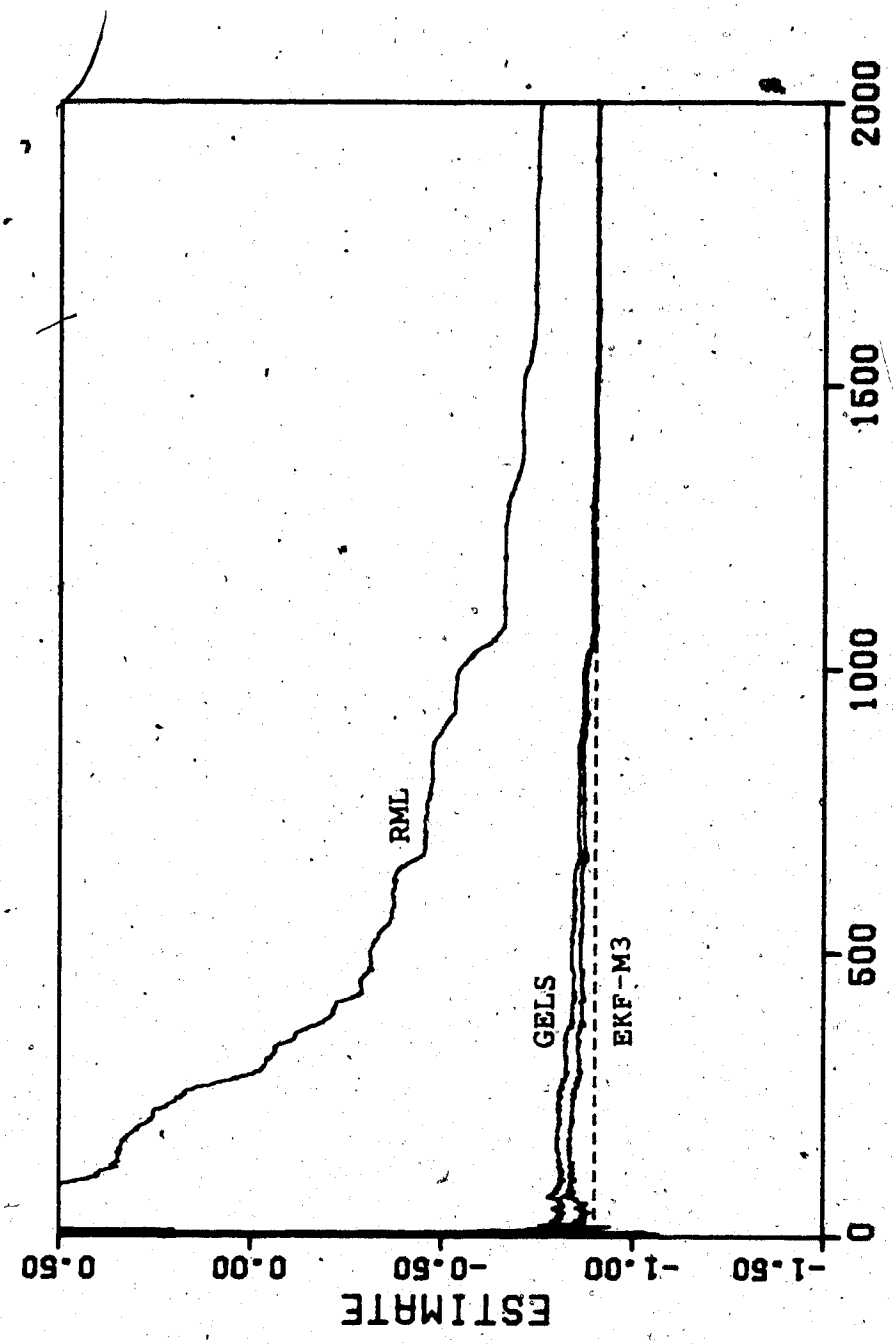
5.1 The output run of the measurement y(k)



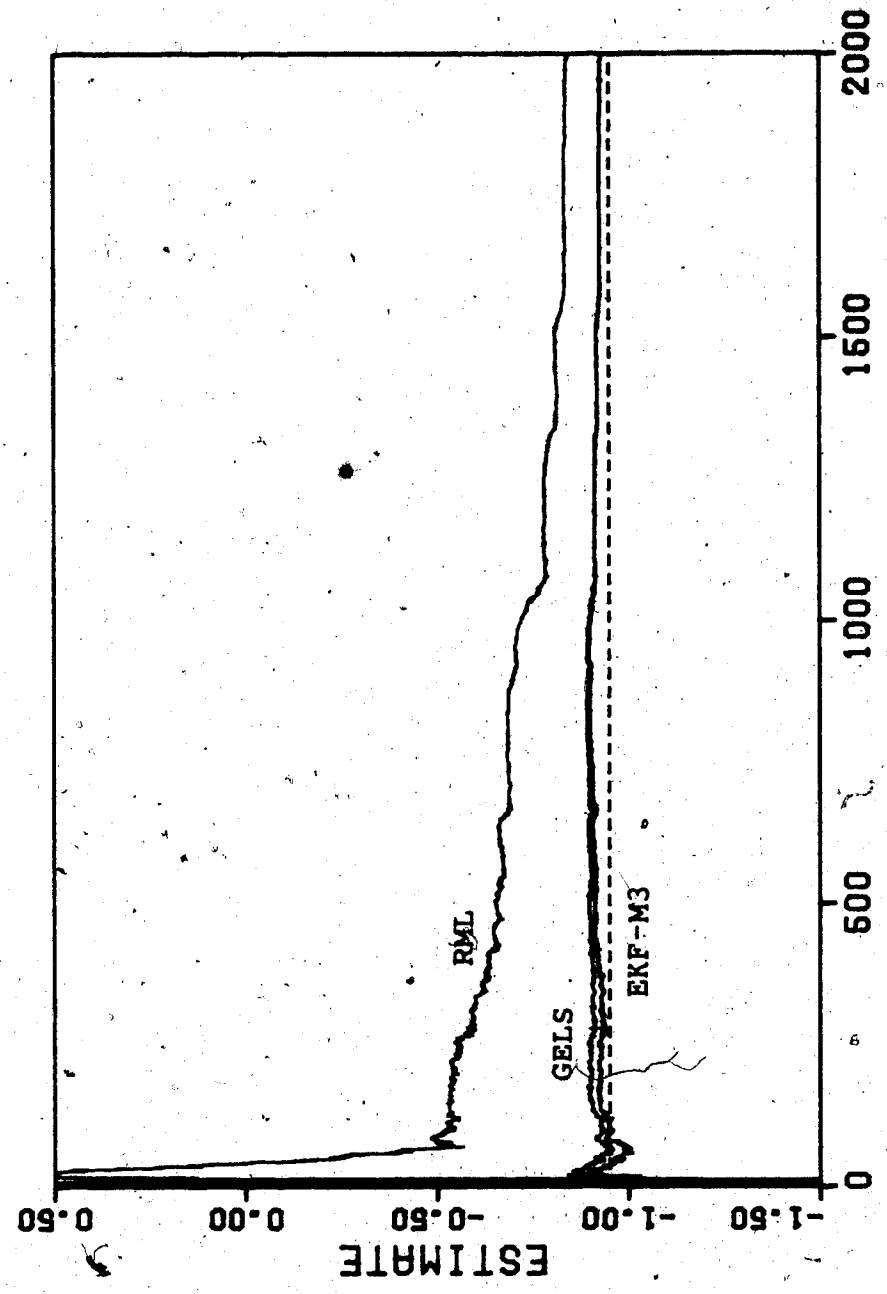
5.2a The estimates of c_1



5.2b The estimates of a_2



5.2c The estimates of a,



5.2d The estimates of a_2

$$x(k+1) = Ax(k) + dv(k)$$

$$y(k) = h'x(k) + v(k)$$

where $d = [0.570 \quad -0.027 \quad 0.306 \quad -0.619]'$

thus, $c = [0.430 \quad -0.153 \quad 0.474 \quad -0.041]'$

and $v(k) \sim N(0.0, 3.988)$

Two features of this system are worth noting. This system is very sensitive to noise (Lee, 1964) and consequently, it is difficult to identify the system parameters without bias (Saridis, 1974). Also

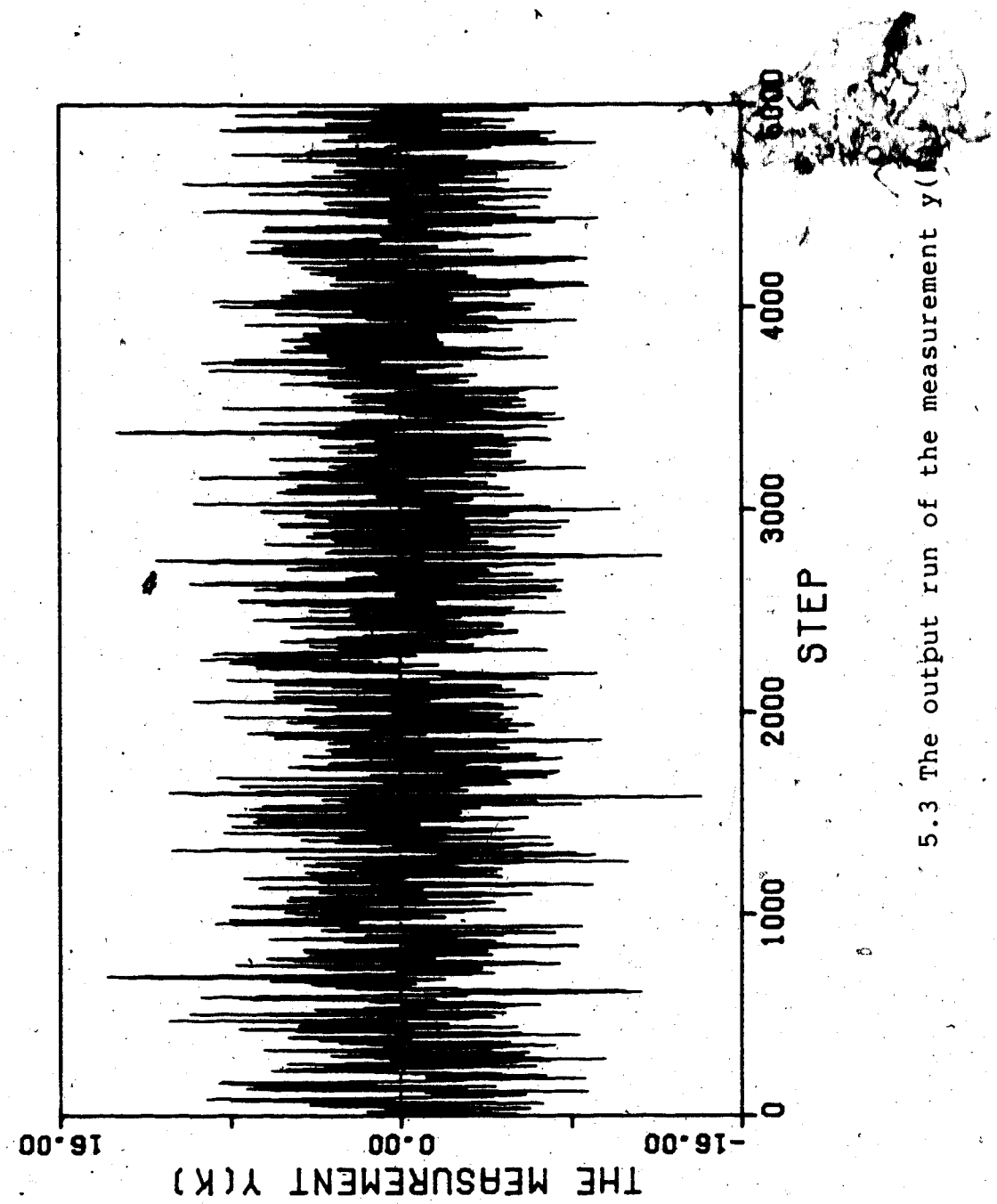
$$\frac{1}{1 - 0.430q^{-1} + 0.153q^{-2} - 0.474q^{-3} + 0.041q^{-4}} - \frac{1}{2}$$

is not strictly positive real.

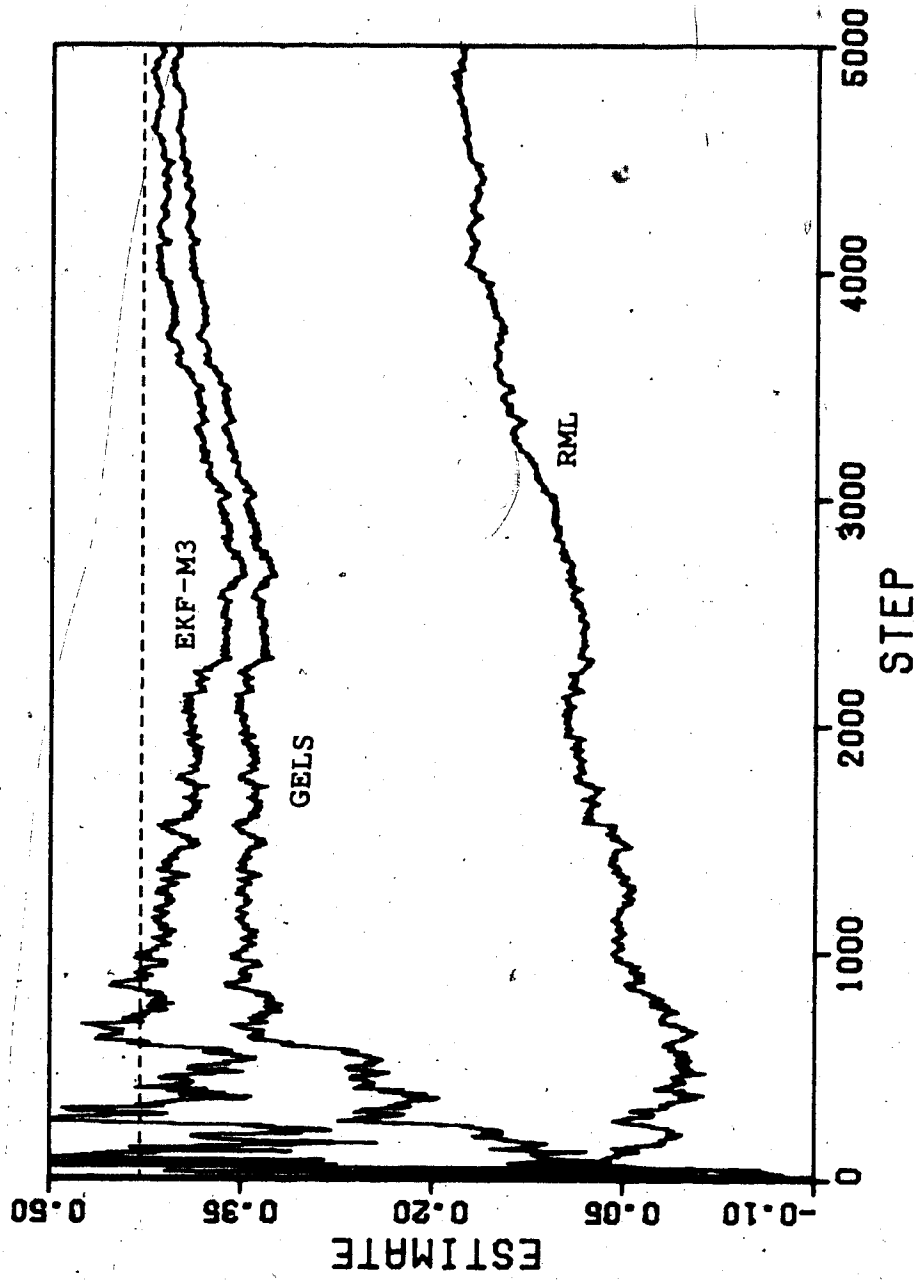
The same three algorithms used in the first simulation study are also used here. For the EKF-M3 and GELS algorithms, the following values were assumed as the à priori informations on the vector parameter c. The same values are also used as the vector f:

$$\hat{c}(0) = f = [0.20 \quad 0.05 \quad 0.30 \quad 0.30]'$$

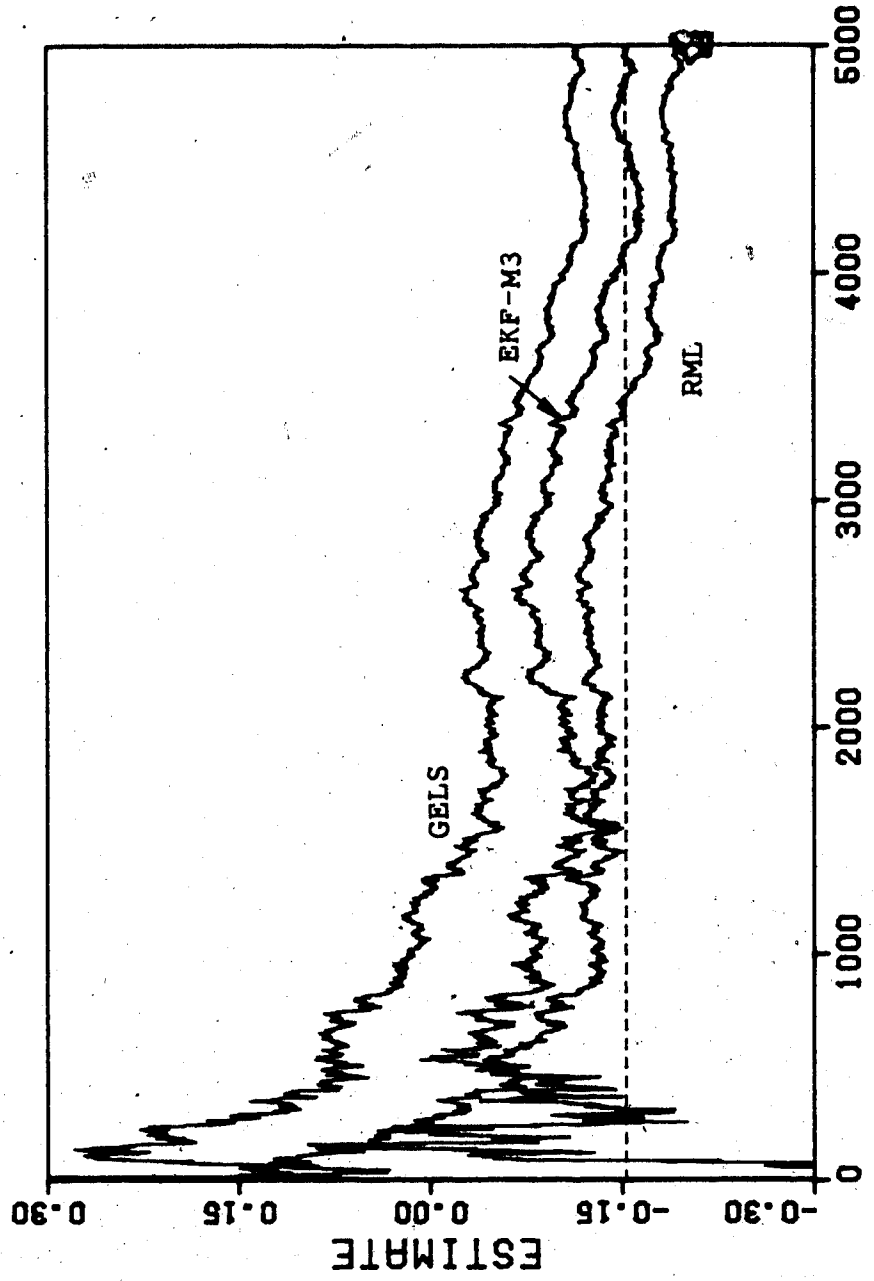
The RML algorithm was also initiated with the same initial values for fair comparisons. The RML algorithm is also complemented with the projection facility as the previous simulation study. The results of the simulation studies are shown in Figure 5.3 and 5.4.



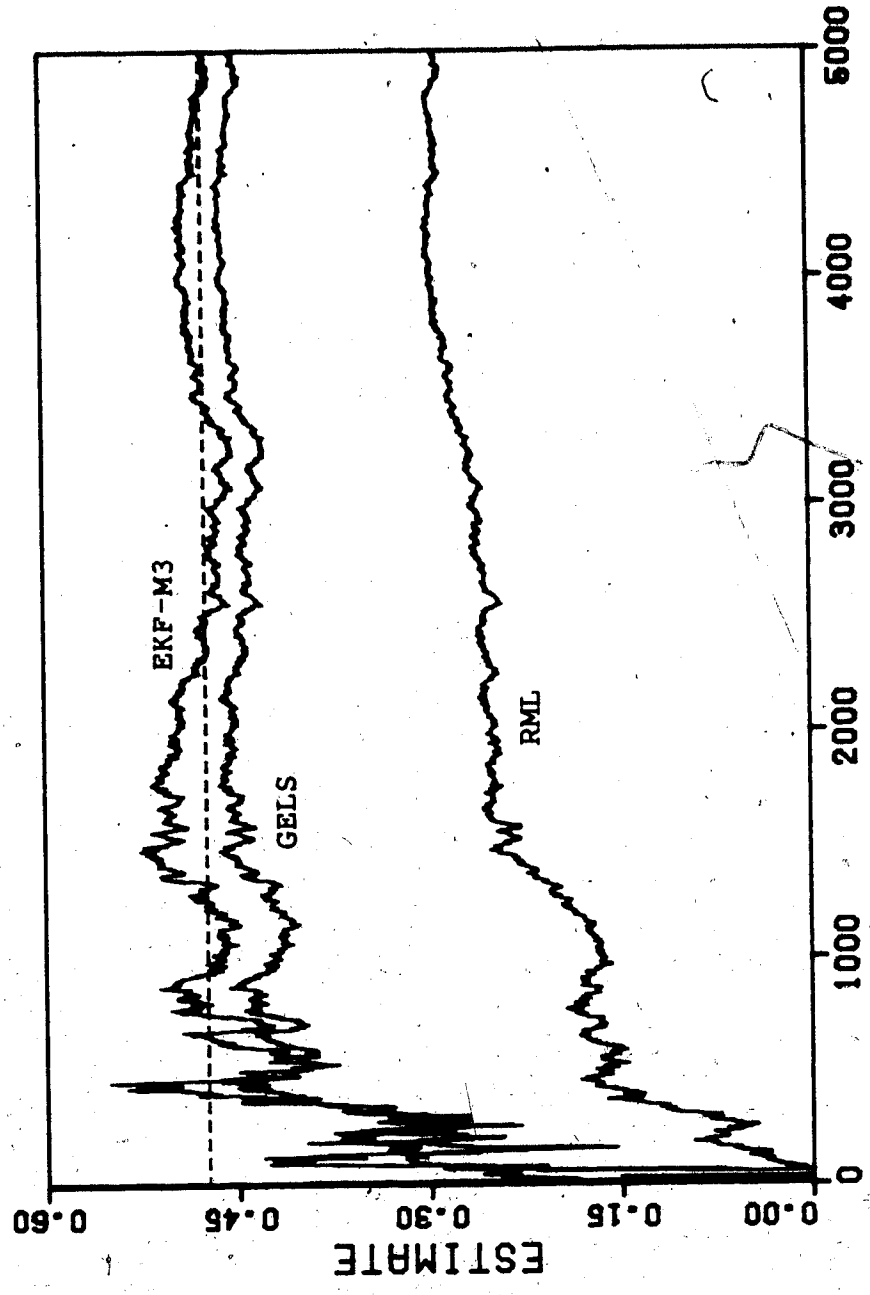
5.3 The output run of the measurement y(



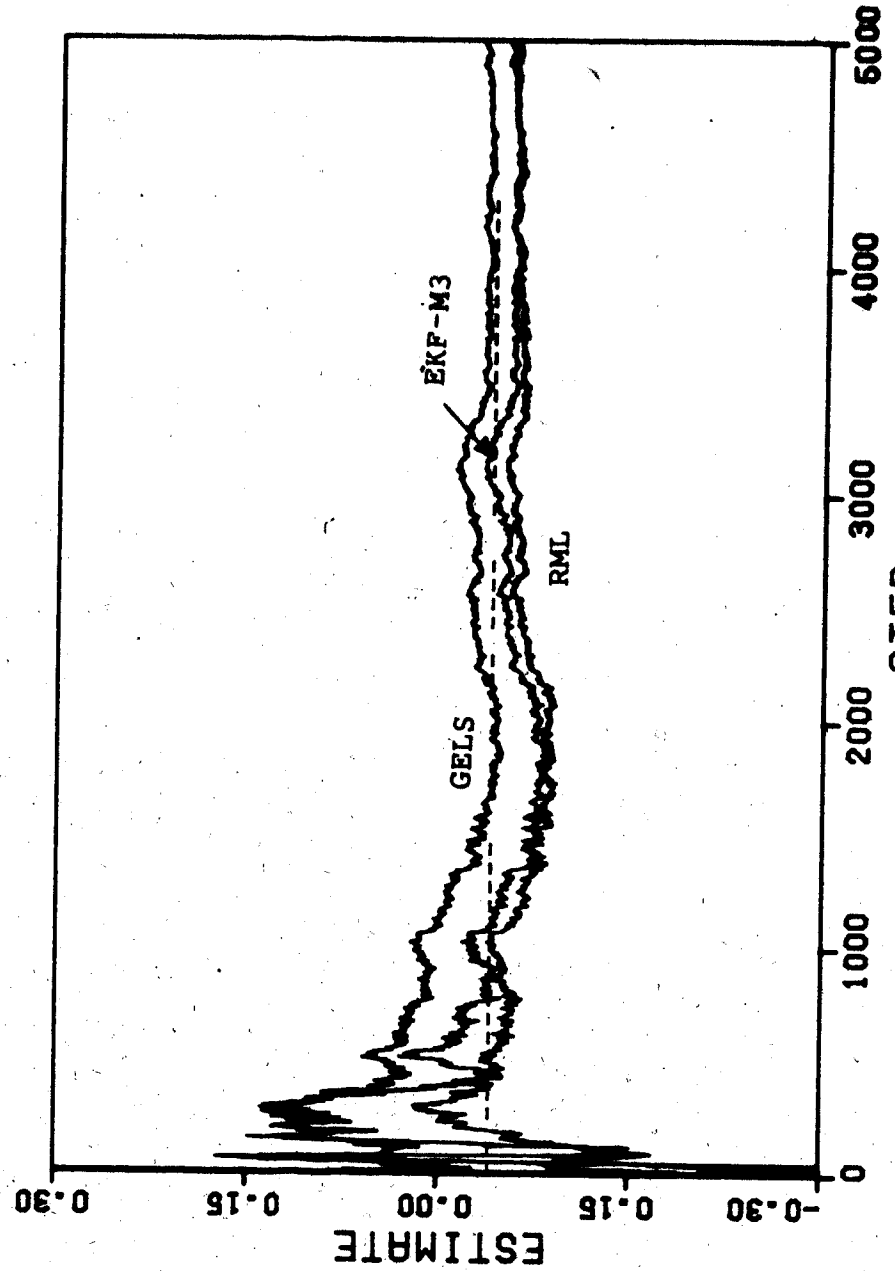
5.4a The estimates of c_1



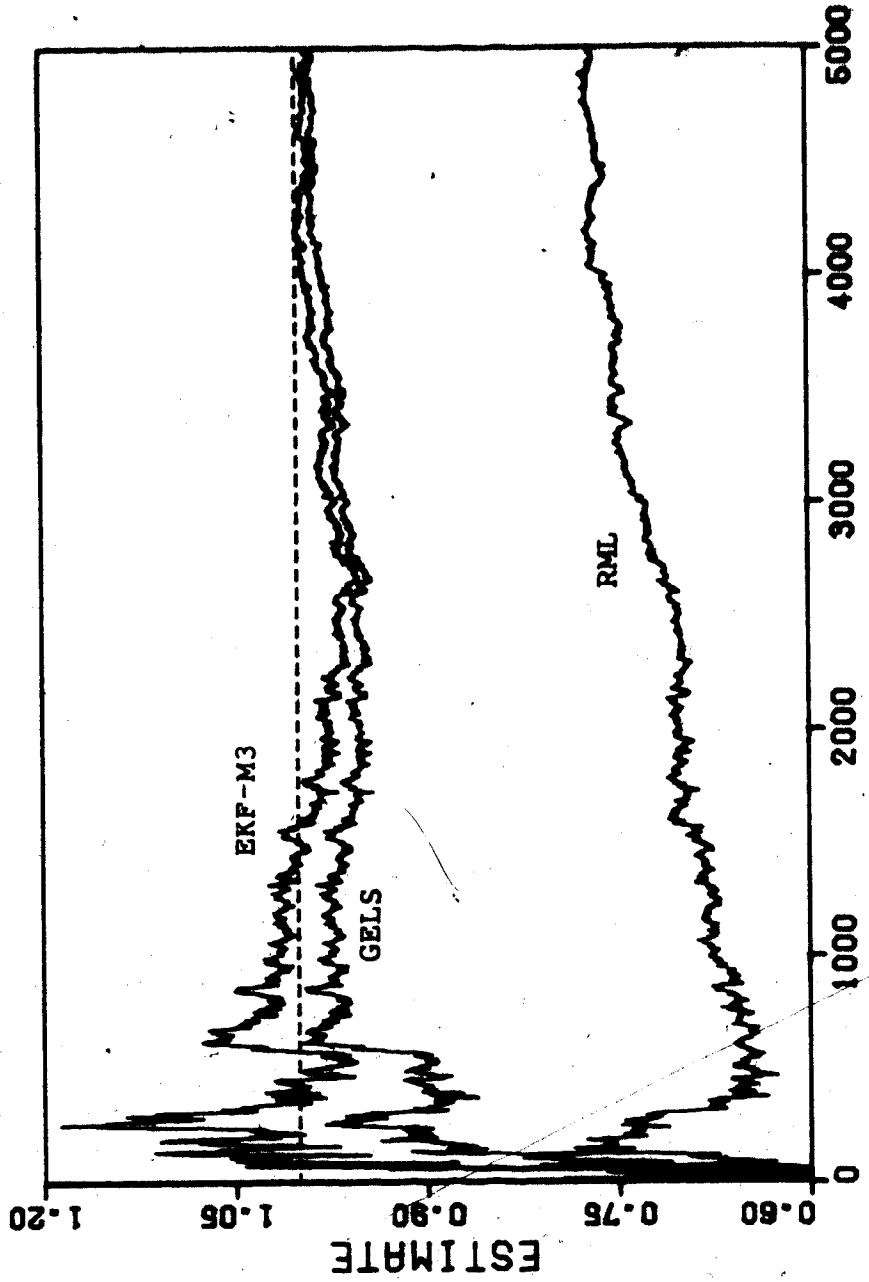
5.4b The estimates of c_2



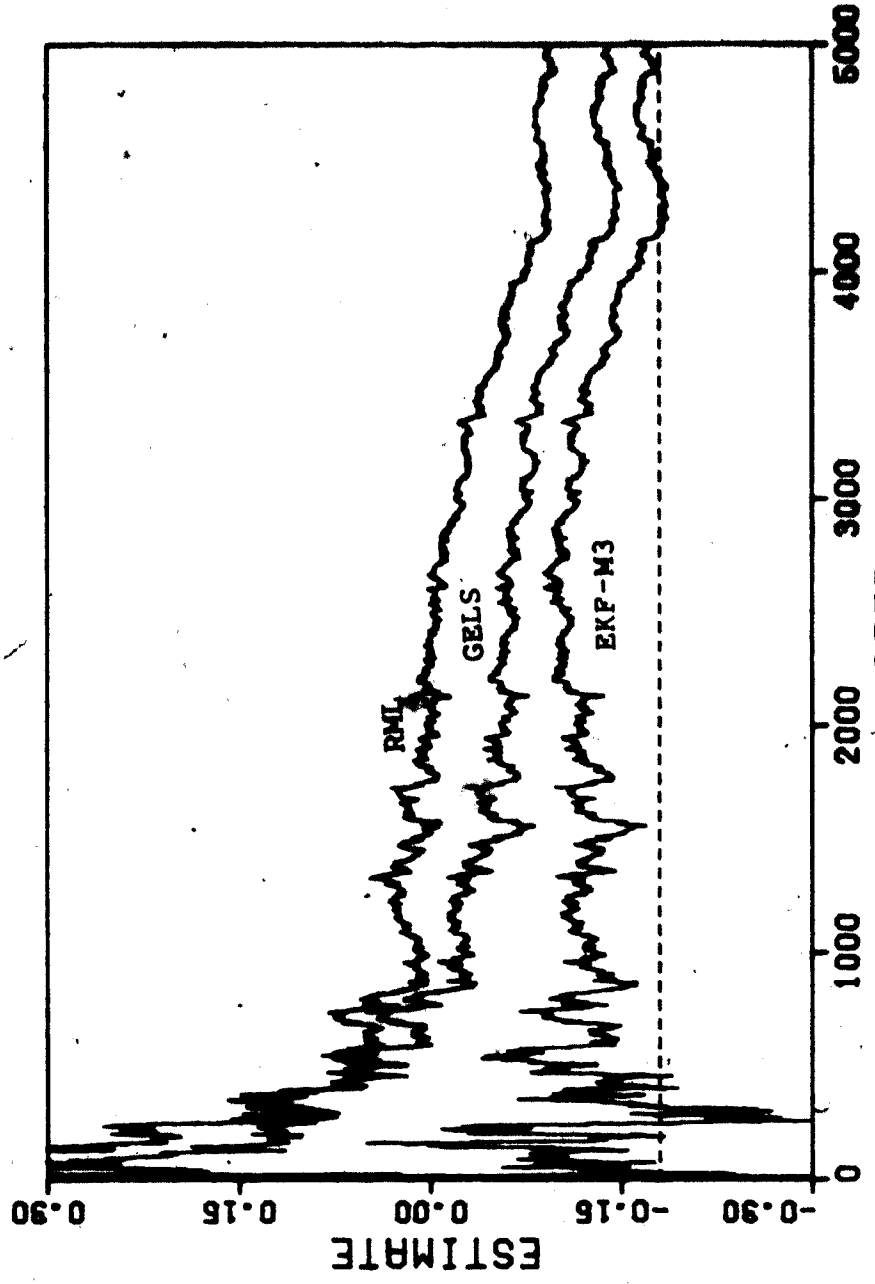
5.4c The estimates of c_3



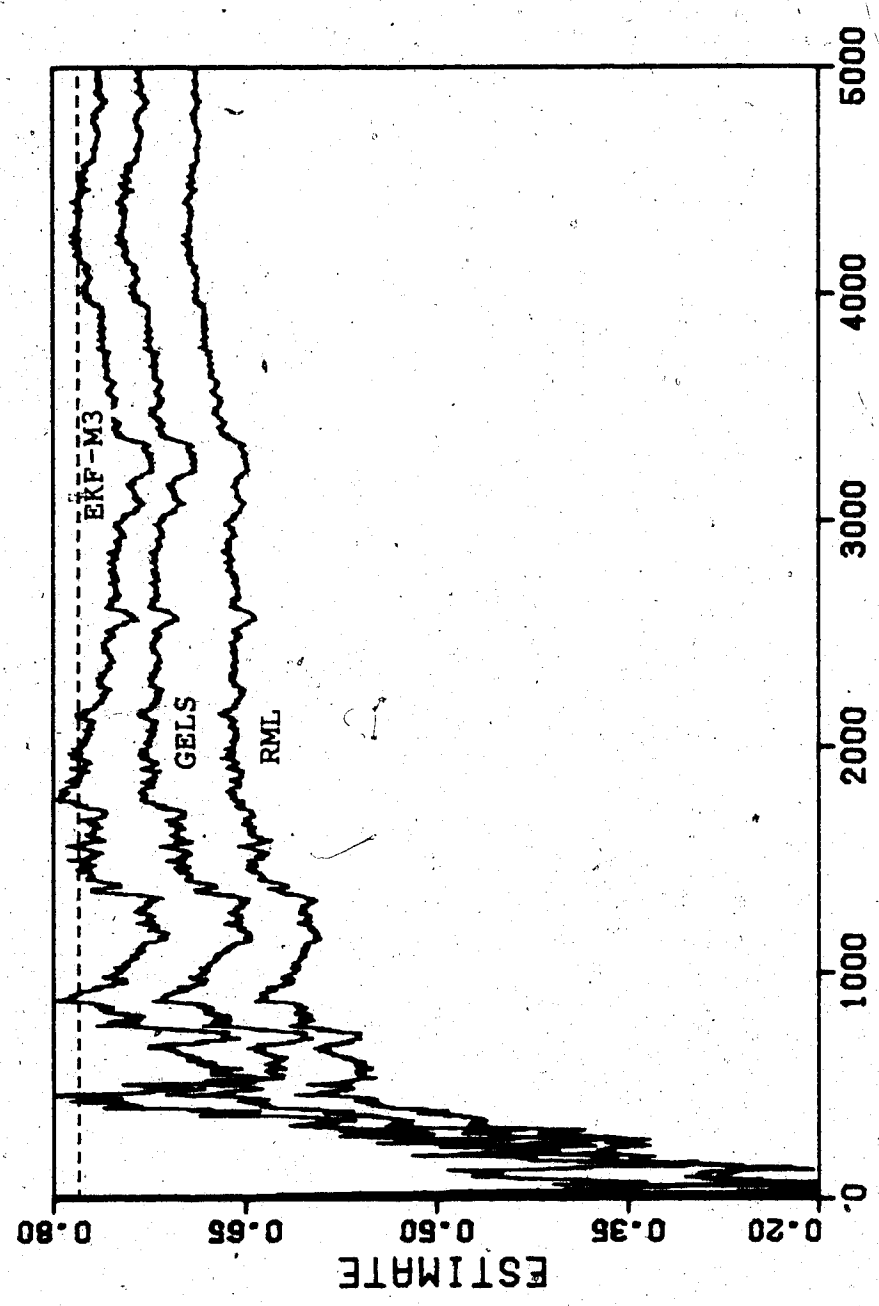
5.4d The estimates of c.



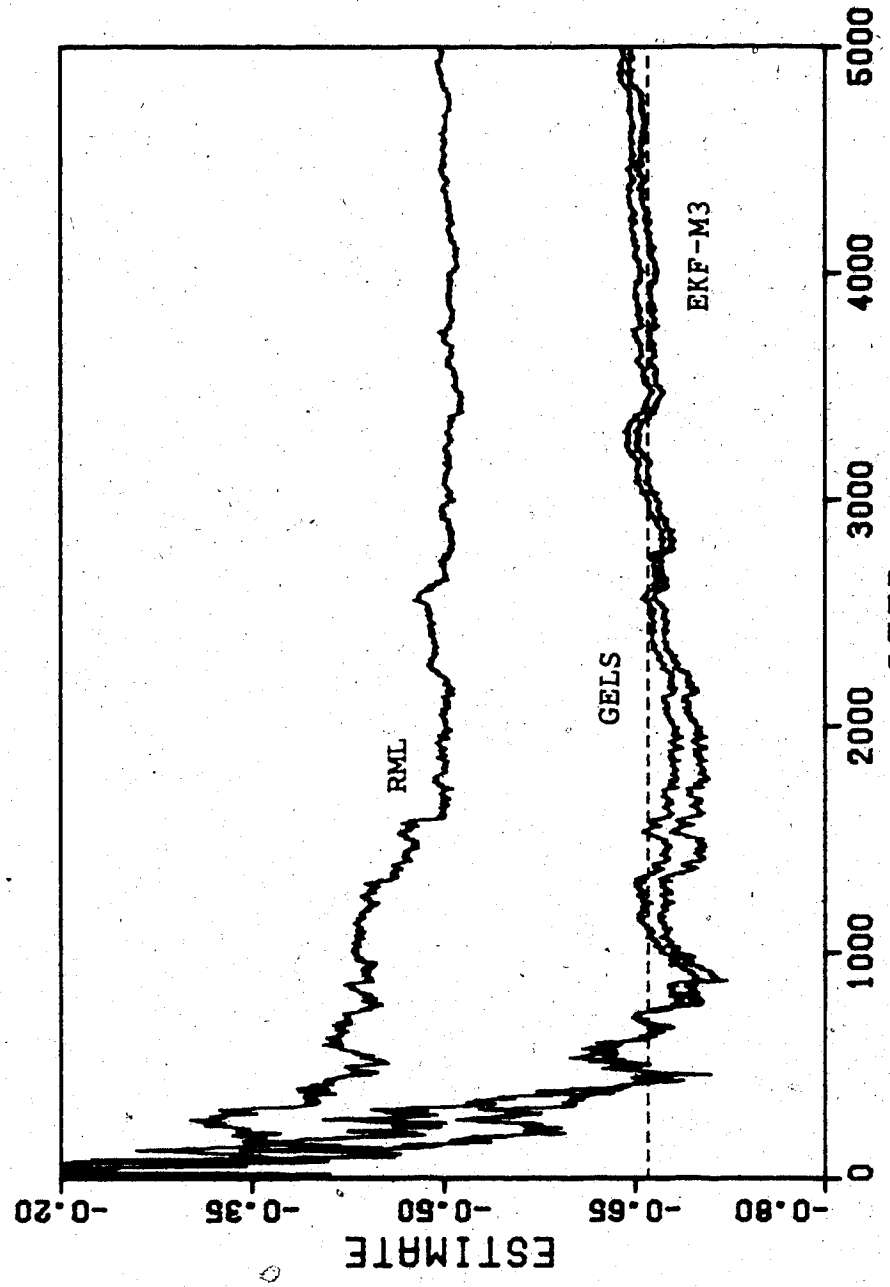
5.4e The estimates of a,



5.4f The estimates of a_1



5.4g The estimates of a,

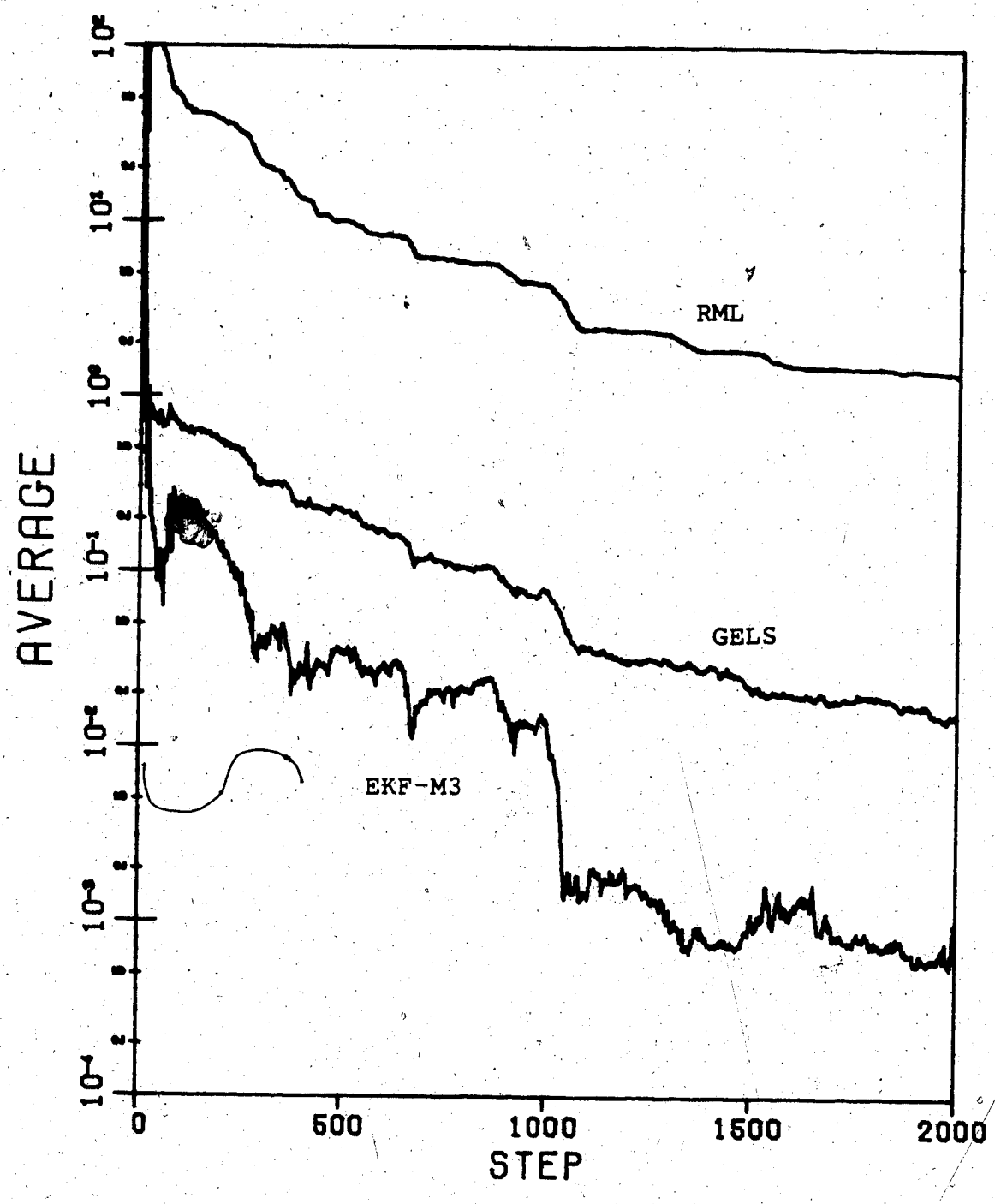


5.4h The estimates of a,

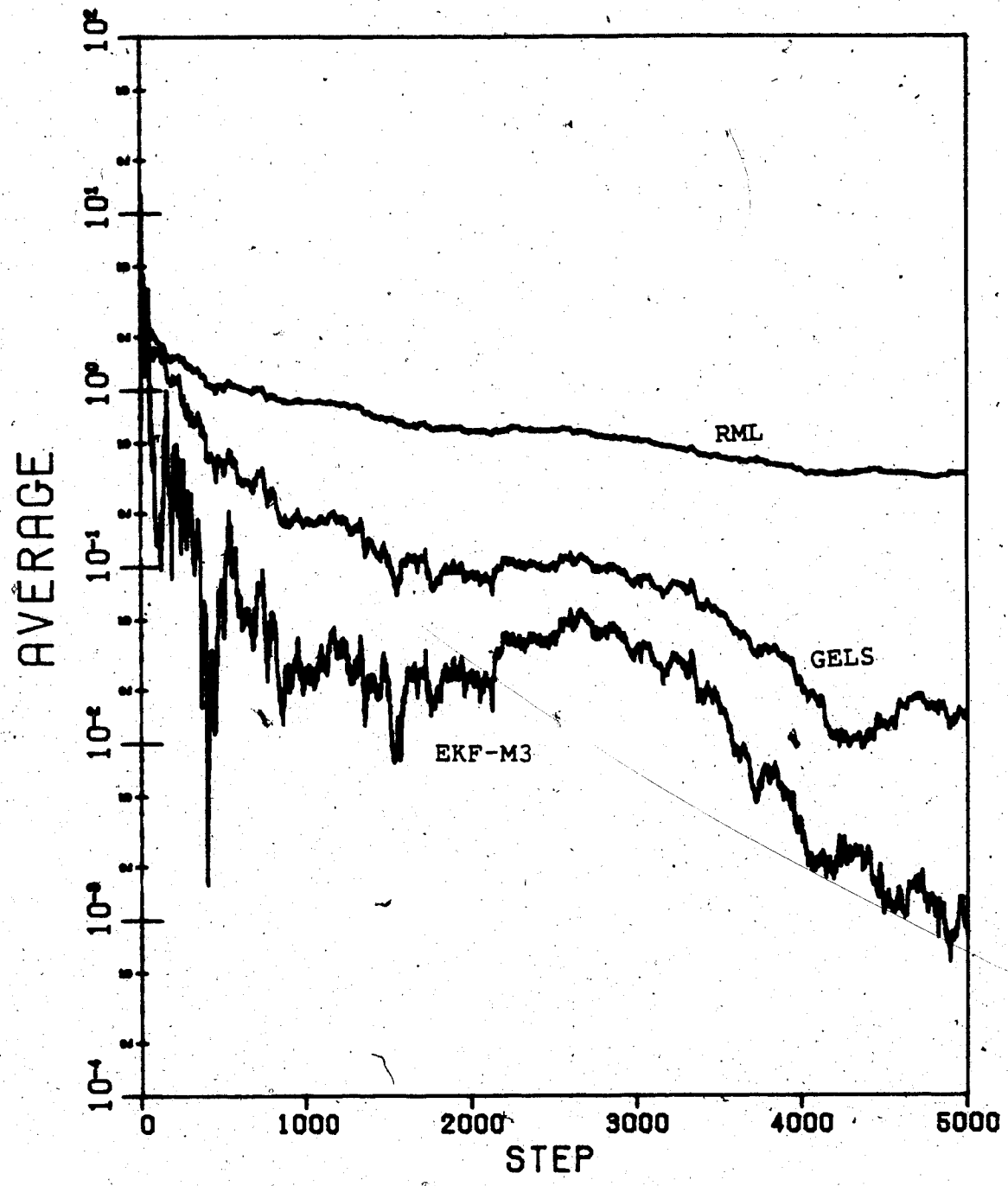
5.4 Remarks

The results of the comparison are summarized in Figures 5-5 and 5-6 for both studies. These figures illustrate the convergence properties of the algorithms. The squared estimation errors were normalized by dividing then by the initial squared errors to give the same weight to each estimate.

Both studies show that the performance of the EKF-M3 was vastly superior to those of the other two algorithms for almost all the the parameters. The performance of the RML algorithm on the other hand varied considerably depending on the parameters. For example, in the second study its convergence rates for estimates $c_1(k)$ and $c_3(k)$ were unacceptably slow while those for $c_2(k)$ and $c_4(k)$ were surprisingly fast. Though the overall convergence rate of the RML algorithm was not much slower than other two algorithms as shown in the figures, this algorithm experienced the worst transient reponse among the three. As a result, even after 5000 iterations the convergence of this algorithm was not conclusive. As expected, the GELS algorithm showed slower convergence than the EKF-M3 algorithm. However the convergence patterns of the both algorithms were quite similar as shown in Figure 5-4 because the gradients of the error correction terms of the both algorithms are the same.



5.5 The second order system identification results



5.6 The fourth order system identification results

APPENDIX 5A

Derivation of the Generalized Extended Least Squares Algorithm (GELS)

As was pointed out in Section 4.5, it should be noted that the difference between the ELS algorithm and EKF-M2 algorithm is that in ELS à priori error $\epsilon(k)$ is used for the generation of $\hat{\theta}_1(k)$ while the EKF-M2 algorithm uses à postériori error $\epsilon_1(k)$. It should also mentioned that the EKF-M3 algorithm was obtained by generalizing EKF-M2 using a canonical form which included a general vector f to avoide the possibility of biased convergence. These two ideas will now be used to obtain a generalized Extended Least Squares algorithm. The GELS algorithm is then of the form:

$$z_2(k+1) = F'z_2(k) + hy(k) \tag{A1.1}$$

$$\hat{\theta}_1(k+1) = F'\hat{\theta}_1(k) + h[y(k) - \hat{y}(k)] \tag{A1.2}$$

$$\hat{p}(k+1) = \hat{p}(k) + P_{22}(k)\phi(k)m_3(k)[y(k) - \hat{y}(k)] \tag{A1.3}$$

$$y(k) = \phi(k)'\hat{p}(k) \tag{A1.4}$$

$$\phi(k) = [\hat{\theta}_1(k)', z_2(k)']' \tag{A1.5}$$

$$P_{22}(k+1) = P_{22}(k) - P_{22}(k)\phi(k)m_3(k)\phi(k)'P_{22}(k) \tag{A1.6}$$

$$m_3(k) = [1 + \phi(k)'P_{22}(k)\phi(k)]^{-1} \tag{A1.7}$$

$$P_{22}(0) = P_{22}(0)' > 0$$

Covergence of GELS

The errors $\tilde{p}(k) = p - \hat{p}(k)$ and $\tilde{\theta}_1(k) = \theta_1(k) - \hat{\theta}_1(k)$ satisfy the following error equations:

$$\tilde{\theta}_1(k+1) = F_c'\tilde{\theta}_1(k) - h\gamma(k) \qquad \text{(A2.1)}$$

$$e(k) = \tilde{\theta}_2'\theta_1(k) + \gamma(k) + v(k) \qquad \text{(A2.2)}$$

$$\tilde{p}(k+1) = \tilde{p}(k) - P_{22}(k)\phi(k)m_3(k)e(k) \qquad \text{(A2.3)}$$

$$\gamma(k) = \phi(k)'\tilde{p}(k) \qquad \text{(A2.4)}$$

$$P_{22}(k+1) = P_{22}(k) - P_{22}(k)\phi(k)m_3(k)\phi(k)'P_{22}(k) \qquad \text{(A2.5)}$$

where $F_c = F - \theta_2 h'$

As was the case for the EKF-M3 algorithm, the GELS algorithm is also assured of unbiased convergence if the errors $\tilde{p}(k)$ and $\tilde{\theta}_1(k)$ converge to the origin in the error vector space. Furthermore, if the errors $\tilde{p}(k)$ and $\tilde{\theta}_1(k)$ converge to the origin without the explicit noise v(k) in the equation (A2.2), one can easily conclude by similarity of the structure of the error equations that the GELS algorithm will yield unbiased estimates of the parameters.

Consequently, the overall unbiased convergence of the GELS algorithm can be established by examining the stability of the error feedback system without the noise term v(k) in equation (A2.2). This stability problem can be stated as follows:

Lemma: Given the following time-varying system:

$$\tilde{\theta}_1(k+1) = F_c'\tilde{\theta}_1(k) - h\gamma(k) \qquad \text{(A3.1)}$$

$$e(k) = \theta_2'\tilde{\theta}_1(k) + \gamma(k) \qquad \text{(A3.2)}$$

$$\tilde{p}(k+1) = \tilde{p}(k) - P_{22}(k)\phi(k)m_3(k)e(k) \qquad \text{(A3.3)}$$

$$\gamma(k) = \phi(k)'\tilde{p}(k) \qquad \text{(A3.4)}$$

$$P_{22}(k+1) = P_{22}(k) - P_{22}(k)\phi(k)m(k)\phi(k)'P_{22}(k) \qquad (A3.5)$$
$$P_{22}(0) = P_{22}(0)' > 0$$

$\tilde{\theta}_1(k)$ and $\tilde{p}(k)$ converge to zero as $k \to \infty$ if $\phi(k)'P_{22}(k)\phi(k) \leq \delta$ for all k where δ is a positive scalar and

$$H(q^{-1}) = 1/2 - \theta_2'(qI - F_c')^{-1}h$$

is strictly positive real.

Proof:

Consider the following quadratic function as a candidate for the Lyapunov function:

$$V(k) = V_1(k) + V_2(k) \qquad (A4)$$
$$V_1(k) = \tilde{\theta}_1(k)'\Sigma\tilde{\theta}_1(k) \qquad (A5)$$
$$V_2(k) = p(k)'P_{22}(k)^{-1}p(k) \qquad (A6)$$

It follows from equations (A3.1-2):

$$\nabla V_1(k) = V_1(k+1) - V_1(k)$$
$$= \tilde{\theta}_1(k)'[F_c\Sigma F_c' - \Sigma]\tilde{\theta}_1(k) - 2\tilde{\theta}_1(k)'F_c\Sigma h\gamma(k)$$
$$+ h'\Sigma h\gamma(k)^2 \qquad (A7)$$

$$\nabla V_2(k) = V_2(k+1) - V_2(k)$$
$$= \gamma(k)^2 - 2\gamma(k)e(k) + \phi(k)'P_{22}(k)\phi(k)m_3(k)e(k)^2 \qquad (A8)$$

In equation (A8), the following matrix inversion lemma has been used:

$$P_{22}(k+1)^{-1} = P_{22}(k)^{-1} + \phi(k)\phi(k)' \qquad (A9)$$

Since $H(q^{-1}) = 1/2 - \theta_2'(qI - F_c')^{-1}h$ is s.p.r., there exist a real vector ξ, a real scalar τ and a real positive definite matrix Σ such that (Hitz, et al., 1969)

$$F_c\Sigma F_c' - \Sigma = -\xi\xi' \qquad (A10)$$

$$F_c\Sigma h = -\theta_2 + \xi\tau \qquad (A11)$$

$$h'\Sigma h = 1 - \tau^2 \qquad (A12)$$

With these relationships, equation (A7) becomes:

$$\nabla V_1(k) = -[\theta_1(k)'\xi + \gamma(k)\tau]^2 + 2\gamma(k)e_1(k) - \gamma(k)^2 \qquad (A13)$$

thus,

$$\nabla V(k) = \nabla V_1(k) + \nabla V_2(k)$$

$$= -\{[\theta_1(k)'\xi + \gamma(k)\tau]^2 - \phi(k)'P_{22}(k)\phi(k)m_3(k)e(k)^2\} \qquad (A14)$$

Since $\phi(k)$ is bound and $P_{22}(k)$ is nonincreasing, there exists a positive scalar δ such that

$$\phi(k)'P_{22}(k)\phi(k) \leq \delta \qquad (A15)$$

in which case $\nabla V(k) \leq 0$,

and the feedback system (A3) is stable.

Consequently, the GELS algorithm is assured of producing unbiased estimates of the parameters. The slower convergence of GELS compared to EKF-M3 can be attributed to this additional restriction (A15).

CHAPTER VI

CONCLUSIONS

6.1 Summary

This thesis is entirely devoted to the development of various recursive identification methods for deterministic as well as stochastic discrete linear time-invariant systems. Though at the beginning of each chapter, there was a brief discussion on the specific subject which was dealt in that chapter, it is the purpose of this section to summarize the salient features of each chapter and the results obtained in this thesis.

In Chapter II, a canonical form of the innovations representation for the multiinput singleoutput system is introduced. The preference for using the innovations representation over other representations, in particular for the identification problem is largely related to the uniqueness of parameter estimates for a given measurement. This uniqueness, which is of great importance for the identification problem, originates from the fact that the innovations representation is the sole minimum phase system for a given rational spectral density (Anderson, et al., 1979). This, in turn, implies that for a given measurement $y(k)$, a stationary process, there exists only one minimum phase system whose output could be regarded as the measurement and this minimum phase system is identical to

the innovations representation. Thus, by adoption of the innovations representation, a risk of ambiguities of identified parameters could be almost eliminated. Furthermore, the structure of the canonical form introduced is suited especially to the least squares method for the identification problem. This form is also flexible enough to facilitate the development of new identification algorithms.

Chapter III presents a fast identification algorithm for the deterministic linear system. This was done by geometrically interpreting the Kudva-Narendra (K-N) identification algorithm and improving the convergence speed by the introduction of orthonormal vectors. The special advantage of this scheme is that the overall computational requirements are considerably reduced in comparison to the K-N algorithm. While the underlying idea of this development is simple and straightforward, the geometric interpretation makes it possible to visualize how the K-N algorithm works and consequently to improve the convergence rate. Furthermore, this approach provided a basis for the direction to be taken for the development of identification algorithms when noise is present, i.e. the stochastic system identification which is the subject of Chapter IV.

A stumbling block in the use of the standard least squares method for the stochastic system identification problem is, in general, the inclination of the least squares

method to bias the estimates (Astrom, et al., 1971) if noise is correlated, as is generally the case. Several techniques have been proposed for the recursive identification problem to get around this drawback. The basic philosophy behind the proposed methods in common is to decompose the correlated noise into a sequence of weighted white noise or innovations, which is often called as Wold decomposition and then to use the standard recursive least squares method to identify the augumented parameters which includes the parameters for weighing the innovations. One of the obvious problems arising in this approach is that the innovations are not amenable to measurement. Consequently, different methods have been suggested to estimate the innovations.

Since the identification problem for the linear stochastic system can be considered as a nonlinear estimation problem, the Extended Kalman Filter algorithm has been quite often used as a starting step for developing new algorithms. In chapter IV, this approach is once again taken to derive several different identification algorithms in order mainly to alleviate the unnecessary computational burden. The modifications are made possible because of the following observations:

1) the EKF algorithm for the identification problem is assured of its unbiased convergence in the neighborhood of the true values of the parameters.
2) the use of the innovations representation combined with

the above observation justifies setting some elements of the approximate covariance matrix of the EKF algorithm to be zero, thus leading to simpler algorithms in term of computational requirements.

The process of simplification has also revealed that some existing algorithms which were developed by other methods are merely modifications of the EKF algorithm. Thus, this helps to understand better the behaviour of these algorithms. This understanding together with an effective use of the canonical form introduced in Chapter II makes it possible to develop a fast but bias-free identification algorithm if some condition is met. The requirement of the condition could be in most cases satisfied with à priori knowledge of the system.

The differences among the algorithms discussed in this thesis, namely, RML, GELS and EKF-M3 are worth noting. While in the RML and GELS algorithms, à priori errors are used to replace the true innovations, à postériori errors are used for the EKF-M3 algorithm. It is this à postériori error that is thought to be responsible for faster convergence of the EKF-M3 algorithm to the true values compared to the RML and GELS algorithms. Secondly, the RML algorithm incorporates a prefilter with time-varying parameters. On the other hand, the GELS and EKF-M3 algorithms employ a prefilter with fixed parameters. Theoretically, identification algorithms with time-varying

prefilter exhibit superior performance compared to algorithms with time-invariant filters, provided the parameters of the time-varying prefilter approach to the final desired values uniformly. In the case of the RML algorithm this uniformity property is not present in the prefilter. Consequently its convergence property is inferior to that of either the GELS or EKF-M3 algorithm as shown by the simulation studies in this thesis. Table 6-1 summarizes the properties of the various recursive algorithms discussed in this thesis.

It should also mentioned that the analysis of Ljung on the convergence of the RML algorithm (Ljung, 1979) though elegant, proves just the mere convergence of the algorithm and does not address the question whether the convergence is either uniform or asymptotic. One of the advantages of using a time-ivariant prefilter as in the EKF-M3 algorithm, is that it makes the analysis of the convergence property of the algorithm straightforward.

Though the most of discussions have been for multiinput singleoutput systems, the extension for multiinput multioutput systems can be easily made (Anderson, 1974; 1977).

METHOD	ADVANTAGE	DISADVANTAGE	CORRECTION	
ELS (Panuska)	Simple	Possible biased convergence Slow	introduction of prefilter	Use of à priori error
GELS	Global convergence with proper f vector	Slow		Time-invariant prefilter Use of à priori error
EKF-M2	Simple	Possible biased convergence	Introduction of prefilters	Use of à postériori error
EKF-M3	Global convergence with proper f vector			Time-invariant prefilter Use of à postériori error
RML (Soderstrom)	Local convergence	No global convergnce Possible biased convergence Slow	Introduction of projectors	Time-varying prefilter Use of à priori error
RML with Projector	Local convergence	Large amount of computation Slow		Time-varying prefilter + projector Use of à priori error

Table 6.1 Comparisons of recursive algorithms

6.2 Suggestions for Further Research

Fast but bias-free identification algorithms for the linear system using a systematic approach have been developed. But the ever-increasing use of on-line identification methods for systems in the various fields substantially expands the range of application of the identification techniques to systems in which more stringent conditions are often encountered; for example, it is often necessary to identify a system with very limited à priori knowledge of the system. Thus there is a need to improve the existing algorithms so as to handle this situation satisfactorily.

The employment of a time-varying prefilter instead of a time-invariant one used for the algorithm developed in this thesis seems extremely promising as an immediate improvement, thus eliminating the requirement of à priori information to guarantee unbiased convergence. One very important problem requiring an immediate attention is the unbiased convergence property of the modified form of the algorithm. This problem also demands a comprehensive study of a more general form of representation for the stochastic system to accomodate the time-varing prefilter.

Another approach which has a great potential is to treat the linear stochastic system as an unknown infinite impulse response (IIR) filter and use the standard least

squares method. This approach will involve the development of an efficient recursive least squares method for an infinite dimensional system and a fast algorithm to transform a polynomial of ever-inceasing order into a rational polynomial with a fixed order.

While this thesis has been devoted to the theoretical development of efficient recursive identification algorithms for stochastic systems, it is appropriate to mention that these algorithms have a wide application in the area of process control (Isermann, ed; 1980), communications (Friedlander, 1982) and processing of seismic data (Robinson, 1957; Mendel 1977; Mahalanabis, et al., 1981). obtained from geophysical exploration to mention a few.

More extensive computer simulations also needed for a more complete evaluation of the algorithms treated in this thesis. A comprehensive assessement of the performance characteristics including sensitivity to error in estimation system order and the effect of unusually large impulse noise, etc., which are of extreme importance in real applications should be made by means of simulation studies before implementation.

REFERENCES

Anderson B. D. O. (1974). Adaptive identification of multiple-input multiple-output plants. Proc. IEEE Conf. Decision and Control, 273-281.

Anderson B. D. O. (1977). An approach to multivariable system identification. Automatica, Vol. 13, 401-408.

Anderson B. D. O. and J. B. Moore (1979). Optimal filtering. Prentice-Hall, New Jersey.

Anderson T. W. (1962). The choice of the degree of a polynominal regression as a multiple decision problem. Ann. math. stat. 33, 255-265.

Aoki M. and R. M. Staley (1970). On input signal synthesis in parameter identification. Automatica, Vol. 6, 430-440.

Astrom K. J. and T. Bohlin (1966). Numerical identification of linear dynamic systems from normal operating records. in Theory of self-adaptive control systems. P. H. Hammond Ed., New York: Plenum

Astrom K. J. (1970). Introduction to stochastic control theory. Academic, New York.

Astrom K. J. and P. Eykhoff (1971). System identification: A survey. Automatica, Vol. 7, 123-162.

Athans M. (1967). The relationship of alternate state-space representation in linear filtering problems. IEEE Trans. Aut. Control AC-12, 775-776.

Bucy R. S. and P. D. Joseph (1968). Filtering for stochastic processes with application to guidance. Interscience, New York.

Bucy R. S. and K. D. Senne (1971). Digital synthesis of nonlinear filters. Automatica, Vol. 7, 287-298.

Chen C. T. (1970). Introduction to linear system theory. Holt, Rinehart and Winston, New York.

Cox H. (1964). On the estimation of state variables and parameters for noisy dynamic systems. IEEE Trans. Aut. Control AC-9, 5-12.

Dugard L. and I. D. Landau (1980). Recursive output error identification algorithms: Theory and evaulation. Automatica, Vol. 16, 443-463.

Egardt B. (1980). Unification of some discrete-time adaptive schemes. IEEE Trans. Aut. Control AC-25, 693-697.

Friedlander B. (1982). System identification technique for daptive signal processing IEEE Trans. Acoust. Speech, Signal Processing. ASSP-30, 240-246.

Geesey R. and T. Kailath (1969). Comment on "The relationship of alternate state-space representation in linear filtering problems". IEEE Trans. Aut. Control AC-14, 113-114.

Goodwin G. C., P. J. Ramdage and P. E. Caines (1979). Discrete time stochastic adaptive control. Proc. 18th IEEE Conf. Decision Contr. 736-739.

Hitz B. E. and B. D. O. Anderson (1969). Discrete positive real functions and their application to system stability. Proc. IEE, Vol. 116, 153-155.

Holst J. (1977). Adaptive prediction and recursive estimation. Div. Aut. Control, Lund Inst. Tec., Lund Sweden, Rep. 1013.

Isermann R., U. Baur, W. Bamberger, P. Kneppo and H. Siebert (1974). Comparison of six on-line identification and parameter estimation methods. Automatica, Vol. 10, 81-103.

Isermann R., ed (1980). IFAC Tutorial on "System Identification". Automatica, Vol. 16, 505-587.

Jazwinski A. H. (1970). Stochastic processes and filtering theory. Academic Press, New York.

Johnson C. R. (1980). An outpur error identification interpretation of model reference adaptive control. Automatica, Vol. 16, 419-421.

Johnson C. R., M. G, Larimore, J. R. Treichler and B. D. O. Anderson (1981). SHARF convergence properties. IEEE Trans. Circuit and Systems CAS-28, 499-510.

Jumarie G. (1979). Hyperstability and average hyperstability conditions for broad class of Gaussian stochastic systems. Information Sci. 17, 23-41.

Kailath T. (1968). An innovations approach to least squares estimation. Part I: Linear filtering in additive white noise. IEEE Trans. Aut. Control AC-13, 646-655.

Kalman R. E. (1960). A new approach to linear filtering and prediction problems. J. Basic Eng., Ser. D, 82, 35-45.

Kalman R. E. (1963). Mathematical description of linear dynamic systems. SIAM J. Control, 152-192.

Kashyap R. L. (1974). Estimation of parameters in a partially whitened representation of a stochastic process. IEEE Trans. Aut. Control AC-19,13-21.

Kim C. G. and V. Gourishankar (1981). A model reference technique of identification for discrete systems. Automatica, Vol. 17, 637-640.

Kopp R. E. and R. J. Orford (1963). Linear regression applied to system identification for adaptive control systems. AIAA J., Vol. 1, 2300-2306.

Kraft L. G. (1976). Controllable form state variables obtained from an arbitrarily fast adaptive observer. Proc. IEEE Conf. Decision and Control, 1088-1094.

Kreisselmeier G. (1977). Adaptive observers with exponential rate of convergence. IEEE Trans. Aut. Control AC-22, 2-8.

Kreisselmeier G. (1980). Adaptive control via adaptive observation and asymptotic feedback matrix synthesis. IEEE Trans. Aut. Control AC-25, 717-722.

Kudva P. and K. S. Narendra (1974). An identification procedure for discrete multivariable systems, IEEE Trans. Aut. Control AC-19, 549-552.

Kushner H. (1967). Stochastic stability and control, Academic Press, New York.

Kushner H. (1971). Introduction to stochastic control, Holt, Rinehart and Winston, New York.

Landau I. D. (1969). A hyperstability criterion for model reference adaptive control systems. IEEE Trans. Aut. Control AC-14, 552-555.

Landau I. D. (1974). A survey of model reference adaptive techniques-theory and and applications. Automatica, Vol. 10, 353-379.

Landau I. D. (1976). Unbiased recursive identification using model reference adaptive techniques. IEEE Trans. Aut. Control AC-21, 194-202.

Landau I. D. (1978). Elimination of the real positivity condition in the design of parallel MRAS. IEEE Trans. Aut. Control AC-23, 1015-1020.

Landau I. D. and H. M. Silveria (1979). A stability thorem with applications to adaptive control. IEEE Trans. Aut. Control AC-24, 305-312.

Landau I. D. (1981). Near supermartingales for convergence analysis of recursive identification and adaptive control schemes. Proc. 20th IEEE Conf. Decision Contr.

Landau I. D. (1982). Martingale convergence analysis of adaptive scheme - A feedback approach. IEEE Trans. Aut. Control AC-27, 716-719.

Lee R. C. K. (1964). Optimal Estimation, Identification and Control. MIT Press, Cambridge, MA.

Lin Y. H. and K. S. Narendra (1978). A new error model for discrete systems and its application to adaptive identification and control. S & IS Report No. 7802.

Lion P. M. (1966). Rapid identification of linear and nonlinear systems. Preprints of JACC, Washington, 605-615.

Ljung L., T. Soderstrom and I. Gustavsson (1975). Counterexamples to general convergence of a commonly used recursive identification method. IEEE Trans. Aut. Control AC-20, 643-652.

Ljung L. (1977a). On positive real transfer function and convergence of some recursive schemes. IEEE Trans. Aut. Control AC-22, 539-551.

Ljung L. (1977b). Analysis of recursive stochastic algorithms. IEEE Trans. Aut. Control AC-22, 551-575.

Ljung L., M. Morf and D. Falconer (1879). Fast calculation of gain matrices for recursive estimation schemes. Int. J. Control, Vol. 27, 1-19.

Ljung L. (1979). Asymptotic behavior of the extended Kalman filter as a parameter estimator of linear systems. IEEE Trans. Aut. Control AC-24, 36-50.

Luder G. and K. S. Narendra (1974). Stable adaptive schemes for state estimation and identification of linear systems. IEEE Trans. Aut. Control AC-19, 841-847.

Mahalanabis A. K., S. Prasad and K. P. Mohandas (1981). A fast optimal deconvolution algorithm for real seismic data using Kalman predictor model. IEEE Trans. Geosci. and Remote Sens. GE-19, 216-221.

Martin-Sanchez J. M. (1976). A new solution to adaptive control. IEEE Proc. Vol. 64, 1209-1218.

Mendel J. M. (1973). Discrete techniques of parameter estimation. New York: Dekker.

Mendel J. M. (1977). White noise estimator for seismic data processing in oil exploration. IEEE Trans. Aut. Control AC-22, 694-707.

Moore J. B. and H. Weiss (1979). Recursive prediction error method for adaptive estimation. IEEE Trans. Sys. Man and Cybern. SMC-9, 197-205.

Moore J. B. and H. Weiss (1980). Recursive prediction error algorithm without a stability test. Automatica, Vol. 16, 683-687.

Narendra K. S. and S. S. Tripathi (1973). Identification and optimization of aircraft dynamics. Journal Aircraft, Vol. 10, pp. 193-199.

Neveu J. (1975). Discrete-Parameter Martingales. Elsevier, New York.

Panuska V. (1968). A stochastic approximation method for identification on linear systems using adaptive filtering. Proc. Joint Aut. Control Conf.

Panuska V. (1969). An adaptive recursive least squares identification algorithm. 8th IEEE Symp. Adaptive Processes.

Panuska V. (1980a). A new form of the extended Kalman filter for parameter estimation in linear systems with correlated noise. IEEE Trans. Aut. Control AC-25, 229-235.

Panuska V. (1980b). Non-convergence of the approximate maximum likelihood identification algorithm. Automatica, Vol. 16, 215-217.

Popov V. M. (1962). Absolute stability of non-linear systems of automatic control. Automn. Remote Control 22, 857-875.

Robinson E. A. (1957). Predictive decomposition of seismic traces. Geophysics Vol. 22, 767-778.

Saridis G. (1974). Comparison of six on-line identification algorithms. Automatica, Vol. 10, 69-79.

Siljak D. D. (1968). Analytic test for absolutely stability. Electron Lett., 4, 358-359.

Siljak D. D. (1970). Algebraic criterion for absolute stability, optimality and passivity of dynamic systems.

Proc. IEE, Vol. 117, 2033-2036.

Soderstrom T. (1973). An on-line algorithm for approximate identification of linear dynamic systems. Div. Aut. Control, Lund Inst. Tec., Lund Sweden, Rep. 7308.

Soderstrom T. (1977). On model structure testing in system identification. Int. J. Control, Vol. 26, 1-8.

Soderstrom T., L. Ljung and I. Gustavsson (1978). A theoretical analysis of recursive identification algorithm. Automatica, Vol. 14, 231-244.

Solo V. (1979). The convergence of AML. IEEE Trans. Aut. Control AC-24, 958-963.

Unbenhauen H. and B. Gohring (1974). Tests for determining model order in parameter estimation. Automatica, Vol. 10. 233-244.

van den Boom A. J. W. and van den Enden (1974). The determination of the order of process and noise dynamics. Automatica, Vol. 10, 245-258.

Westphal L. C. (1978). An improved adaptive identification for discrete multivariable linear systems. IEEE Trans. Aut. Control AC-23, 860-865.

Whittaker H. P. (1958). Design of a model reference adaptive control for aircraft. Instrumentation Lab., MIT, Boston, Rep. R-164.

Young P. C. (1974). Recursive approaches to time-series analysis. Bull. Inst. Math. Applic., 10, 209-224.

Young P. C., A. Jakeman and R. McMurtrie (1980). An instrumental variable method for model order identification. Automatica, Vol. 16, 281-294.

Yuan J. S-C. and W. M. Wonham (1977). Probing signals for model reference identification. IEEE Trans. Aut. Control AC-22, 530-538.

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

Zames G. (1966). On the input-output stability of time varying non-linear feedback systems. IEEE Trans. Aut. Control AC-11, 228-238.

COMPUTER PROGRAMS

Following are the listings of the computer programs used in simulation studies of Chapter V. Since, as mentioned previously, the EKF-M3 and GELS algorithms differ only in the use of à postériori or à priori errors, these two algorithms are combined together into one and selected by a flag **ICH** of the routine **GELS**. The calculation of the gain vector of the correction term of the parameter estimation is performed with the fast algorithm proposed by Ljung, et al. (1978). The program **GELS** includs this fast algorithm. The rest of **GELS** program and entire **RML** program follow exactly the same notations as used in Chapter IV.

```
      SUBROUTINE GELS(YU,FV,P,ITER,IYU,NORD,NORDD,GAIN,ICH)
C
C
C     THESE ARE THE GENERALIZED EXTENDED LEAST SQUARES AND
C     EKF-M3 IDENTIFICATION ALGORITHMS FOR THE SYSTEM
C     WITHOUT CONTROL INPUT, i.e. TIME-SERIES, OR WITH
C     CONTROL INPUT.
C     THE CONTROL FLAG 'ICH' SWITCHES BETWEEN GELS AND
C     EKF-M3.
C     THE RESULTS WILL BE STORED IN A TEMPORARY FILE "-GELS"
C     IF GELS IS RUN OR "-EKF-M3" IF EKF-M3 IS RUN.
C
C
C     YU:   OUTPUT OF THE SYSTEM WHOSE PAPRAMETERS ARE TO BE
C           IDENTIFIED AND THEN THE CONTROL INPUTS ARE
C           IMMEDIATELY FOLLOWED IF THEY EXIST.
C     FV:   SYSTEM VECTOR OF THE SYSTEM MATRIX F OF THE
C           IDENTIFIER
C
C                                    T
C                          |      I  |
C           F =            | FV |  -  |
C                          |      0  |
C
C
C     P:      THIS VECTOR CONTAINS THE INITIAL ESTIMATES FOR
C             c, a and b VECTORS
C     ITER:   THE NUMBER OF ITERARTION DESIRED.
C     IYU:    THE NUMBER OF ENTRIES IN "YU" IF CONTROL INPUTS
C             EXISTS THEN IYU=ITER*2, OTHERWISE IYU=ITER
C     NORD:   ORDER OF THE SYSTEM
C     NORDD:  2*NORD IF NO INPUTS
C             3*NORD IF INPUTS EXIST
C     GAIN:   SCALAR VALUE FOR THE INITIAL GAIN MATRIX P22
C             P22(0) WILL BE GAIN*IDENTITY MATRIX
C     ICH:    SELECT GELS       IF ICH=0
C                    EKF-M3     IF ICH=1
C
C
      LOGICAL*1 DATA1(8),DATA2(8)
      REAL K(15),X(15),A(15,3),FV(NORD),D(15,3),XI(3),
     $     PRMT(15),YU(IYU),EPSI(3),WKA(3,4),ISIG(3,3),
     $     AMAX(15),AMIN(15),P(NORDD),EPSI0(3)
C
      DATA DATA1/'-','G','E','L','S',3*' '/
      DATA DATA2/'-','E','K','F','-','M','3',' '/
C
C
C     ATTACH THE PROPER TEMPORARY FILE FOR THE OUTPUTS
C
      IF(ICH .EQ. 0)CALL FTNCMD('ASSIGN 3=?;',0,DATA1)
      IF(ICH .EQ. 1)CALL FTNCMD('ASSIGN 3=?;',0,DATA2)
C
C     DETERMINE IF THERE ARE INPUTS TO THE SYSTEM
```

```
C
      NDATA=NORDD/NORD
C
C     SET SIG(0)=GAIN*IDENTITY MATRIX. THIS IS EQUIVALENT TO
C     SETTING P22(0)=GAIN*IDENTITY MATRIX. THUS, THE INVERSE
C     OF SIG(0), ISIG(0)=IDENTITY MATRIX/GAIN
C
      DO 1 I=1,NDATA
      DO 2 J=1,NDATA
      ISIG(I,J)=0.0
      IF (I .EQ. J) ISIG(I,J)=1.0/GAIN
    2 CONTINUE
    1 CONTINUE
C
C     SET  A(0) AND D(0) EQUAL TO ZERO
C
      DO 5 I=1,NORDD
      DO 6 J=1,NDATA
      A(I,J)=0.0
    6 D(I,J)=0.0
    5 CONTINUE
C
C     SET K(0) AND X(0) EQUAL TO ZERO
C
      DO 7 I=1,NORDD
      K(I)=0.0
    7 X(I)=0.0
C
C     SET INITIAL ESTIMATES P,  P=[FV-C,A-FV,B]
C
      DO 8 I=1,NORD
      P(I)=FV(I)-P(I)
    8 P(I+NORD)=P(I+NORD)-FV(I)
C
C     SET INITIAL VALUE OF AMAX AND AMIN EQUAL TO ZERO
C
      DO 9 I=1,NORDD
      AMAX(I)=-100.0
    9 AMIN(I)= 100.0
C
C
C
C
C
      DO 1000 KKK=1,ITER
C
C
      YH=0.0
      L=1
      DO 10 I=1,NDATA
      DO 11 J=1,NORD
      YH=YH+P(L)*X((J-1)*NDATA+I)
   11 L=L+1
   10 CONTINUE
```

```
C
C
      ERR=YU(KKK)-YH
C
C     SEE Eqn. (4.13) in Ljung's Paper in IJC
C
      AH=1.0
      IF (ICH .EQ. 0)GOTO 12
      AM3=0.0
      DO 15 I=1,NORDD
   15 AM3=AM3+K(I)*X(I)
      AH=AH-AM3
C
C
   12 DO 17 I=1,NORD
      PRMT(I)=FV(I)-P(I)
      PRMT(I+NORD)=FV(I)+P(I+NORD)
      IF (NDATA .EQ. 2)GOTO 17
      PRMT(I+NORD*2)=P(I+NORD*2)
   17 CONTINUE
C
C
      DO 18 I=1,NORDD
      AMAX(I)=AMAX1(AMAX(I),PRMT(I))
   18 AMIN(I)=AMIN1(AMIN(I),PRMT(I))
C
C
      WRITE(3,200)(PRMT(J),J=1,NORDD)
  200 FORMAT(1X,15F12.4)
C
C
      L=1
      DO 20 I=1,NDATA
      DO 21 J=1,NORD
      P(L)=P(L)+ERR*K((J-1)*NDATA+I)
   21 L=L+1
   20 CONTINUE
C
C     CALCULATION OF NEW COMER FOR SHIFT REGISTER "X"
C
      DO 25 I=1,NDATA
      S=0.0
      DO 26 J=1,NORD
   26 S=S+FV(J)*X((J-1)*NDATA+I)
      IF (I .EQ. 1)HIN=ERR*AH
      IF (I .EQ. 2)HIN=YU(KKK)
      IF (I .EQ. 3)HIN=YU(ITER+KKK)
   25 XI(I)=S+HIN
C
C  ************************************************************
C  *  THE ITERATIVE METHOD OF CALCULATION OF "K"              *
C  *  BY Ljung IS USED. SEE Int. J. of Contr. Vol. 27 No 1*
C  *  pp 1 19, 1978.                                          *
C  ************************************************************
```

```
C
C       CALCULATION OF "EPSILON0" OF Eqn. 4.1
C
        DO 30 I=1,NDATA
        S=0.0
        DO 31 J=1,NORDD
   31   S=S+A(J,I)*X(J)
   30   EPSI0(I)=XI(I)+S
C
C       UPDATE "A" OF Eqn. 4.2
C
        DO 33 I=1,NORDD
        DO 34 J=1,NDATA
   34   A(I,J)=A(I,J)-K(I)*EPSI0(J)
   33   CONTINUE
C
C       CALCULATION OF "EPSILON"  OF Eqn. 4.3
C
        DO 40 I=1,NDATA
        S=0.0
        DO 41 J=1,NORDD
   41   S=S+A(J,I)*X(J)
   40   EPSI(I)=XI(I)+S
C
C       UPDATE "ISIG" OF Eqn. 4.4. HERE THE MATRIX INVERSION
C       LEMMA HAS BEEN USED.
C
        S=1.0
        DO 90 I=1,NDATA
        WKA(I,1)=0.0
        WKA(I,2)=0.0
        DO 91 J=1,NDATA
        WKA(I,1)=WKA(I,1)+ISIG(I,J)*EPSI(J)
   91   WKA(I,2)=WKA(I,2)+EPSI0(J)*ISIG(J,I)
   90   S=S+WKA(I,2)*EPSI(I)
C
        DO 95 I=1,NDATA
        DO 96 J=1,NDATA
   96   ISIG(I,J)=ISIG(I,J)-WKA(I,1)*WKA(J,2)/S
   95   CONTINUE
C
C       CALCULATION OF K-bar OF Eqn. 4.5
C
        DO 45 I=1,NDATA
        S=0.0
        DO 46 J=1,NDATA
   46   S=S+ISIG(I,J)*EPSI(J)
   45   WKA(I,1)=S
C
        DO 48 I=1,NORDD
        S=0.0
        DO 49 J=1,NDATA
   49   S=S+A(I,J)*WKA(J,1)
   48   K(I)=K(I)+S
```

```
C
C  ************************************************
C  *  HEREAFTER "EPSI0" WILL BE  USED AS "ETA0", *
C  *            "XI" AS "XI(t-n)" in Eqn. 4.7    *
C  *            "EPSI" AS "MU" in Eqn. 4.6       *
C  *            "K" AS "M(t)" in Eqn. 4.6        *
C  ************************************************
C
C     UPDATE "X" AND COMPUTE "M(t)" and "MU(t)" in Eqn. 4.6
C
      NN=NORDD-NDATA
      DO 50 I=1,NDATA
      EPSI(I)=K(NN+I)
   50 WKA(I,2)=X(NN+I)
      DO 51 I=1,NN
      K(NORDD-I+1)=K(NN-I+1)
   51 X(NORDD-I+1)=X(NN-I+1)
      DO 52 I=1,NDATA
      X(I)=XI(I)
      K(I)=WKA(I,1)
   52 XI(I)=WKA(I,2)
C
C     CALCULATION OF "ETA0" OF Eqn. 4.7
C
      DO 55 I=1,NDATA
      S=0.0
      DO 56 J=1,NORDD
   56 S=S+D(J,I)*X(J)
   55 EPSI0(I)=XI(I)+S
C
C     UPDATE "D" OF Eqn. 4.8
C     FIRSTLY, THE SECOND TERM OF THE RIGHT SIDE IN Eqn. 4.8
C     AGAIN THE MATRIX INVERSION LEMMA HAS BEEN USED.
C
      S=1.0
      DO 58 I=1,NDATA
   58 S=S-EPSI(I)*EPSI0(I)
C
      DO 60 I=1,NDATA
      DO 61 J=1,NDATA
      WKA(I,J)=EPSI(I)*EPSI0(J)/S
      IF (I .EQ. J)WKA(I,J)=WKA(I,J)+1.0
   61 CONTINUE
   60 CONTINUE
C
C     THEN, THE FIRST TERM OF THE RIGHT SIDE IN Eqn. 4.8
C
      DO 65 I=1,NORDD
      DO 66 J=1,NDATA
   66 D(I,J)=D(I,J)-K(I)*EPSI0(J)
   65 CONTINUE
C
C     COMBINE THE ABOVE TWO TERMS
C
```

```
      DO 70 L=1,NORDD
      DO 71 I=1,NDATA
      WKA(I,4)=0.0
      DO 72 J=1,NDATA
   72 WKA(I,4)=WKA(I,4)+D(L,J)*WKA(J,I)
   71 CONTINUE
      DO 73 M=1,NDATA
   73 D(L,M)=WKA(M,4)
   70 CONTINUE
C
C     UPDATE "K"
C
      DO 75 I=1,NORDD
      S=0.0
      DO 76 J=1,NDATA
   76 S=S+D(I,J)*EPSI(J)
   75 K(I)=K(I)-S
C
C
C
C
 1000 CONTINUE
C
C
      WRITE(3,200)(AMAX(J),J=1,NORDD)
      WRITE(3,200)(AMIN(J),J=1,NORDD)
C
C
      WRITE(3,300)
  300 FORMAT(1X,'ESTIMATE OF C VECTOR IS FOLLOWED BY THAT
     $ OF A')
      WRITE(3,301)
  301 FORMAT(1X,'IN FORMAT F12.4')
      WRITE(3,302)
  302 FORMAT('  ')
C
      IF (ICH .EQ. 0)GOTO 303
      WRITE(3,304)
  304 FORMAT(1X,'THESE ARE OUTPUT OF EKF-M3')
      RETURN
  303 WRITE(3,305)
  305 FORMAT(1X,'THESE ARE OUTPUT OF GELS')
C
      RETURN
      END
```

```
      SUBROUTINE RML(YU,FV,P,ITER,IYU,NORD,NORDD,GAIN,ICH)
C
C
C     THIS IS THE RECURSIVE MAXIMUM LIKELIHOOD
C     IDENTIFICATION ALGORITHM COMPLIMENTED WITH
C     PROJECTOR WHICH PROJECTS THE C ESTIMATES INTO THE
C     PREDETERMINED AREA FOR THE SYSTEM WITHOUT CONTROL
C     INPUT, i.e. TIME-SERIES OR WITH CONTROL INPUT.
C     THE RESULTS WILL BE STORED IN A TEMPORARY FILE "-RML"
C
C
C     YU:     OUTPUT OF THE SYSTEM WHOSE PAPRAMETERS ARE TO
C             BE IDENTIFIED AND THEN THE CONTROL INPUTS ARE
C             IMMEDIATELY FOLLOWED IF THEY EXIST.
C     FV:     THE VECTOR FV IS ALL ZERO FOR RML, THUS FORCED
C             TO BE ZERO.
C     P:      THIS VECTOR CONTAINS THE INITIAL ESTIMATES FOR
C             c, a and b VECTORS
C     ITER:   THE NUMBER OF ITERARTIONS REQUIRED. i.e. THE
C             LENGTH OF ARRAY OF Y
C     IYU:    THE NUMBER OF ENTRIES IN "YU" IF CONTROL
C             INPUTS EXIST THEN IYU=ITER*2,
C                     OTHERWISE IYU=ITER
C     NORD:   ORDER OF THE SYSTEM
C     NORDD:  2*NORD IF NO INPUTS
C             3*NORD IF THERE ARE INPUTS
C     GAIN:   SCALAR VALUE FOR THE INITIAL GAIN MATRIX P22
C             P22(0) WILL BE EQUAL TO GAIN*IDENTITY MATRIX
C     ICH:    ICH IS ADDED FOR THIS ROUTINE TO BE FORMAT
C             COMPATIBLE WITH GELS
C
C
      LOGICAL*1 DATA1(8)
      REAL F0(5,5),H(5),P22(15,15),UP(5,15),UN(5,15),
     $     Z1(5),Z2(5),Z3(5),PRMT(15),FY(15),S(15),T(5),
     $     YU(IYU),TMAX(15),TMIN(15),FV(NORD),P(NORDD),
     $     CINIT(5),FC(5,5)
      DATA H/1.0,4*0.0/
      DATA UP/75*0.0/
      DATA DATA1/'-','R','M','L',4*' '/
C
C
C     ATTACH THE PROPER TEMPORARY FILE FOR THE OUTPUTS
C
      CALL FTNCMD('ASSIGN 3=?;',0,DATA1)
C
C     DETERMINE IF THERE ARE INPUTS TO THE SYSTEM
C
      INPUT=0
      IF (NORDD .EQ. 3*NORD) INPUT=1
C
C     SET UP F0
C
      DO 600 I=1,NORD
```

```
      DO 610 J=1,NORD
      F0(I,J)=0.0
      IF(I .EQ. J+1) F0(I,J)=1.0
 610  CONTINUE
 600  CONTINUE
C
      DO 620 I=1,NORD
 620  FV(I)=0.0
C
C     SET INITIAL GAIN MATRIX P22(0)=GAIN*IDENTITY MATRIX
C
      DO 1 I=1,NORDD
      DO 2 J=1,NORDD
      P22(I,J)=0.0
      IF (I .EQ. J) P22(I,J)=GAIN
   2  CONTINUE
   1  CONTINUE
C
C
      DO 5 I=1,NORD
      Z3(I)=0.0
      Z2(I)=0.0
   5  Z1(I)=0.0
C
C     SET INITIAL ESTIMATES P,  P=[FV-C,A-FV,B]
C
      DO 7 I=1,NORD
      P(I)=FV(I)-P(I)
   7  P(I+NORD)=P(I+NORD)-FV(I)
C
C     STORE THE INITIAL VALUES FOR ESTIMATE C
C
      DO 8 I=1,NORD
   8  CINIT(I)=P(I)
C
C     DO 650 I=1,NORDD
      TMAX(I)=-100.0
 650  TMIN(I)= 100.0
C
C
C
      DO 777 K=1,ITER
C
C
C
C
      DO 10 I=1,NORD
      FY(I)=Z1(I)
      FY(I+NORD)=Z2(I)
      IF (INPUT .EQ. 0) GOTO 10
      FY(I+NORD*2)=Z3(I)
  10  CONTINUE
C
C
```

```
      YH=0.0
      DO 20 I=1,NORDD
   20 YH=YH+FY(I)*P(I)
C
C
      ERR=YU(K)-YH
C
C     CALCULATION OF FC = F0 - H*HTHETA2
C
      DO 40 I=1,NORD
      DO 42 J=1,NORD
   42 FC(I,J)=F0(I,J)-H(I)*P(J)
   40 CONTINUE
C
C     CALCULATION OF UP(k+1) in eqn. (4.24)
C
      DO 410 I=1,NORD
      DO 420 J=1,NORDD
      UN(I,J)=0.0
      DO 430 L=1,NORD
  430 UN(I,J)=UN(I,J)+FC(I,L)*UP(L,J)
  420 UN(I,J)=UN(I,J)-H(I)*FY(J)
  410 CONTINUE
C
C     CALCULATION OF S = P22(k)UP(k+1)'h in eqn. (4.28d)
C
      DO 450 I=1,NORDD
      S(I)=0.0
      DO 460 J=1,NORDD
  460 S(I)=S(I)+P22(I,J)*UN(1,J)
  450 CONTINUE
C
C     CALCULATION OF m1(k) with w=1 in eqn. (4.27)
C
      AM1=1.0
      DO 58 I=1,NORDD
   58 AM1=AM1+UN(1,I)*S(I)
      AM1=1.0/AM1
C
C     UPDATE Z1(k)
C
      DO 70 I=1,NORD
      T(I)=0.0
      DO 72 J=1,NORD
   72 T(I)=T(I)+F0(I,J)*Z1(J)
   70 T(I)=T(I)+H(I)*ERR
C
      DO 74 I=1,NORD
   74 Z1(I)=T(I)
C
C     UPDATE Z2(k)
C
      DO 80 I=1,NORD
      T(I)=0.0
```

```
      DO 82 J=1,NORD
   82 T(I)=T(I)+F0(I,J)*Z2(J)
   80 T(I)=T(I)+H(I)*YU(K)
C
      DO 84 I=1,NORD
   84 Z2(I)=T(I)
C
C     UPDATE Z3(k)
C
      DO 86 I=1,NORD
      T(I)=0.0
      DO 87 J=1,NORD
   87 T(I)=T(I)+F0(I,J)*Z3(J)
   86 T(I)=T(I)+H(I)*YU(K+ITER)
C
      DO 88 I=1,NORD
   88 Z3(I)=T(I)
C
C     UPDATE THE ESTIMATE P(k) in eqns (4.28a) and (4.28d)
C
      DO 90 I=1,NORDD
   90 P(I)=P(I)-S(I)*AM1*ERR
C
C     TEST FOR ESTIMATE FOR VECTOR C
C
      CALL RAIBLE(P,NORD,ITEST)
      IF(ITEST .EQ. 0)GOTO 102
      DO 110 I=1,NORD
  110 P(I)=CINIT(I)
C
C
  102 DO 550 I=1,NORD
      PRMT(I)=FV(I)-P(I)
      PRMT(I+NORD)=FV(I)+P(I+NORD)
      IF (INPUT .EQ. 0)GOTO 550
      PRMT(I+NORD*2)=P(I+NORD*2)
  550 CONTINUE
C
C
      DO 146 I=1,NORDD
      TMAX(I)=AMAX1(TMAX(I),PRMT(I))
  146 TMIN(I)=AMIN1(TMIN(I),PRMT(I))
C
C
      WRITE(3,200)(PRMT(J),J=1,NORDD)
  200 FORMAT(1X,15F12.4)
C
C     UPDATE THE GAIN MATRIX P22(k)
C
      DO 120 I=1,NORDD
      DO 122 J=1,NORDD
  122 P22(I,J)=P22(I,J)-S(I)*S(J)*AM1
C
C     MOVE U(k+1) TO U(k)
```

```
C
      DO 124 L=1,NORD
  124 UP(L,I)=UN(L,I)
  120 CONTINUE
C
C
C
C
  777 CONTINUE
C
C
      WRITE(3,200)(TMAX(J),J=1,NORDD)
      WRITE(3,200)(TMIN(J),J=1,NORDD)
C
C
      WRITE(3,300)
  300 FORMAT(1X,'ESTIMATE OF C VECTOR IS FOLLOWED BY
     $ THAT OF A')
      WRITE(3,301)
  301 FORMAT(1X,'IN FORMAT F12.4')
      WRITE(3,302)
  302 FORMAT('  ')
C
      WRITE(3,305)
  305 FORMAT(1X,'THESE ARE OUTPUT OF RML')
C
      RETURN
      END
C
C
C
C
      SUBROUTINE RAIBLE(A,N,ITEST)
C
C
C     THIS PROGRAM TESTS THE STABILITY OF THE POLYNOMIAL
C     F(Z) IN THE FORM
C
C           N      N-1
C          Z  + A1Z   + ..... + An
C
C     A:      AN ARRAY CONTAINING THE COEFFICIENTS OF F(Z)
C     N:      THE DEGREE OF THE POLYNOMIAL F(Z)
C     ITEST:  0  F(Z) IS STABLE
C             1  F(Z) IS UNSTABLE OR SIGULAR CASE
C
C
C
C
      REAL A(N),AA(6),BB(5)
      NN=N+1
C
      DO 10 I=2,NN
   10 AA(I)=A(I-1)
```

```
      AA(1)=1.0
C
      DO 70 K=1,N
      AK=AA(NN)/AA(1)
      NN=NN-1
C
      DO 30 I=1,NN
   30 BB(I)=-AK*AA(NN-I+2)
C
      DO 40 I=1,NN
   40 AA(I)=AA(I)+BB(I)
C
      IF(AA(1) .GT. 0.0) GOTO 70
      ITEST=1
      RETURN
C
   70 CONTINUE
C
      ITEST=0
C
C
      RETURN
      END
```