THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

NON-LINEAR ADAPTIVE FILTERING AND CONTROL WITH APPLICATION TO CHEMICAL REACTOR SYSTEMS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

WARREN HWA-NAN HUANG

Norman, Oklahoma

1971

NON-LINEAR ADAPTIVE FILTERING AND CONTROL

WITH APPLICATION TO CHEMICAL REACTOR SYSTEMS

APPROVED BY

Michael Fu

James A. Payne

Kenneth E. Starling

James Christensen

Harry H West

DISSERTATION COMMITTEE

DEDICATED TO

Chi-Shen Shih Huang, my Mother, in memoriam

ACKNOWLEDGEMENTS

I am grateful to the persons who have helped me in various ways:

Professor M. L. McGuire -- for his advice and encouragement in the course of this research.

Professor J. H. Christensen -- for his helpful suggestions in real time implementation of this work.

Dr. H. H. West -- for his generous assistance during my stay at the University of Oklahoma.

Professors K. E. Starling, J. A. Payne and H. H. West -- for serving on my advisory committee.

Messers. K. Sanders and W. Woodard -- for their helpful discussions.

Chem Share Corporation and Computer Design Data, Inc. -- for their financial support through part-time work.

Time Sharing Terminal of Chemical Engineering Department -- for providing part of calculation of this work.

University of Oklahoma Computer Center -- for donating computer time.

My parents -- for their love, constant encouragement and sacrifice.

My wife -- for her devotion, patience and understanding.

Mrs. Mary Lou Stokes -- for her typing of the dissertation.

TABLE OF CONTENTS

		Page
ACKNOWLEDGEMENTS		iii
LIST OF TABLES		vi
LIST OF ILLUSTRATIONS		vii
Chapter		
I.	INTRODUCTION	1
	1. Introduction	1
	2. Scope and Objective	3
II.	FILTERING THEORY	5
	1. Introduction	5
	2. Discrete Linear Kalman Filter	6
	3. Non-linear Discrete Filtering	16
	4. Numerical Examples and Discussions	21
III.	COMBINED FILTERING AND CONTROL	44
	1. Introduction	44
	2. Combined Linear Filtering and Linear Control with A Quadratic Performance Index	49
	3. Suboptimal Non-linear Stochastic Control	53
IV.	MODELING ERROR COMPENSATION AND ADAPTIVE FILTERING	64
	1. Introduction	64
	2. Modeling Errors and Filter Divergence	65
	3. Modeling Error Compensation Techniques	70
	4. On-line Adaptive Filtering	80
	5. Numerical Examples and Discussions	96
V.	OBSERVABILITY, CONTROLABILITY AND INFORMATION MATRIX	107
	1. Introduction	107
	2. Information Matrix and Observability	108
	3. Controlability	109
	4. Bounds for the Estimation Error Covariance Matrix	112
	5. Numerical Examples and Discussions	113
VI.	STOCHASTIC STABILITY	122
	1. Introduction	122
	2. Some Forms of Stochastic Functions	124

TABLE OF CONTENTS (CONT'D)

3. Properties of the Riccati Equation 126
4. Numerical Examples 130

VII. ERROR SENSITIVITY ANALYSIS AND FILTER ADAPTIVITY. . . 133

1. Introduction 133
2. Development of General Sensitivity Analysis Algorithm 135
3. Error Sensitivity Due to Prior Statistics. 140
4. Sensitivity Analysis Due to Model Errors 148
5. Error Sensitivity Due to Measurement Error 156
6. Relation to Adaptivity 165
7. Numerical Examples and Discussions 166

VIII. ERRORS IN FILTERING 167

1. Introduction 167
2. Modeling Error and Model Approximation 168
3. Round Off Errors in Matrix Inversion 172
4. Numerical Examples and discussions 173

IX. REAL TIME IMPLEMENTATION OF FILTERING 187

X. SUMMARY AND CONCLUSIONS 192

BIBLIOGRAPHY 194

APPENDICES 198

A. PROGRAM FLOW CHART 198

B. SAMPLE COMPUTER OUTPUT 204

LIST OF TABLES

Table		Page
II-1.	Filtering results of falling body system	11
IV-1.	Approximate the 4th order model by a second order using Huang's II (iterated fixed point smoothing) sample output for smoothing the initial condition uncertainties.	93
IV-2.	Model error compensation technique and adaptive filtering	106
VIII-1.	Summary of adaptive filters.	191

LIST OF ILLUSTRATIONS

Figure		Page
2.1	Second order non-linear filter, P11(0) = 0.005, R11(k) = 0.01, plot of X1.	25
2.2	Second order non-linear filter P11(0) = 0.01, R11(k) = 0.01, plot of X4.	26
2.3	Second order non-linear filter P11(0) = 0.01, R11(k) = 0.01, plot of X1.	27
2.4	Second order non-linear filter P11(0) = 0.015, R11(k) = 0.01, plot of X1.	28
2.5	Second order non-linear filter P11(0) = 0.015, R11(k) = 0.01, plot of X4.	29
2.6	Extended Kalman filter, P11(0) = 0.015, R11(k) = 0.01, plot of X1.	30
2.7	Extended Kalman filter, P11(0) = 0.005, R11(k) = 0.01, plot of X4.	31
2.8	Extended Kalman filter, P11(0) = 0.005, R11(k) = 0.01, plot of X1.	32
2.9	Second order non-linear filter, P11(0) = 0.01, R11(k) = 0.03, plot of X1.	37
2.10	Second order non-linear filter, P11(0) = 0.01, R11(k) = 0.03, plot of X4.	38
2.11	Extended Kalman filter, P11(0) = 0.005, R11(k) = 0.01, plot of X4.	39
2.12	Extended Kalman filter, P11(0) = 0.01, R11(k) = 0.03, plot of X1.	40
2.13	Double the sampling period, second order non-linear filter. Plot of X3.	42
2.14	Double the sampling period, second order non-linear filter. Plot of X4.	43

LIST OF ILLUSTRATIONS (CONT'D)

3.1 Bang-bang control, second order non-linear filter and Huang's I plot of X3 60

3.2 Bang-bang control, second order non-linear filter and Huang's I plot of X4 61

3.3 Bang-bang control, extended Kalman filter, with Rk = 6.0, plot of X1 62

3.4 Bang-bang control, extended Kalman filter, with Rk = 6.0, plot of X4 63

4.1 Filter divergence, extended Kalman filter, with 20 per cent rate constant error, without compensation plot of X1. 67

4.2 Extended Kalman filter with 20 per cent rate constant error, without compensation, plot of X4. 68

4.3 Second order non-linear filter with 20 per cent rate constant error, plot of X2 69

4.4 Huang's adaptive III to compensation for 20 per cent rate constant error, plot of X1. 77

4.5 Huang's adaptive III to compensate for model errors, plot of X2 78

4.6 Huang's adaptive III, plot of X4. 79

4.7 Huang's adaptive IV, to compensate for 20 per cent rate constant error, plot of X1 98

4.8 Huang's adaptive IV to compensate for 20 per cent rate constant error, plot of X4 99

4.9 Schmidt filter to compensate for model error, plot of X1. 100

4.10 Model error compensation by Wolf's II, plot of X4 101

4.11 Model error compensation by constant noise, plot of X2. . 102

4.12 Model error compensation by constant noise, plot of X4. . 103

4.13 Model error compensation by Wolf's I, plot of X1. 104

4.14 Model error compensation by Wolf's I, plot of X4. 105

LIST OF ILLUSTRATIONS (CONT'D)

5.1 Error bound plot for P11 filter divergence. 114

5.2 Error bound plot for P22 filter divergence. 115

5.3 Error bound plot for P33 filter divergence. 116

5.4 Error bound plot for P44 filter divergence. 117

5.5 Error bound plot for P11 constant noise 118

5.6 Error bound plot for P22 constant noise 119

5.7 Error bound plot for P33 constant noise 120

5.8 Error bound plot for P44 constant noise 121

7.1 Sensitivity plot for P11 due to prior statistics. 142

7.2 Sensitivity plot for P44 due to prior statistics. 143

7.3 Sensitivity plot for P44 due to measurement noise error . 144

7.4 Sensitivity plot for P33 due to measurement noise error . 145

7.5 Sensitivity plot for P44 due to model and prior statistics error 146

7.6 Sensitivity plot for P11 due to model and prior statistics error 147

7.7 Sensitivity plot for P11 due to integration error 151

7.8 Sensitivity plot for P22 due to integration error 152

7.9 Sensitivity plot for P33 due to integration error 153

7.10 Sensitivity plot for P44 due to integration error 154

7.11 Sensitivity plot for P11 due to model error (constant noise) 158

7.12 Sensitivity plot for P11 due to model error (without compensation). 159

7.13 Sensitivity plot for P22 (constant noise) 160

7.14 Sensitivity plot for P44 (dimension reduction). 161

LIST OF ILLUSTRATIONS (CONT'D)

7.15 Sensitivity plot for P33 due to measurement matrix error. . 162

7.16 Sensitivity plot for P22 due to measurement matrix error. . 163

7.17 Sensitivity plot for P44 due to measurement matrix error. . 164

8.1 This work I supplied to treat integration error, plot of X1 170

8.2 Double the sampling period using filter decomposition plot of X1 178

8.3 Use filter decomposition to treat model error plot of X1. . 179

8.4 This work I applied to treat integration error, plot of X4 180

8.5 This work III applied to model dimension reduction, plot of X4. 181

8.6 This work III applied to dimension reduction, plot of X1. . 182

8.7 Constant noise compensation to perform dimension reduction, plot of X1 183

8.8 This work IV (overweight by stochastic approximation) applied to dimension reduction plot of X1. 184

8.9 This work IV applied to dimension reduction, plot of X4 . . 185

ABSTRACT

The increasing demand in modern system design and control to intelligently use information from observations in order to make better decisions is stimulated by the better understanding of the theory of stochastic process and computer hardware improvements.

This research presents an extensive study of both the extended non-linear Kalman filter and the second order approximate filter. Theoretical and practical aspects are emphasized in the areas such as comparisons of non-linear filter techniques, combined optimal filtering and stochastic control, filter asymptotic stability, error sensitivity analysis, bound of estimation error covariance matrix, off-line model error compensation and on-line adaptive filtering and filter decomposition and its applications to real-time filtering for large complex systems.

As a result of this study, a sophisticated computer program is developed to treat each of the above subjects and applied to the filtering of a stirred tank reactor for both steady state feed-forward control and optimum bang-bang control. The reactor is modeled so that it has four state variables.

Conclusive results are obtained on each of the subject. After this extensive study, we are more confident of applying the non-linear filtering to actual process. This includes the implementation of real-time filtering since our filter decomposition algorithm can

ABSTRACT

The increasing demand in modern system design and control to intelligently use information from observations in order to make better decisions is stimulated by the better understanding of the theory of stochastic process and computer hardware improvements.

This research presents an extensive study of both the extended non-linear Kalman filter and the second order approximate filter. Theoretical and practical aspects are emphasized in the areas such as comparisons of non-linear filter techniques, combined optimal filtering and stochastic control, filter asymptotic stability, error sensitivity analysis, bound of estimation error covariance matrix, off-line model error compensation and on-line adaptive filtering and filter decomposition and its applications to real-time filtering for large complex systems.

As a result of this study, a sophisticated computer program is developed to treat each of the above subjects and applied to the filtering of a stirred tank reactor for both steady state feed-forward control and optimum bang-bang control. The reactor is modeled so that it has four state variables.

Conclusive results are obtained on each of the subject. After this extensive study, we are more confident of applying the non-linear filtering to actual process. This includes the implementation of real-time filtering since our filter decomposition algorithm can

treat any large complex system by partitioning into several small manageable subsystems, then apply different adaptive filtering techniques to the individual subsystem. The error sensitivity analysis and estimation error covariance bound calculation not only provide us with the interactions among the variables and parameters but also lead to better understanding of the structure of the system which is essential to filter decomposition or model dimension reduction (approximate a higher dimension complex system by a lower order).

With the aid of the adaptive filters developed by this work, we are able to perform the filtering under various uncertainties.

We believe that this study will serve a useful guide for the future theoretical and practical development of non-linear filtering and control.

NONLINEAR ADAPTIVE FILTERING AND CONTROL WITH APPLICATION TO CHEMICAL REACTOR SYSTEMS

CHAPTER I

INTRODUCTION

I.1 Introduction.

Filtering is the estimation of the state variables or parameters of a system based on the measurement observations when random disturbances appear in both the system and the measurements.

Following the work of Kalman and Bucy [K1] in linear filtering, many papers have appeared treating linear filtering via 'least squares', i.e., the estimate which minimizes the mean square error. Others have used the'maximum likelihood' principle which maximizes the likelihood function (the probability density function for a parameter given the measurement of that parameter). In addition, there have been many studies of the extension of non-linear filters [S6], [B1], [B8].

The reason that the so-called Kalman-Bucy filter has received a great deal of attention, both in the academic and aerospace fields, is due to its sequential and recursive computational approach which allows off-line digital calculations as well as on-line filtering of observations.

In spite of the wide spread applications in the aerospace field, there are still differences in deriving a Kalman filter. As stated by Jazwinski [J2], this is probably due to a lack of a thorough understanding of Kalman's original work [K3]. Such important points in linear theory as stability and error sensitivity analysis modeling techniques have been ignored. Due to the urgent need existing in missile guidance, a great deal of emphasis has been placed on the practical aspects, while neglecting the fundamental probabilistic structuring of the filtering. This drawback has become increasingly important in dealing with non-linear filtering.

A good review made by Jazwinski [J2] presents a complete coverage of filtering theory and error analysis together with applications to guidance.

Despite the widespread attention to filtering in the aerospace industry, it has been only within the last two years that filtering techniques have stimulated the interest of chemical engineers. This has been due to the complexity of a chemical process (which has predominantly non-linear characteristics) coupled with the lack of progress in the theory of non-linear filtering. However, there have been many recent developments in the theory of approximate non-linear filtering. However, no conclusive results are available, either in the theory or in applications to general problems. Bucy [B8] made a recent contribution in developing some asymptotic properties and introduced a partial solution to non-linear filtering problem.

Wells [W1] and Mehra[M3] have studied filtering and adaptive techniques. The former has been applied to chemical reactor modeling and control while the latter was applied only to a linear model. This limitation was due to the large amount of computation in generating the noise covariance and filter gain matrix.

However, there is an urgent need to conduct a systematic study by applying Kalman's original theory and its extensions to the approximate non-linear filtering problem, i.e., studying asymptotic stability, error sensitivity analysis and various aspects in modeling errors, and error compensation. These should be carried out in order to achieve a unified approach and conclusive results. Hopefully this will provide a guide to the applications of non-linear filtering.

Thus, with this aim, this study is an attempt to apply the above aspects to the filtering and control of a reactor. Both the theory and numerical results are emphasized in order to obtain the unified conclusions.

Our final goal in this study is to develop an online non-linear adaptive filter which will achieve optimal filtering with linear control according to the separation principle and suboptimal filtering with non-linear control.

I.2 Scope and Objective.

The objective in conducting this extensive study is:

1. To introduce error compensation techniques into the non-linear extended Kalman filter and the second order filter to compensate for model errors due to uncertain parameters, deliberate model simplifications and any error due

to nonlinear approximation in filter model. All of these aspects are applied to the chemical reactor model.

2. To extend the results obtained in linear filtering to the non-linear filtering problem in a unified manner. This includes such topics in linear filtering as stability, model error sensitivity, error bounds all with applications to a chemical reactor operating under steady state control or optimum bang-bang control.

3. To present a state-of-the-art review of the non-linear filtering theory. This includes the development of the second order non-linear filter and comparison of this filter with non-linear extended Kalman filter when applied to a chemical reactor.

4. To develop a sophisticated filter decomposition algorithm to partition a large complex system into several subsystems and applying the adaptive filtering techniques developed by this work to the individual subsystems thus providing a unique approach to the on-line filtering of large system.

CHAPTER II

FILTERING THEORY

II.1 Introduction.

Most of the studies of chemical process dynamics and control have been directed toward deterministic systems, assuming precise knowledge of the system exists by describing the system with a specific mathematical model with a set of physical parameters. However, in actual operation, uncertainties arise because of random fluctuations in feed compositions, other changes in the input such as temperatures, and imprecise knowledge of physical parameters such as chemical reaction rate constants. Very often, due to the corruption of measurement by inaccurate sensor instruments, experimental errors can affect the process dynamics significantly.

The extended Kalman estimation technique has been applied successfully to many aerospace systems such as the Apollo project. Recent feasibility studies by Larson [L2], in applying these techniques to the power industry, have shown encouraging results. This has stimulated academic interest in the chemical engineering field.

Before applying the estimation technique to a real chemical process plant, it is necessary to acquire a sound knowledge of filtering theory, i.e., from the study of basic linear filtering, stability, and error bounds, as well as its applicability to the extensions to nonlinear systems.

II.2 Discrete Linear Kalman Filter.

Since most studies of chemical process simulation have been conducted on digital computer which operates with discrete data we will introduce the discrete version of the Kalman filter [J2].

Given the dynamic system in the state variable form:

$$\underline{x}(k+1) = \Phi(k+1,k)\underline{x}(k) + \Gamma(k)\underline{w}(k+1). \qquad \text{II-1}$$

Where

$\underline{x}(k)$ = nx1 state vector at t_k

$\Phi(k+1,k)$ = nxn state transition matrix (from k to k+1)

$\Gamma(k)$ = nxr system noise gain matrix

$w(k)$ = rx1 Gaussian white system noise

and the observation

$$\underline{z}(k) = H(k)\underline{x}(k) + \underline{v}(k) \qquad \text{II-2}$$

with Gaussian white noise on the system and observation, i.e.,

$$E[\underline{w}(k)] = 0 \quad \text{and} \quad E[\underline{w}(k)\underline{w}'(j)] = Q(k)\delta_{kj}$$

$$E[\underline{v}(k)] = 0 \quad \text{and} \quad E[\underline{v}(k)\underline{v}'(j)] = R(k)\delta_{kj} .$$

Where

$\underline{z}(k+1)$ = mx1 Measurement observation at t_{k+1}

$\underline{v}(k+1)$ = mx1 Gaussian white measurements noise

$Q(k+1)$ = rxr process noise covariance matrix

$R(k+1)$ = mxm Measurement noise covariance matrix

$H(k+1)$ = mxn Measurement matrix .

Then the new estimate and their error covariance matrices are calculated as follows:

1. Prediction Stage:

   a) State Prediction:

$$\hat{x}(k+1,k) = \Phi(k+1,k)\hat{x}(k,k) \qquad \text{II-3}$$

where

$\hat{x}(k+1,k)$ = nx1 State estimate at t_{k+1}

   b) Error Covariance Matrix:

$$P(k+1,k) = \Phi(k+1,k)P(k,k)\Phi'(k+1,k) + \Gamma(k)Q(k+1)\Gamma'(k). \qquad \text{II-4}$$

2. Process the observation - and make correction on the estimate

   a) Filter Gain Matrix:

$$K(k+1) = P(k+1,k)H'(k+1)[H(k+1)P(k+1,k)H'(k+1) + R(k+1)]^{-1} \qquad \text{II-5}$$

   b) Correction of the estimate:

$$\underline{\hat{x}}(k+1,k+1) = \underline{\hat{x}}(k+1,k) + K(k+1)[\underline{\hat{z}}(k+1)-H(k+1)\underline{\hat{x}}(k+1,k)] \qquad \text{II-6}$$

   c) New Error Covariance Matrix Calculation:

$$P(k+1,k+1) = [I-K(k+1)H(k+1)]P(k+1,k) \qquad \text{II-7}$$

or

$$P(k+1,k+1) = [I-K(k+1)H(k+1)]P(k+1,k)[I-K(k+1)H(k+1)] + K(k+1)R(k+1)K'(k+1) . \qquad \text{II-8}$$

Where

$\underline{x}(k+1,k+1)$ = nx1 state estimate at t_{k+1} given $\underline{z}(k)$

P(k,k) = nxn covariance matrix of the error in $\hat{\underline{x}}(k,k)$

P(k+1,k) = nxn covariance matrix of the error in $\hat{\underline{x}}(k+1,k)$

K(k+1) = nxm Kalman filter gain matrix at t_{k+1} .

Proof of the Equivalance of Equations II-7 and II-8

If we drop the time index k on the right hand side of Equation II-8 for convenience, then expand and regroup

$$P(k+1,k+1) = (P-KHP)(I-KH)' + KRK'$$
$$= P - KHP - PH'K' + KHPH'K' + KRK'$$
$$= (I-KH)P - PH'K' + K(HPH'+R)K' \quad .$$

From the definition of steady-state Kalman filter gain matrix

$K(HPH'+R) = PH'$, it follows that

$$P(k+1,k+1) = (I-KH)P - PH'K' + PH'K'$$
$$= [I-K(k+1)H(k+1)]P(k+1,k)$$

which is Equation II-7.

Equations II-7 and II-8 are equivalent. However, II-8 is better conditioned in the sense of maintaining the symmetric and positive definite properties of the error covariance matrix and is less sensitive to errors in calculating the gain matrix. This can be seen from II-7 that

$$\delta P(k+1,k+1) = - \delta K(k+1)H(k+1)P(k+1,k) \ ,$$

while from II-8

$$\delta P(k+1,k+1) = 0 \ .$$

However, this is not necessarily true in the non-linear filtering case based on our simulation experience, II-8 often causing filter divergence.

The following system is used to illustrate the formulation. For a noise-free, second order system representing a falling body in a constant gravitational field the describing equations are

$$\ddot{x} = - g(t \geq 0)$$

which become in state space form:

$$\dot{\underline{x}} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ -g \end{bmatrix} ; \qquad F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

where $\underline{x} = (x_1, x_2)'$; $x_1 = x$; $x_2 = \dot{x}$. The state transition matrix is

$$\Phi(t,\tau) = \begin{bmatrix} 1 & t-\tau \\ 0 & 1 \end{bmatrix}$$

so that

$$\underline{x}(t) = \Phi(t,\tau)\underline{x}(\tau) + \int_{\tau}^{t} \begin{bmatrix} 0 & t-\tau \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -g \end{bmatrix} d\tau$$

and

$$\underline{x}(t+1) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \underline{x}(t) - g \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}; \qquad t = 0,1 \ \ldots \ .$$

The scalar observation of position is

$$\underline{z}(t) = (1 \quad 0)\underline{x}(t) + \underline{v}(t); \qquad t = 1,\ldots 6; \qquad M = (1 \quad 0)$$

with error covariance R = 1.0. The initial estimate is $\underline{\hat{x}}(0) = \begin{pmatrix} 95 \\ 1 \end{pmatrix}$ and the error covariance is

$$P(0) = \begin{pmatrix} 10 & 0 \\ 0 & 1 \end{pmatrix} .$$

The system is observable since the information matrix is non-singular.

$$I(t,t-1) = \sum_{i=t-1}^{t} \Phi'(i,t)H'(i)R^{-1}(i)H(i)\Phi(i,i)$$

$$I_n = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} .$$

Let g = 1 and the true initial condition is $x_1(0) = z(0) = 100$, $x_2(0) = 0.0$.

A complete output summary is listed in Table II-1. As stated by Jazwinski [J2], the velocity estimation error is somewhat outside 1 standard deviation is due to the particular choice of x(0) and the measurements. This initial transient will disappear eventually; we observe that the position error drops fast as soon as the first observation is processed, while the velocity error drops until the second observation. This occurs because two position observations are required to determine both components of the state vector. The dynamics of this system are such that velocity affects the position but not vice versa.

The Gaussian Markov property of the linear Kalman filter make the transients due to the initial uncertainty in P(0) and x(0) disappear eventually and the estimation error approaches zero as the filter

Table II-1. Filtering results of falling body system.

time t	0	1	2	3	4	5
$x_1(t)$	100.0	99.600	98.00	95.50	92.00	82.00
$x_2(t)$	0.0	-1.000	-2.00	-3.00	-4.00	-6.00
$z(t)$	0.0	100.000	97.90	94.40	92.70	82.10
$x_1(t,t)$	95.0	99.600	98.40	95.20	92.30	82.10
$x_2(t,t)$	1.0	0.370	-1.15	-2.90	-3.69	-5.80
$K_1(t)$		0.920	0.66	0.65	0.61	0.51
$K_2(t)$		0.083	0.33	0.31	0.23	0.15
$P_{11}(t,t)$	10.0	0.920	0.66	0.66	0.61	0.51
$P_{22}(t,t)$	1.0	0.920	0.58	0.29	0.16	0.05

approaches a steady state.

For the example problem with a different value for P(0)

$$P'(0) = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$$

our initial difference in P is

$$P(0) - P'(0) = \begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix} .$$

The difference at the end of the 6th period is

$$P(0)-P'(0) = \begin{bmatrix} 0.02 & 0.01 \\ 0.01 & 0.005 \end{bmatrix} .$$

Different approaches have been used by various authors to derive the above Kalman-Bucy filter. A brief list is as follows:

1. Minimum variance [B3], [S6]: This approach minimizes the estimate error variance. For example

   $$\text{Min } E\left([\underline{x}(k)-\underline{\hat{x}}(k)]'[\underline{x}(k)-\underline{\hat{x}}(k)]\right) .$$

   This can be done easily by minimizing each component of the expected error through the orthogonal projection principle, i.e., the prediction error is orthogonal to the observation.

2. Recursive Least Squares: This method has been used by Bryson and Ho [H2] and others [J2] to derive the discrete Kalman-Bucy filter. This can be done by converting the filtering problem to a least squares problem.

$$\text{Min } \tfrac{1}{2}[\underline{x}(0)-\hat{\underline{x}}(0)]P_o^{-1}[\underline{x}(0)-\hat{\underline{x}}(0)] +$$

$$\tfrac{1}{2}\sum_{i=1}^{k}[\underline{z}(i)-H(i)\underline{x}(i)]'R(i)^{-1}[\underline{z}(i)-H(i)\underline{x}(i)]$$

subject to the constraints

$$\underline{x}(i+1) = \Phi(i+1,i)\underline{x}(i); \qquad i = 0,\ldots k-1 \quad .$$

This is accomplished by setting the gradients equal to zero.

3. Maximum Likelihood: This method was introduced by Ho, Schmidt [H2] and others [J2] and consists of maximizing the conditional density $p[x(k),k|z(1),\ldots z(k)]$ or maximizing the likelihood function.

In the linear case the conditional density satisfies Kolmogorov's forward equation [B8].

$$\partial p/\partial t = -p\ tr(F) - p_x'Fx + \tfrac{1}{2}tr(GQG'p_{xx}) \qquad \text{II-9}$$

for the linear dynamics

$$d\underline{x} = F(t)\underline{x}\ dt + G(t)d\underline{v}$$
$$E(d\underline{v}d\underline{v}') = Q(t)\ . \qquad \text{II-10}$$

The solution to (II-9) for the linear system (II-10) with Gaussian noise is

$$\mathbf{p}[\underline{x}(k),k|z(1),\ldots z(k)] = \frac{|HPH'+R|^{\frac{1}{2}}}{(2x3.1416)^{\frac{1}{2}}|R|^{\frac{1}{2}}|P|^{\frac{1}{2}}}\exp[-\tfrac{1}{2}(A)] \qquad \text{II-11}$$

where

$$\mathbf{A} = (\underline{z}-H\hat{\underline{x}})'R^{-1}(\underline{z}-H\hat{\underline{x}}) + (\underline{x}-\hat{\underline{x}})'P^{-1}(\underline{x}-\hat{\underline{x}}) - (\underline{z}-H\hat{\underline{x}})'(HPH'+R)^{-1}(\underline{z}-H\hat{\underline{x}})$$

Therefore the maximum likelihood estimator is equivalent to minimizing

A. See Jazwinski [J2] for detailed derivations.

Variations of the Linear Kalman Filter.

In deriving the linear Kalman filter we assume both the system and measurement have the additive Gaussian white noise property, i.e. it is not correlated. However, this restriction can be relaxed in the following different ways:

1. Correlated system and measurement noise [K1].

$$E[\underline{w}(k)\underline{v}(j)'] = C(k)\delta_{kj} \neq 0 \quad .$$

In this case, only the Kalman filter gain is modified as

$$K(k) = [P(k,k-1)H'(k) + \Gamma(k-1)C(k)][H(k)P(k,k-1)H'(k) + H(k)\Gamma(k-1)C(k) + C'(k)\Gamma'(k-1)H'(k) + R(k)]^{-1} \qquad \text{II-12}$$

and

$$P(k,k) = P(k-1,k) - K(k)[H(k)P(k,k-1) + C'(k)\Gamma'(k-1)] \quad . \qquad \text{II-13}$$

2. Sequentially correlated (colored) measurement noise [J2].

Consider the measurement noise is correlated as $\underline{v}(k+1) = \Psi(k+1,k)\underline{v}(k) + \underline{u}(k)$, assuming $\underline{u}(k)$ is white noise (not correlated). Defining the augmented system

$$\bar{\underline{x}}(k) = [\underline{x}(k), \underline{v}(k)]'$$

and the matrices

$$\bar{\Phi} = \begin{bmatrix} \Phi & 0 \\ 0 & \Psi \end{bmatrix} \qquad \bar{\Gamma} = \begin{bmatrix} \Gamma & 0 \\ 0 & I \end{bmatrix}$$

$$\bar{\underline{w}}'(k) = [\underline{w}(k), \underline{u}(k)]'$$

$$\bar{H}(k) = (H,I) \qquad \bar{Q} = \begin{pmatrix} Q(k) & 0 \\ 0 & S(k) \end{pmatrix} .$$

The augmented dynamic system is

$$\begin{aligned} \bar{x}(k+1) &= \bar{\Phi}(k+1,k)\bar{x}(k) + \bar{\Gamma}(k)\bar{w}(k) \\ z(k) &= \bar{H}(k)\bar{x}(k) \end{aligned} \qquad \text{II-14}$$

3. Multiplicative noise [S1].

This method treats the observation as being contaminated with multiplicative noise

$$z(k) = [1+m(k)]h[x(k),k] + v(k) \qquad \text{II-15}$$

where m(k) is related to

$$\begin{aligned} m(k) &= C(k)y(k) \\ y(k+1) &= D(k)y(k) + G(k)u(k) . \end{aligned} \qquad \text{II-16}$$

The distributions of $x(0)$, $y(0)$, $u(k)$, $w(k)$, $v(k)$ are Gaussian and have zero means and uncorrelated. Here again, we can use the augumented system by defining:

$$\bar{x}(k) = \begin{pmatrix} x(k) \\ y(k) \end{pmatrix} \qquad \bar{w}(k) = \begin{pmatrix} w(k) \\ u(k) \end{pmatrix}$$

$$\bar{\Phi}(k) = \begin{pmatrix} \Phi(k) & 0 \\ 0 & D(k) \end{pmatrix} \qquad \bar{\Gamma}(k) = \begin{pmatrix} \Gamma(k) & 0 \\ 0 & G(k) \end{pmatrix}$$

$$h(k) = [1+C(k)y(k)]h[x(k),k] . \qquad \text{II-17}$$

See Sage [S1] for detailed derivations and extensions to non-linear filtering via linearization.

II.3. Nonlinear Discrete Filteirng

Let the nonlinear discrete model be

$$\underline{x}(k+1) = \underline{f}[x(k)] + \Gamma[\underline{x}(k)]\underline{w}(k) \qquad \text{II-18}$$

$$\underline{x}(0) = \underline{c}$$

and the nonlinear measurement be

$$\underline{z}(k+1) = \underline{h}[x(k+1)] + \underline{v}(k+1) \qquad \text{II-19}$$

where v(k) and w(k) are uncorrelated Gaussian sequences such that

$$E[\underline{w}(k)\underline{w}'(j)] = Q(k)\delta_{kj}$$

$$E[\underline{v}(k)\underline{v}'(j)] = R(k)\delta_{kj} \quad .$$

Then the filter estimate conditional probability density function contains all the information concerning the processes.

Let this conditional density be denoted by $P_n(y)$ then, the following iterative formula determines $P_n(y)$

$$C_n P_{n+1}(y) = \int \exp[z(k),h(a)]_{R^{-1}} - \tfrac{1}{2}||h(a)||^2_{R^{-1}} P_n(a)$$

$$\times \frac{[\det\Gamma(a)Q(n-1)\Gamma(a)]^{-\frac{1}{2}}}{(2\times 3.1416)^{m/2}} \exp[-\tfrac{1}{2}||y-f(a)||^2_{A^{-1}}]da \qquad \text{II-20}$$

where

$$C_n = \int \exp[\left(z(k),h(a)\right)_{R^{-1}} - \tfrac{1}{2}||h(a)||^2_{R^{-1}}]P_n da$$

with

$$P_o(a) = \frac{(\det K)^{-\frac{1}{2}}}{(2\pi)^{m/2}} \exp(-\tfrac{1}{2}||a||^2_{K^{-1}}) \quad .$$

Here the state has m dimensions. Unfortunately, this approach involves

massive computation. See Bucy [B8] for a detailed derivation. Starting from the conditional probability density function, various approximations are made by different authors in order to achieve a computationally feasible filter algorithm. Among these, Bucy and Bass [B8] use a Taylor series expansion neglecting orders higher than second and thus arrive at the so-called second order approximate non-linear filter. Jazwinski [J2] applied this approach to continuous processes and with discrete measurement, he arrived at a continuous-discrete filter. Lo [L4] made a recent modification and presented the idea of continuous-discrete smoothing.

An alternative to this approach is skipping the conditional probability density for the nonlinear model. The nonlinear model is linearized and we can use the existing Kalman filter for the linearized model. This approach is the so-called non-linear extended Kalman filter.

Variations of these two types of nonlinear filters can be found in Jazwinski and Schwartz's works [J2]. Sorenson [S11] made comparisons of four different non-linear filters on their performance on probability density functions. He concluded that using the extended Kalman filter assuming the conditional density function was approximated by Gaussian density showed inadequacy in a nonlinear, nongaussian system. The second order nonlinear filter proved to be better than the Kalman type for a nonlinear nongaussian model. In the same paper, Sorenson introduced his Gaussian sum filter by approximating the conditional density as the sum of weighted gaussian densities. He concluded that in some instances, his gaussian sum approximation is

superior to that of the second order method. However, one drawback is that his new nonlinear filter requires significantly greater computational load than both the extended Kalman filter and second order filter.

Extended Kalman Filter

The extended Kalman filter is essentially the same as the linear Kalman filter, except the transition matrix and measurement matrix are replaced by the linearized model of the original nonlinear process. Thus, linear filtering theory can be extended to the treatment of nonlinear approximate filtering.

Let the nonlinear system be given as

$$\underline{x}(k+1) = \underline{f}[\underline{x}(k),\underline{u}(k)] + \underline{g}(k)\underline{w}(k) \qquad \text{II-21}$$

$$\underline{z}(k) = \underline{h}[\underline{x}(k)] + \underline{v}(k) \qquad \text{II-22}$$

where $\underline{f}$ and $\underline{h}$ are nonlinear process and measurement vector functions and $\underline{w}(k)$, $\underline{v}(k)$ are both zero mean gaussian white noise.

The extended Kalman filter is evaluated by replacing

$$\Phi_{k-1}(\underline{f}_{k-1})_{\hat{\underline{x}}(k-1|k-1)} = \left.\frac{\partial \underline{f}[\underline{x}(k-1),\underline{u}(k-1)]}{\partial \underline{x}(k-1)}\right|_{\hat{\underline{x}}(k-1|k-1)} \qquad \text{II-23}$$

$$H(k) = \left.\frac{\partial \underline{h}[\underline{x}(k)]}{\partial \underline{x}(k)}\right|_{\hat{\underline{x}}(k|k)} . \qquad \text{II-24}$$

Therefore in the prediction stage, the filter estimate is

$$\hat{\underline{x}}(k|\,k-1) = \underline{f}[\hat{\underline{x}}(k-1|\,k-1),\underline{u}(k)] \qquad \text{II-25}$$

and the observation is

$$\underline{\hat{z}}(k|\,k-1) = \underline{h}[\hat{\underline{x}}(k|\,k-1)] \qquad \text{II-26}$$

In order to make the extended Kalman filter applicable to a nonlinear process, P(0), Q(k), R(k) cannot be too large to violate linearity assumptions. Note that the Gaussian and Markov properties no longer hold globally for non-linear filtering and the initial prior statistics are very important and make the filter performance more sensitive.

Nonlinear Second Order Approximate Optimal Filtering

Schwartz and Bass [B2] approximated the optimal non-linear filter using a Taylor Series expansion with terms carried out to second order.

Let the nonlinear model be represented by

$$d\underline{x} = f(\underline{x})dt + \Gamma(\underline{x})dv \qquad \text{II-27}$$

$$d\underline{z} = h(\underline{x}) + dw \qquad \text{II-28}$$

where $\underline{x}$ is the state and $\underline{z}$ is the observation.

Using Taylor expansions of $f(\underline{x})$ and $h(\underline{x})$ around the estimate $\hat{\underline{x}}$

$$\underline{f}(\underline{x}) = \underline{f}(\hat{\underline{x}}) + f_x(\hat{\underline{x}})(\underline{x}-\hat{\underline{x}}) + \tfrac{1}{2}f_{xx}(\hat{\underline{x}}) : (x-\hat{\underline{x}})(x-\hat{\underline{x}})' \qquad \text{II-29}$$

$$\underline{h}(\underline{x}) = \underline{h}(\hat{\underline{x}}) + h_x(\hat{\underline{x}})(\underline{x}-\hat{\underline{x}}) + \tfrac{1}{2}h_{xx}(\hat{\underline{x}}) : (x-\hat{\underline{x}})(x-\hat{\underline{x}})' \qquad \text{II-30}$$

where

$$f_{xx}(x) : S = \begin{vmatrix} \text{trace } [f_{1xx}(x)S] \\ \vdots \\ \text{trace } [f_{nxx}(x)S] \end{vmatrix} .$$

Taking the expected value of Equations II-29 and II-30, an approximate non-linear filter is: [since $E(x) = \hat{x}$]

$$d\hat{x} = f(\hat{x}) + \tfrac{1}{2}f_{xx}(\hat{x}) : P\,dt + P\,h_x'(\hat{x})R^{-1}(t) \qquad \text{II-31}$$
$$[dz - h(\hat{x})dt - \tfrac{1}{2}h_{xx}(\hat{x}) : P\,dt] \quad .$$

The approximate equation for the error covariance matrix is:

$$dP = Pf_x'(\hat{x})dt + f_x(\hat{x})Pdt - P[h_x'(\hat{x})R^{-1}(t)h_x(\hat{x})]Pdt +$$

$$\Gamma(\hat{x})Q(t)\Gamma'(\hat{x})dt + \tfrac{1}{2}[\Gamma(\hat{x})Q(t)\Gamma'(\hat{x})]_{xx} : Pdt -$$

$$\tfrac{1}{2}(P : h_{xx}'(\hat{x})R^{-}(t)[dz - h(\hat{x})dt - \tfrac{1}{2}h_{xx}(\hat{x}) : P)$$

with

$$P(0) = E\left([x(0) - \hat{x}(0)][x(0) - \hat{x}(0)]'\right) . \qquad \text{II-32}$$

Where $M_{xx} : P$ denote a matrix and is a dyadic product

$$(M_{xx} : P)_{ij} = \text{trace } (M_{ijxx}P) .$$

Q(t) and R(t) are the process and measurement noise covariances respectively.

The above approximate non-linear filter shows marked differences to the extended Kalman filter which is obtained by linearization of the process model and then applying linear filtering theory. This approximate filter also represents a major simplification over the direct calculation of the conditional distribution.

The Taylor series approximation techniques represents only one approach to the practical synthesis of a filtering scheme for non-linear systems. There still needs to be a great deal of effort, both in the theoretical and practical aspects of non-linear filtering. This work tries to fill a gap in this area by extensive evaluation and comparison of extended Kalman filters utilizing a chemical reactor example suggested by Wells [W1].

We observe that all of the terms in Equations II-31 and II-32 including Pf_{xx} are used to correct the process nonlinearity while P^2h_{xx} corrects the measurement nonlinearity.

In our numerical example, we have a linear measurement system, therefore all of the terms with P^2h_{xx} drop out.

II.4. Numerial Examples and Discussions.

Non-linear filtering of a Stirred Tank Reactor system [W1]. Given a chemical reactor, the describing equation are:

a. Energy balance: (reactor) (Reaction: $2A \xrightarrow{k} B$)

$$\rho V C_p \frac{dT}{dt} = \rho F C_p (T_i - T) + \Delta H K V \exp(-E/RT) C_A^2 - hA(T - T_w) \qquad \text{II-23}$$

b. Energy balance on reactor wall

$$\rho_w V_w C_{pw} \frac{dT_w}{dt} = hA(T - T_w) + h_w A_w (T_w - T_c) \qquad \text{II-34}$$

c. Energy balance on jacket side

$$\rho_c V_c C_{pc} \frac{dT_c}{dt} = \rho_c F_c C_{pc} (T_{c_i} - T_c) + h_w A_w (T_w - T_c) \qquad \text{II-35}$$

d. Mass balance on reactor contents

$$V \frac{dc_A}{dt} = F(C_{Ai} - C_A) - KV \exp(-E/RT) C_A^2 \ . \qquad \text{II-36}$$

Following the same normalization procedures as outlined by Wells' and requiring the reactor to be maintained at steady state, the state space estimation equations are

$$\dot{x}_1 = -(c_1 + c_4)x_1 + c_3(1 + x_4)^2 \exp[K_1 x_1/(1 + x_1)] + c_4 x_2 - c_3 \qquad \text{II-37}$$

$$\dot{x}_2 = -(c_5 + c_6)x_2 + c_5 x_1 + c_6 x_3 \qquad \text{II-38}$$

$$\dot{x}_3 = -(c_7+c_8)x_3 + c_8x_2 \qquad \text{II-39}$$

$$\dot{x}_4 = -c_1x_4 - c_2(1+x_4)^2\exp[K_1x_1/(1+x_1)] + c_2 \qquad \text{II-40}$$

while the measurement equations are

$$\underline{y}_i(k) = \underline{x}_i(k) + \underline{v}_i(k); \quad i = 1,2,3 . \qquad \text{II-41}$$

The linear measurement matrix $M = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{vmatrix}$.

$$\bar{y} = M\bar{x} + \bar{v} \quad .$$

The transition matrix for the linearized system is related to the Jacobian of the above system equations

$$\Phi = \begin{vmatrix} f_{11} & f_{12} & 0 & f_{14} \\ f_{21} & f_{22} & f_{23} & 0 \\ 0 & f_{32} & f_{33} & 0 \\ f_{41} & 0 & 0 & f_{44} \end{vmatrix}$$

with process parameters given below

$$c_1 = c_2 = 0.2; \qquad c_3 = 1.0$$

$$c_4 = c_5 = c_6 = c_8 = 0.5; \qquad c_7 = 0.05 \quad .$$

Wells did not point out his K_1 value where $K_1 = E/RT_s$. For the response given by his Fig. 2 it was computed by this simulation that K_1 is approximately 1.0.

The components of the Jacobian matrix are:

$$f_{11} = (c_1+c_4) + c_3B^2K_1DE$$

$$f_{12} = c_4; \quad f_{14} = 2c_3BE$$

$$f_{21} = c_5$$

$$f_{22} = -(c_5+c_6); \quad f_{23} = c_6$$

$$f_{32} = c_8$$

$$f_{33} = -(c_7+c_8)$$

$$f_{41} = -c_2B^2K_1DE$$

$$f_{44} = -c_1 - 2c_2BE \quad .$$

Where

$$B = (1+x_4)$$

$$E = \exp[K_1x_1/(1+x_1)]$$

$$D = (1+x_1)^{-2} \quad .$$

For the second order filter we need to evaluate the dyadic product of the second order partial derivative and error covariance matrix. The second order partial derivatives are defined as:

$$f_{1xx} = \frac{\partial^2 f_1}{\partial x_i \partial x_j} \qquad i = 1,4 \quad . \quad j = 1,4$$

$$f_{1xx} = \begin{vmatrix} f_{1x_1x_1} & 0 & 0 & f_{1x_1x_4} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ f_{1x_4x_1} & 0 & 0 & f_{1x_4x_4} \end{vmatrix} \quad . \qquad \text{II-42}$$

Where

$$f_{1x_1x_1} = c_3B^2D^2K_1^2E - 2c_3B^2DK_1E/(1+x_1)$$

$$f_{1x_4x_1} = f_{1x_1x_4} = 2c_3BDEK_1$$

$$f_{1x_4x_4} = 2c_3E$$

$$f_{2xx} = \frac{\partial^2 f_2}{\partial x_i \partial x_j} \quad , \qquad \text{II-43}$$

$$f_{3xx} = \frac{\partial^2 f_3}{\partial x_i \partial x_j} \quad . \qquad \text{II-44}$$

The matrices f_{2xx} and f_{3xx} are identically zero.

$$f_{4xx} = \frac{\partial^2 f_4}{\partial x_i \partial x_j} \qquad \text{II-45}$$

then

$$f_{4xx} = \begin{vmatrix} f_{4x_1x_1} & 0 & 0 & f_{4x_1x_4} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ f_{4x_4x_1} & 0 & 0 & f_{4x_4x_4} \end{vmatrix} ,$$

where

$$f_{4x_1x_1} = - c_2B^2D^2K_1E + 2c_2B^2DEK_1/(1+x\ )$$

$$f_{4x_4x_1} = f_{4x_1x_4} = - 2c_2BDEK_1$$

$$f_{4x_4x_4} = -2c_2E \ .$$

Substituting the above relations into the extended Kalman filter and the second order filter with the x's evaluated at the estimate at every sample period, we are able to carry out the sequential filtering results as shown on Fig. 2.1, through Fig. 2.4 for the second order filtering and Fig. 2.5 to Fig. 2.8 for the extended Kalman filter. We adopt Wells' Nomenclature for the constants c(i), i = 1,8 and the physical parameters. These values are outlined in Table II-2.

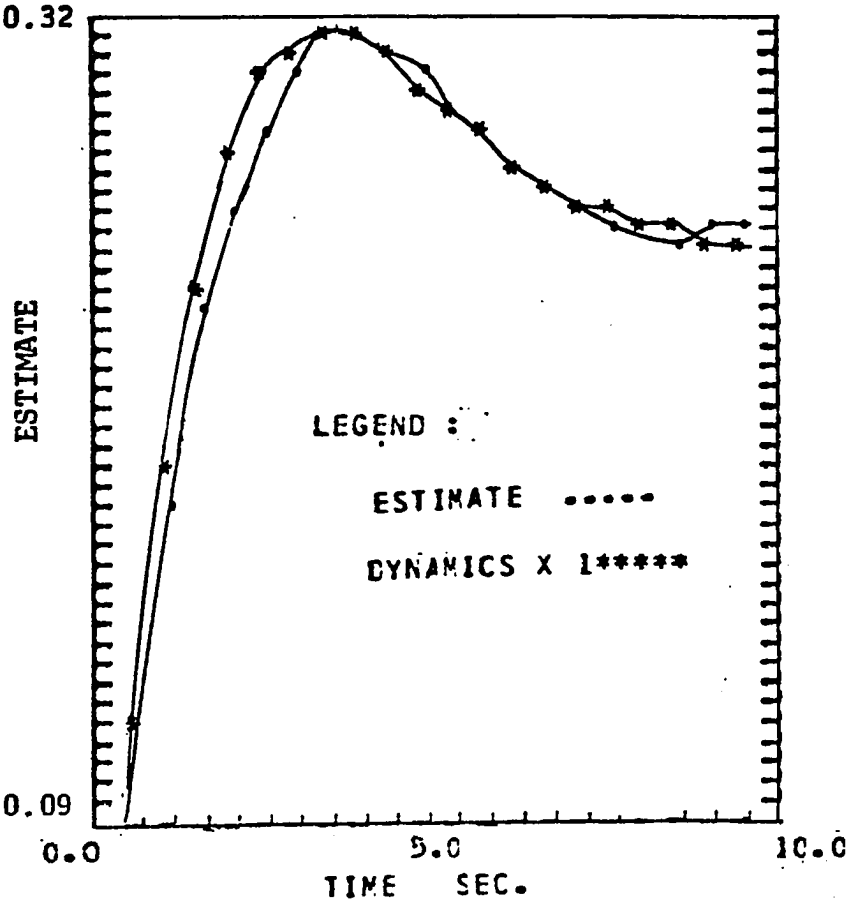


FIG. 2.1 EFFECT OF DIFFERENT P(0)
SECOND ORDER NONLINEAR FILTER
CONSTANT NOISE P11(0) = 0.005

TIME SEC.	X 1	ESTIMATE
0.5	0.0988	0.0741
1.0	0.1810	0.1704
1.5	0.2427	0.2324
2.0	0.2845	0.2653
2.5	0.3097	0.2946
3.0	0.3222	0.3121
3.5	0.3255	0.3236
4.0	0.3228	0.3242
4.5	0.3165	0.3220
5.0	0.3083	0.3112
5.5	0.2995	0.2990
6.0	0.2907	0.2904
6.5	0.2827	0.2795
7.0	0.2755	0.2733
7.5	0.2695	0.2701
8.0	0.2645	0.2627
8.5	0.2607	0.2619
9.0	0.2578	0.2564
9.5	0.2559	0.2608
10.0	0.2547	0.2592

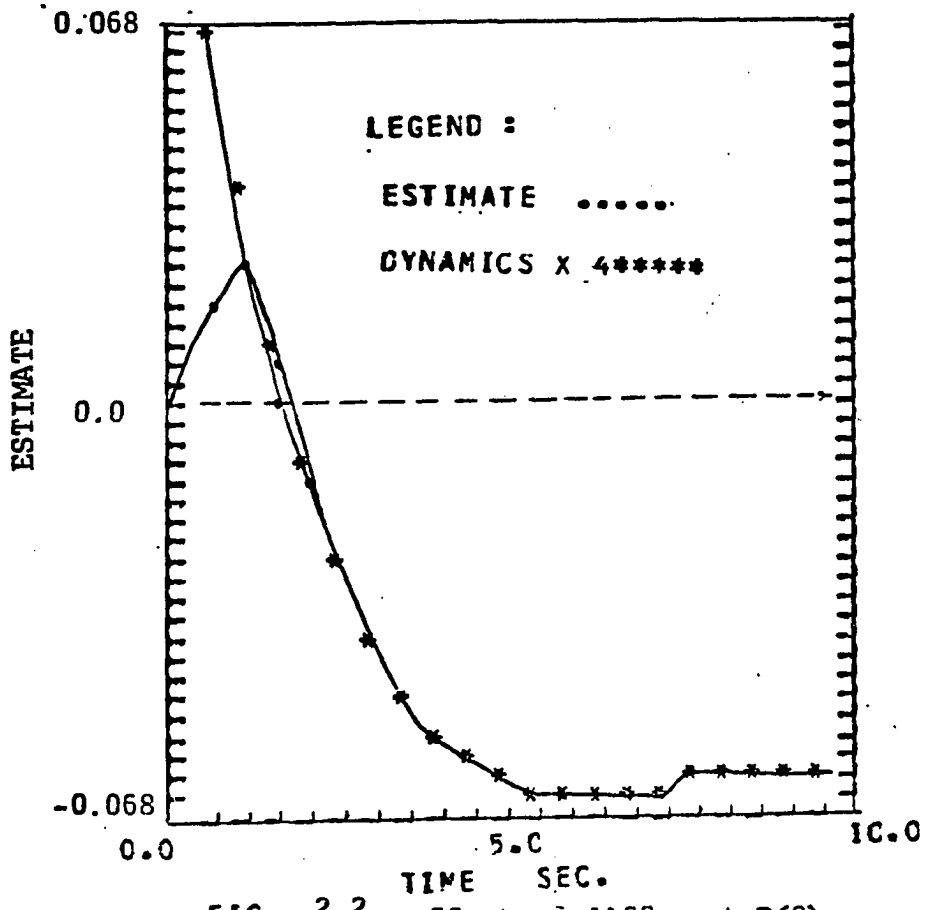


FIG. 2.2 effect of different P(0)
SECOND ORDER NONLINEAR FILTER
CONSTANT NOISE P11(0) = 0.01

TIME SEC.	X 4	ESTIMATE
0.5	0.0683	0.0192
1.0	0.0377	0.0232
1.5	0.0107	0.0070
2.0	-0.0120	-0.0138
2.5	-0.0301	-0.0297
3.0	-0.0440	-0.0428
3.5	-0.0543	-0.0531
4.0	-0.0616	-0.0612
4.5	-0.0666	-0.0666
5.0	-0.0698	-0.0703
5.5	-0.0716	-0.0725
6.0	-0.0725	-0.0729
6.5	-0.0727	-0.0730
7.0	-0.0724	-0.0722
7.5	-0.0718	-0.0714
8.0	-0.0711	-0.0708
8.5	-0.0703	-0.0701
9.0	-0.0696	-0.0697
9.5	-0.0689	-0.0689
10.0	-0.0684	-0.0689

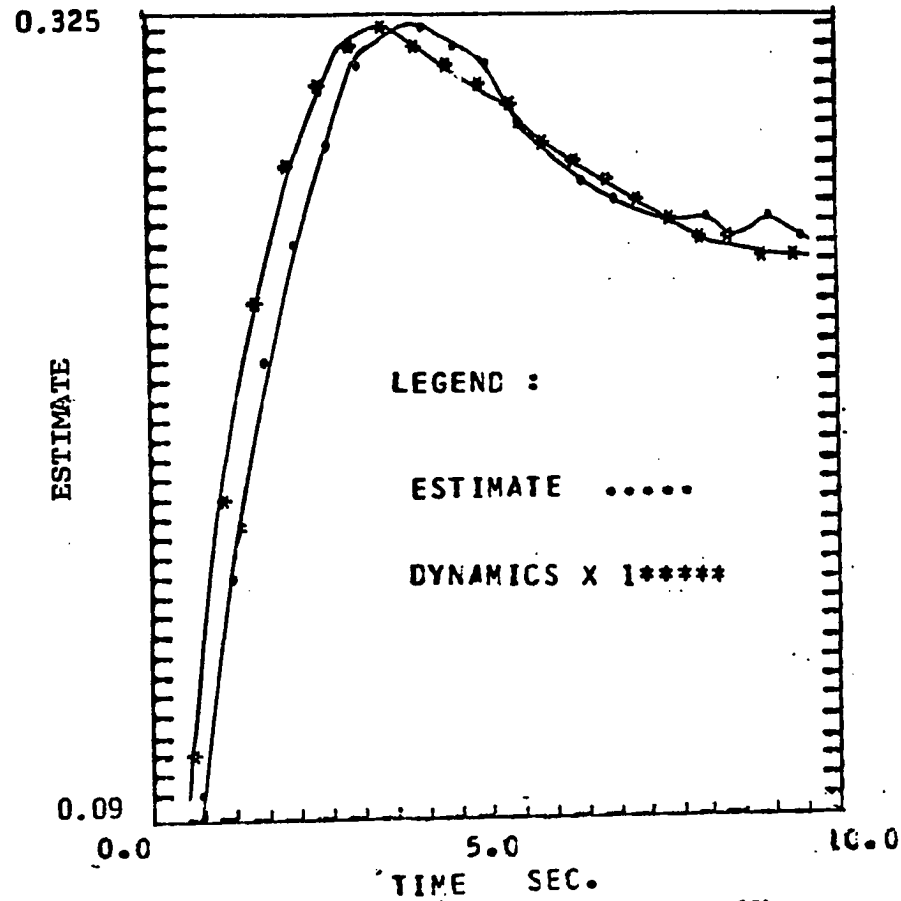


FIG. 2.3 effect of different P(0)
SECOND ORDER NONLINEAR FILTER
CONSTANT NOISE P11(0) = 0.01

TIME SEC.	X 1	ESTIMATE
0.5	0.0988	0.0877
1.0	0.1810	0.1546
1.5	0.2427	0.2238
2.0	0.2845	0.2588
2.5	0.3097	0.2919
3.0	0.3222	0.3135
3.5	0.3255	0.3265
4.0	0.3228	0.3257
4.5	0.3165	0.3229
5.0	0.3083	0.3127
5.5	0.2995	0.2960
6.0	0.2907	0.2893
6.5	0.2827	0.2758
7.0	0.2755	0.2717
7.5	0.2695	0.2697
8.0	0.2645	0.2648
8.5	0.2607	0.2628
9.0	0.2578	0.2587
9.5	0.2559	0.2638
10.0	0.2547	0.2620

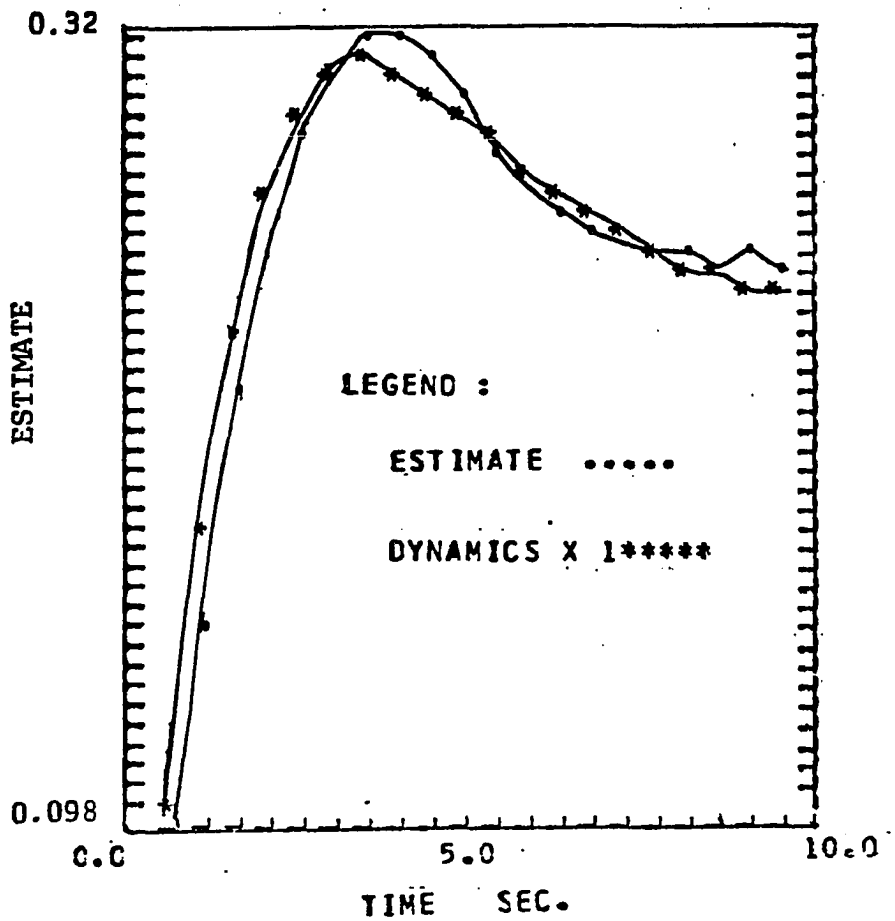


FIG. 2.4 EFFECT OF DIFFERENT P(0)
SECOND ORDER NONLINEAR FILTER
CONSTANT NOISE P11(0) = 0.015

TIME SEC.	X 1	ESTIMATE
0.5	0.0988	0.1010
1.0	0.1810	0.1541
1.5	0.2427	0.2250
2.0	0.2845	0.2650
2.5	0.3097	0.3008
3.0	0.3222	0.3217
3.5	0.3255	0.3321
4.0	0.3228	0.3299
4.5	0.3165	0.3260
5.0	0.3083	0.3147
5.5	0.2995	0.2970
6.0	0.2907	0.2899
6.5	0.2827	0.2761
7.0	0.2755	0.2717
7.5	0.2695	0.2695
8.0	0.2645	0.2646
8.5	0.2607	0.2635
9.0	0.2578	0.2584
9.5	0.2559	0.2635
10.0	0.2547	0.2618

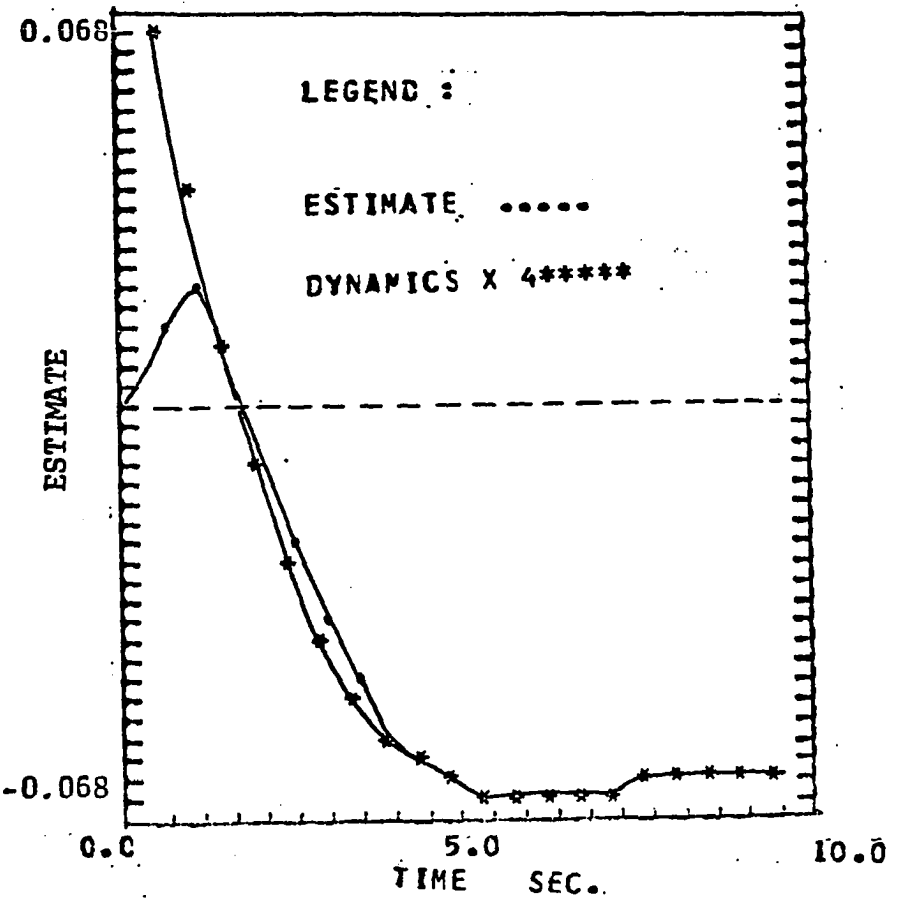


FIG. 2.5 EFFECT OF DIFFERENT P(0)
SECOND ORDER NONLINEAR FILTER
CONSTANT NOISE P11(0) = 0.015

TIME SEC.	X 4	ESTIMATE
0.5	0.0683	0.0147
1.0	0.0377	0.0228
1.5	0.0107	0.0108
2.0	-0.0120	-0.0104
2.5	-0.0301	-0.0276
3.0	-0.0440	-0.0420
3.5	-0.0543	-0.0531
4.0	-0.0616	-0.0615
4.5	-0.0666	-0.0671
5.0	-0.0698	-0.0708
5.5	-0.0716	-0.0728
6.0	-0.0725	-0.0732
6.5	-0.0727	-0.0732
7.0	-0.0724	-0.0724
7.5	-0.0718	-0.0715
8.0	-0.0711	-0.0709
8.5	-0.0703	-0.0701
9.0	-0.0696	-0.0697
9.5	-0.0689	-0.0689
10.0	-0.0684	-0.0688

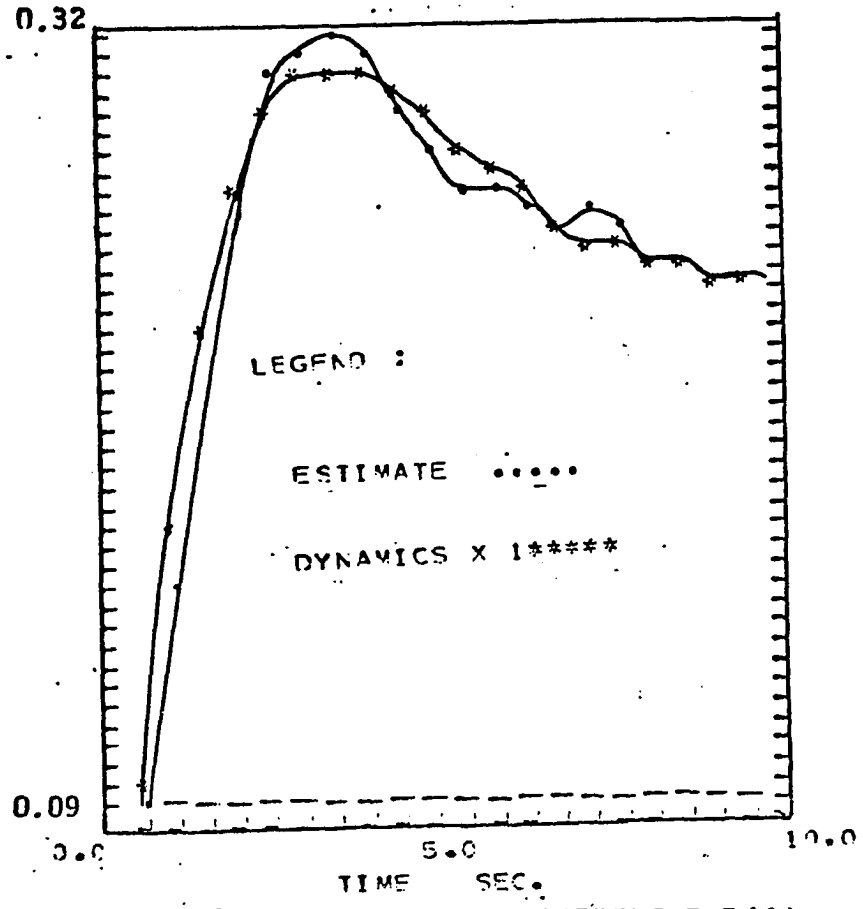


FIG. 2.6 EFFECT OF DIFFERENT P(0)
NONLINEAR EXTENDED KALMAN FILTER
CONSTANT NOISE P11(0) = 0.015

TIME SEC.	X 1	ESTIMATE
0.5	0.0988	0.0937
1.0	0.1810	0.1603
1.5	0.2427	0.2430
2.0	0.2845	0.2882
2.5	0.3097	0.3206
3.0	0.3222	0.3293
3.5	0.3255	0.3349
4.0	0.3228	0.3265
4.5	0.3165	0.3131
5.0	0.3083	0.2983
5.5	0.2995	0.2884
6.0	0.2907	0.2859
6.5	0.2827	0.2769
7.0	0.2755	0.2755
7.5	0.2695	0.2779
8.0	0.2645	0.2721
8.5	0.2607	0.2635
9.0	0.2578	0.2610
9.5	0.2559	0.2566
10.0	0.2547	0.2562

ACCUMULATED MEAN SQUARE ERROR = 0.5878E-04

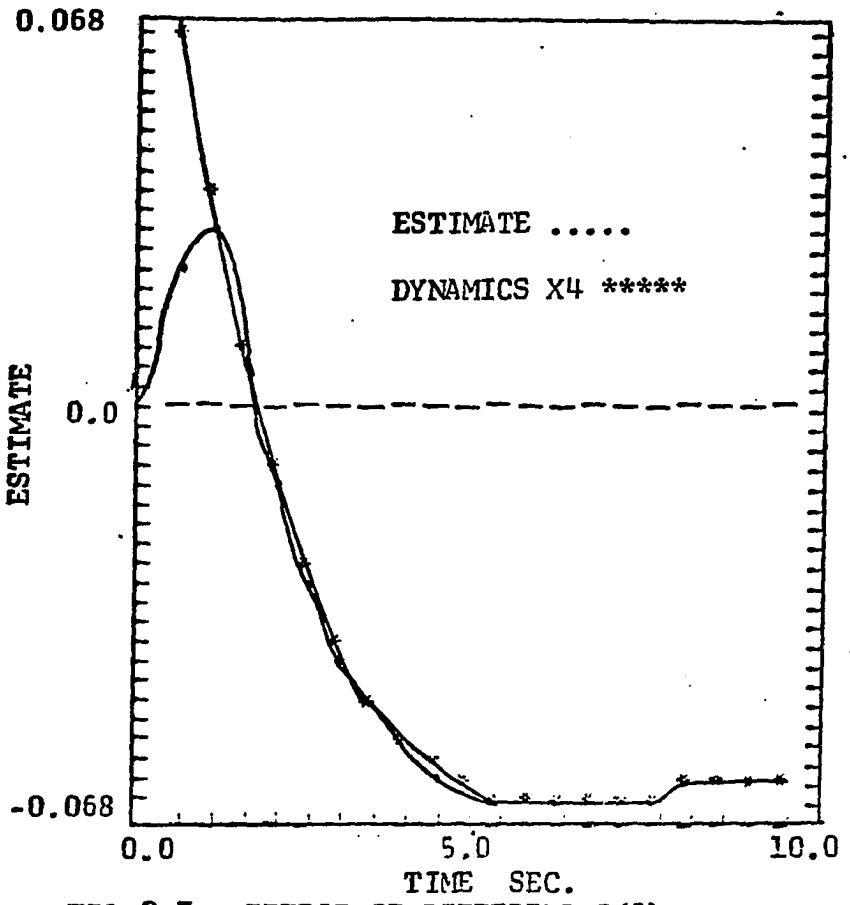


FIG. 2.7 EFFECT OF DIFFERENT P(0)
NONLINEAR EXTENDED KALMAN FILTER WITH
CONSTANT NOISE P11(0) = 0.005

TIME SEC.	X 4	ESTIMATE
0.5	0.0683	0.0231
1.0	0.0377	0.0306
1.5	0.0107	0.0059
2.0	-0.0120	-0.0164
2.5	-0.0301	-0.0331
3.0	-0.0440	-0.0463
3.5	-0.0543	-0.0558
4.0	-0.0616	-0.0629
4.5	-0.0665	-0.0674
5.0	-0.0698	-0.0700
5.5	-0.0716	-0.0711
6.0	-0.0725	-0.0717
6.5	-0.0727	-0.0719
7.0	-0.0724	-0.0715
7.5	-0.0718	-0.0713
8.0	-0.0711	-0.0713
8.5	-0.0703	-0.0708
9.0	-0.0696	-0.0698
9.5	-0.0689	-0.0692
10.0	-0.0684	-0.0682

ACCUMULATED MEAN SQUARE ERROR = 0.1077E-03

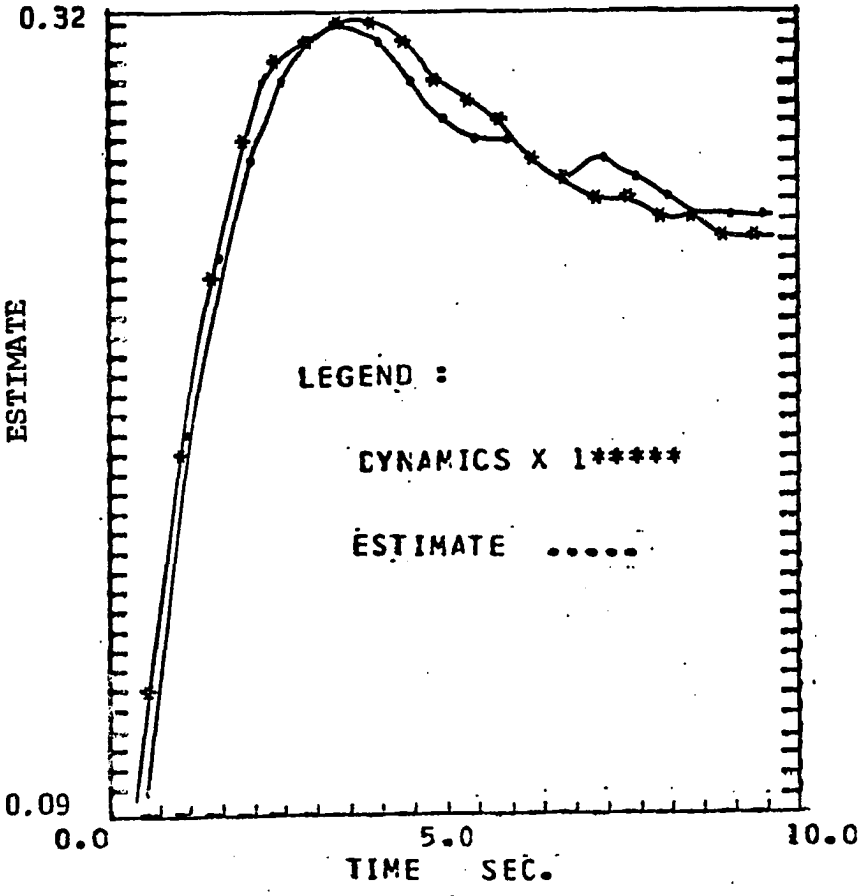


FIG. 2.8 EFFECT OF DIFFERENT P(0)
NONLINEAR EXTENDED KALMAN FILTER
CONSTANT NOISE P11(0) = 0.005

TIME SEC.	X 1	ESTIMATE
0.5	0.0988	0.0689
1.0	0.1810	0.1852
1.5	0.2427	0.2480
2.0	0.2845	0.2807
2.5	0.3097	0.3078
3.0	0.3222	0.3166
3.5	0.3255	0.3231
4.0	0.3228	0.3171
4.5	0.3165	0.3066
5.0	0.3083	0.2940
5.5	0.2995	0.2856
6.0	0.2907	0.2844
6.5	0.2827	0.2763
7.0	0.2755	0.2758
7.5	0.2695	0.2787
8.0	0.2645	0.2731
8.5	0.2607	0.2644
9.0	0.2578	0.2618
9.5	0.2559	0.2572
10.0	0.2547	0.2566

Table II-2. Nomenclature.

Symbol	Definition	Units
ρ	average density of reactor contents	lb/ft^3
ρ_w	average density of reactor wall	lb/ft^3
ρ_c	average density of coolant	lb/ft^3
V	reactor volume	ft^3
V_w	reactor wall volume	ft^3
V_c	cooler volume	ft^3
T_i	feed temperature	°R
T	reactor content temperature	°R
T_w	reactor wall temperature	°R
T_c	cooler temperature	°R
C_p	average heat capacity of reactor contents	BTU/lb-°R
C_{pw}	average heat capacity of wall	BTU/lb-°R
C_{pc}	average heat capacity of coolant	BTU/lb-°R
F	volumetric flow	ft^3/sec
F_c	volumetric coolant flow rate	ft^3/sec
K	preexponential rate constant frequency factor	1/sec
H	heat of reaction	BTU/lb
h	film coefficient of reactor content	BTU/ft^2-°F-sec
h_w	film coefficient between wall and coolant	BTU/ft^2-°F-sec
C_A	weight concentration of reactant A	lb/ft^3
C_{A_i}	feed concentration of A	lb/ft^3
E	activation energy for reaction	$atm-ft^3/mole$
R	gas constant	$atm-ft^3$/°R-mol
t		time sec.

Table II-2 Continued

A	reactor wall area	ft^2
A_w	wall cooler surface area	ft^2

Dimensionless Constants

β = time scale factor = 1/240

$K_1 = E/RT_s$

$c_1 = F/V$

$c_2 = (K/\beta)C_s^2 \exp(-E/RT_s)$

$c_3 = \Delta H K \exp(-K_1) C_s^2/(\rho C_p T_s)$

$c_4 = hA/(\rho V C_p \beta)$

$c_5 = hA/(\rho_w V_w C_{pw} \beta)$

$c_6 = (h_w A_w)/(C_{pw} \rho_w V_w \beta)$

$c_7 = F_c/(V_c \beta)$

$c_8 = h_w A_w/(\rho_c V_c C_{pc} \beta)$

C_s steady state concentration

T_s steady state temperature

$x_1 = \frac{T-T_s}{T_s}$ = Normalized reactor temperature.

$x_2 = \frac{T_w-T_s}{T_s}$ = Normalized reactor wall temperature.

$x_3 = \frac{T_c-T_s}{T_s}$ = Normalized coolant temperature.

$x_4 = \frac{C_A-C_s}{C_s}$ = Normalized reactant concentration.

We assume the following prior statistics:

$$P(0) = \begin{vmatrix} 0.01 & 0 & 0 & 0 \\ 0 & 0.01 & 0 & 0 \\ 0 & 0 & 0.01 & 0 \\ 0 & 0 & 0 & 0.01 \end{vmatrix}$$

$$Q(k) = \text{constant} = \begin{vmatrix} 0.0001 & 0 & 0 & 0 \\ 0 & 0.0001 & 0 & 0 \\ 0 & 0 & 0.0001 & 0 \\ 0 & 0 & 0 & 0.0001 \end{vmatrix}$$

$$R(k) = \text{constant} = \begin{vmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.01 \end{vmatrix}.$$

with initial process conditions $x_1 = x_2 = x_3 = 0$; $x_4 = 0.1$ and the initial estimate $\hat{x}_1 = \hat{x}_2 = \hat{x}_2 = \hat{x}_4 = 0$.

As stated by Wells the estimate $\hat{x}_1$, $\hat{x}_2$, $\hat{x}_3$ approach the true states very well. However, x_4 requires several observations and is caused by the fact that the actual error between the estimates is zero for x_1, x_2, x_3 and 0.1 for x_4. The overshoot at the peaks of each of the states is due to integration error. Euler's integration rule was used.

Figure 2.1 shows the filtering on reduced temperature by second order non-linear filter with lower initial estimation error covariance P11(0) leads to a low estimate for the temperature, however, the transient due to the initial uncertainty of P11(0) disappears after the fourth sampling (Gaussian Markov property) where the filter approaches the steady state. Figure 2.2 and Fig. 2.5 shows its effect on the estimate of concentration. Figure 2.4 shows, the delayed approach to steady state due to a higher value of P11(0), even though it predicts very well in the early part of the filtering. It has been found in

this simulation that with P11(0) greater than 0.02, the filter will diverge after the third period. This shows a very sensitive nature of the reactor model to different P(0). Similar results for the extended Kalman filter is shown in Fig. 2.6 to Fig. 2.8.

The degrading of the reactor temperature estimate due to raising the error variance of its measurement is shown in Fig. 2.9 for the second order filter and Fig. 2.12 for the extended Kalman filter. It is shown that the non-linear second order filter did a better job than the extended Kalman filter. Figure 2.10 shows that increasing the error variance of the reactant temperature has little effect on the estimate of reactant concentration.

Discussion:

The reason that x_4 requires more observations is not only due to the existence of initial transients, but also to the fact that $\hat{x}_4$ is not a measured variable. Here we are facing the same situation as in the linear filtering example. Therefore, it requires at least four observations to completely define the 4 state variables, and the estimate of x_4 is not expected to be good until 4th observation. The solution to the oscillations and overshooting of this particular example is to introduce the second order approximate filter.

One approach used to reduce the overshooting due to integration errors is to introduce a fictitious noise to compensate for the integration errors. See Chapter IV for a detailed discussion and numerical examples.

Extensive numerical studies were performed on Wells' reactor model using both the non-linear extended Kalman filter and the second

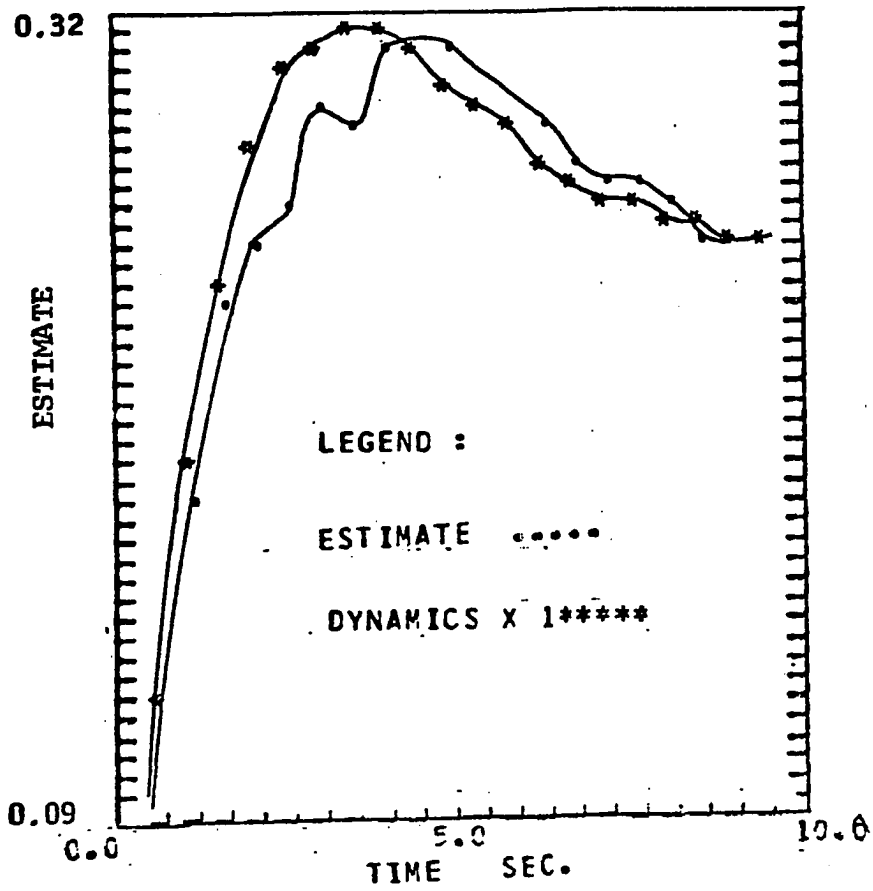


FIG. 2.9 EFFECT OF DIFFERENT R(K)
SECOND ORDER NONLINEAR FILTER
CONSTANT NOISE R11(K) = 0.03

TIME SEC.	X 1	ESTIMATE
0.5	0.0988	0.0686
1.0	0.1810	0.1696
1.5	0.2427	0.2311
2.0	0.2845	0.2500
2.5	0.3097	0.2686
3.0	0.3222	0.2976
3.5	0.3255	0.2930
4.0	0.3228	0.3169
4.5	0.3165	0.3176
5.0	0.3083	0.3167
5.5	0.2995	0.3023
6.0	0.2907	0.2954
6.5	0.2827	0.2955
7.0	0.2755	0.2801
7.5	0.2695	0.2733
8.0	0.2645	0.2760
8.5	0.2607	0.2649
9.0	0.2578	0.2551
9.5	0.2559	0.2555
10.0	0.2547	0.2546

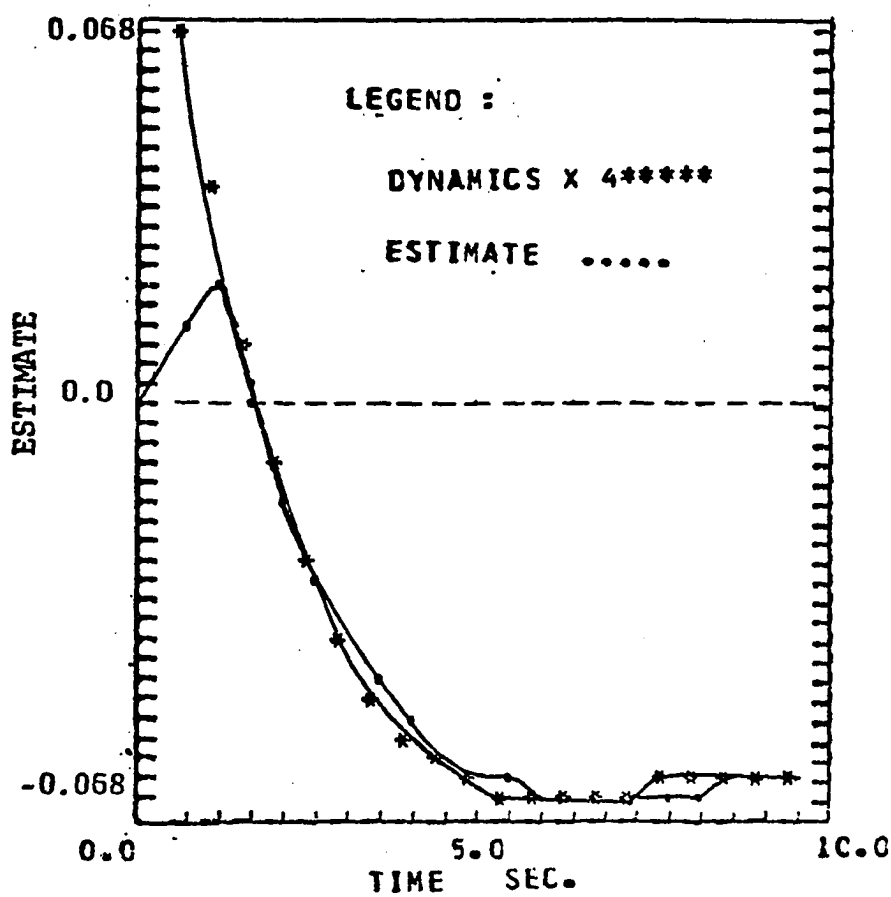


FIG. 2.10 EFFECT OF DIFFERENT R(K)
SECOND ORDER NONLINEAR FILTER
CONSTANT NOISE R(K) = 0.03

TIME SEC.	X 4	ESTIMATE
0.5	0.0683	0.0152
1.0	0.0377	0.0205
1.5	0.0107	0.0024
2.0	-0.0120	-0.0182
2.5	-0.0301	-0.0325
3.0	-0.0440	-0.0435
3.5	-0.0543	-0.0527
4.0	-0.0616	-0.0593
4.5	-0.0666	-0.0650
5.0	-0.0698	-0.0690
5.5	-0.0716	-0.0716
6.0	-0.0725	-0.0727
6.5	-0.0727	-0.0733
7.0	-0.0724	-0.0734
7.5	-0.0718	-0.0728
8.0	-0.0711	-0.0721
8.5	-0.0703	-0.0716
9.0	-0.0696	-0.0707
9.5	-0.0689	-0.0696
10.0	-0.0684	-0.0688

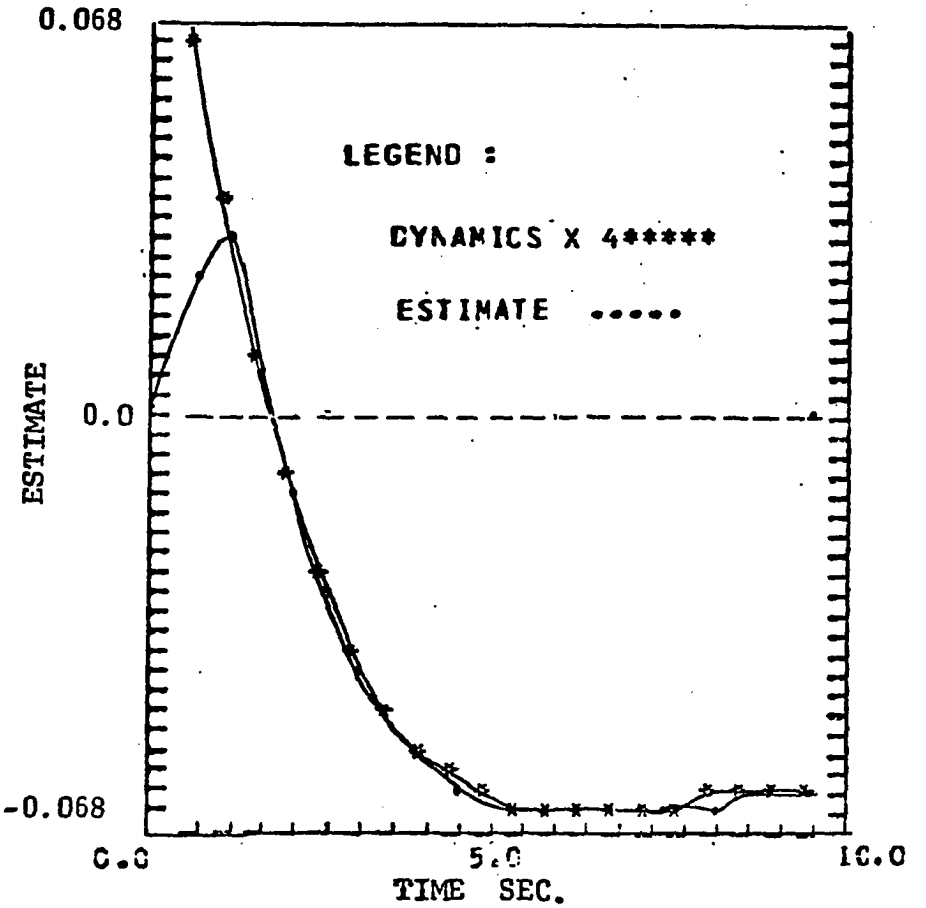


FIG. 2.11 EFFECT OF DIFFERENT P(0)
NONLINEAR EXTENDED KALMAN FILTER
CONSTANT NOISE P11(0) = 0.005

TIME SEC.	X 4	ESTIMATE
0.5	0.0683	0.0248
1.0	0.0377	0.0306
1.5	0.0107	0.0061
2.0	-0.0120	-0.0162
2.5	-0.0301	-0.0330
3.0	-0.0440	-0.0462
3.5	-0.0543	-0.0557
4.0	-0.0616	-0.0628
4.5	-0.0666	-0.0674
5.0	-0.0698	-0.0701
5.5	-0.0716	-0.0711
6.0	-0.0725	-0.0713
6.5	-0.0727	-0.0715
7.0	-0.0724	-0.0711
7.5	-0.0718	-0.0709
8.0	-0.0711	-0.0710
8.5	-0.0703	-0.0709
9.0	-0.0696	-0.0702
9.5	-0.0689	-0.0696
10.0	-0.0684	-0.0689

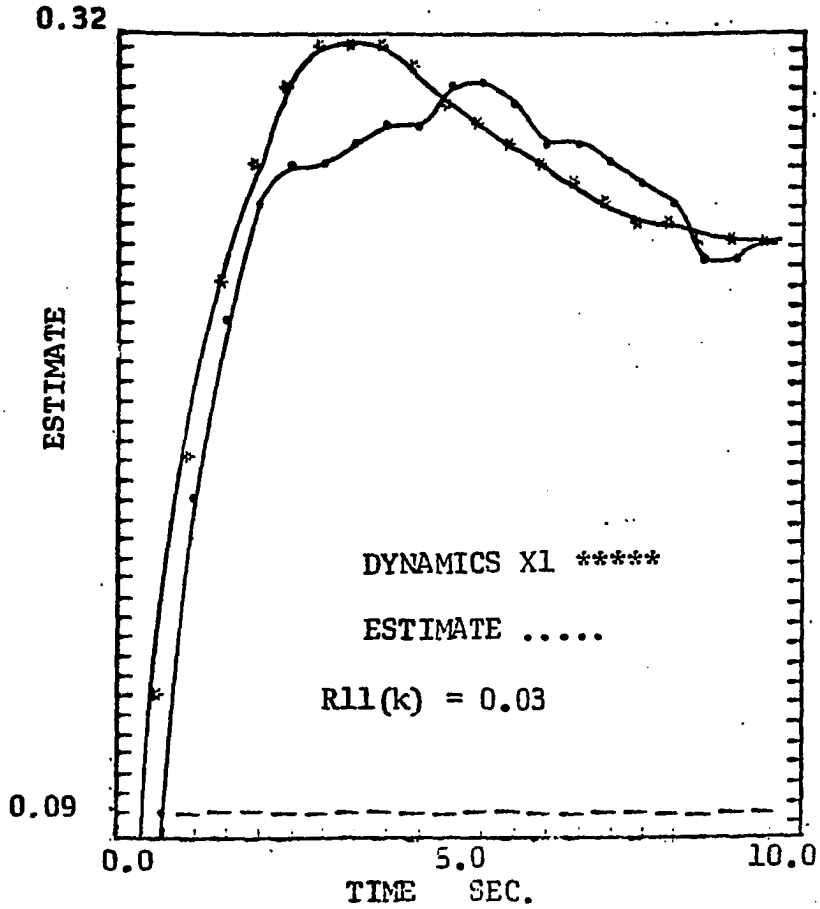


FIG. 2.12 EFFECT OF DIFFERENT R(K)
NONLINEAR EXTENDED KALMAN FILTER & CONSTANT NOISE

TIME SEC.	X 1	ESTIMATE
0.5	0.0988	0.0587
1.0	0.1810	0.1569
1.5	0.2427	0.2313
2.0	0.2845	0.2693
2.5	0.3097	0.2812
3.0	0.3222	0.2864
3.5	0.3255	0.2936
4.0	0.3228	0.2966
4.5	0.3165	0.2993
5.0	0.3083	0.3104
5.5	0.2995	0.3095
6.0	0.2907	0.3078
6.5	0.2827	0.2884
7.0	0.2755	0.2931
7.5	0.2695	0.2822
8.0	0.2645	0.2746
8.5	0.2607	0.2680
9.0	0.2578	0.2530
9.5	0.2559	0.2528
10.0	0.2547	0.2558

ACCUMULATED MEAN SQUARE ERROR = 0.3668E-03

order filter. All of the results shown improvement over his particular example and are probably due to the fact that we are using a better matrix inversion routine [G1] to reduce the possible round-off errors in the gain matrix calculation.

Generally speaking, the second order filter gives improved filtering results when compared to the nonlinear extended Kalman filter in correcting the system nonlinearity. This improvement will disappear as the sampling period is reduced to the point that the linearization error in carrying out the extended Kalman filter is not observable.

In our reactor model, with a sampling period of 0.5 second, the system nonlinearity is almost neglegible, thus the performance of nonlinear extended Kalman is essentially the same as nonlinear second order filter. However, the system nonlinearity error is so serious it can cause the divergence of nonlinear extended Kalman filter, while the nonlinear second order filter gives satisfactory results as shown in Fig. 2.14, and 2.15. Therefore, in the online implementation of nonlinear filtering, we suggest using nonlinear extended Kalman filtering only when the computational environment is suitable for small sampling periods. The linearization error will be acceptable within the smaller sampling periods but requires more frequent sampling and faster computational speed. On the other hand, the second order filter, with its correction terms for the system and measurement nonlinearity, proved to be more flexible for real time filtering.

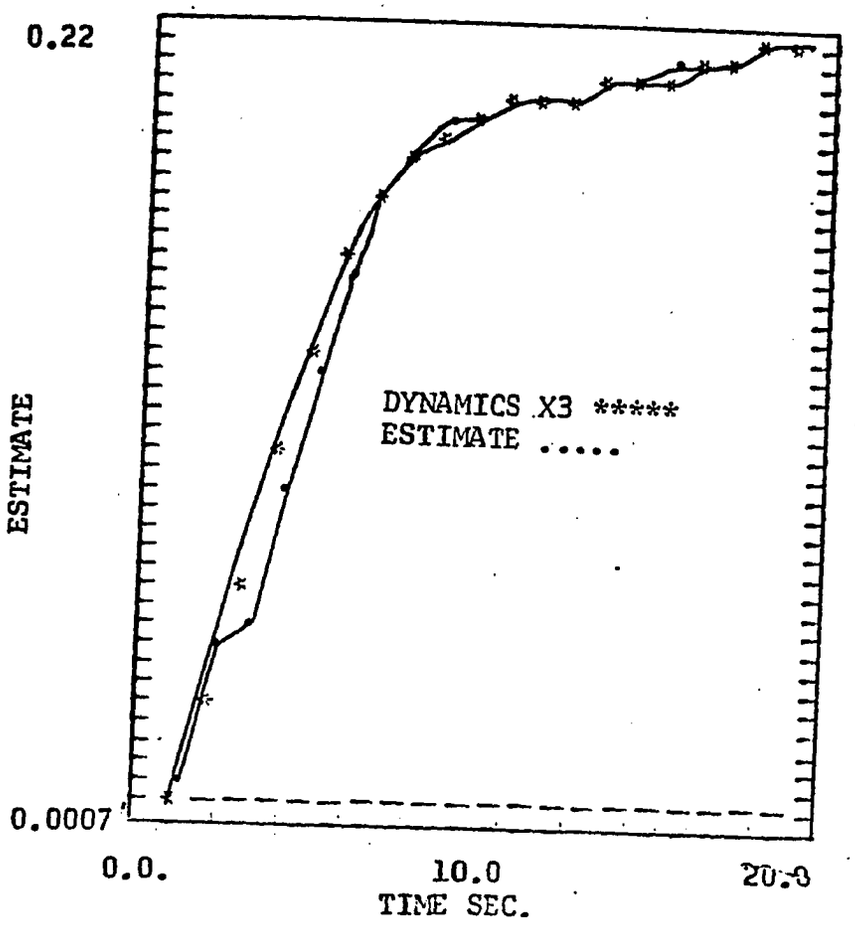


FIG. 2.13 DOUBLE THE SAMPLING PERIOD
SECOND ORDER NONLINEAR FILTER
CONSTANT NOISE

TIME SEC.	X 3	ESTIMATE
1.0	0.0071	0.0104
2.0	0.0321	0.0527
3.0	0.0677	0.0586
4.0	0.1042	0.0932
5.0	0.1355	0.1296
6.0	0.1594	0.1569
7.0	0.1764	0.1774
8.0	0.1877	0.1896
9.0	0.1949	0.1962
10.0	0.1995	0.2005
11.0	0.2025	0.2032
12.0	0.2046	0.2044
13.0	0.2065	0.2054
14.0	0.2082	0.2074
15.0	0.2100	0.2113
16.0	0.2119	0.2140
17.0	0.2139	0.2153
18.0	0.2159	0.2169
19.0	0.2180	0.2189
20.0	0.2201	0.2207

ACCUMULATED MEAN SQUARE ERROR = 0.3503E-04

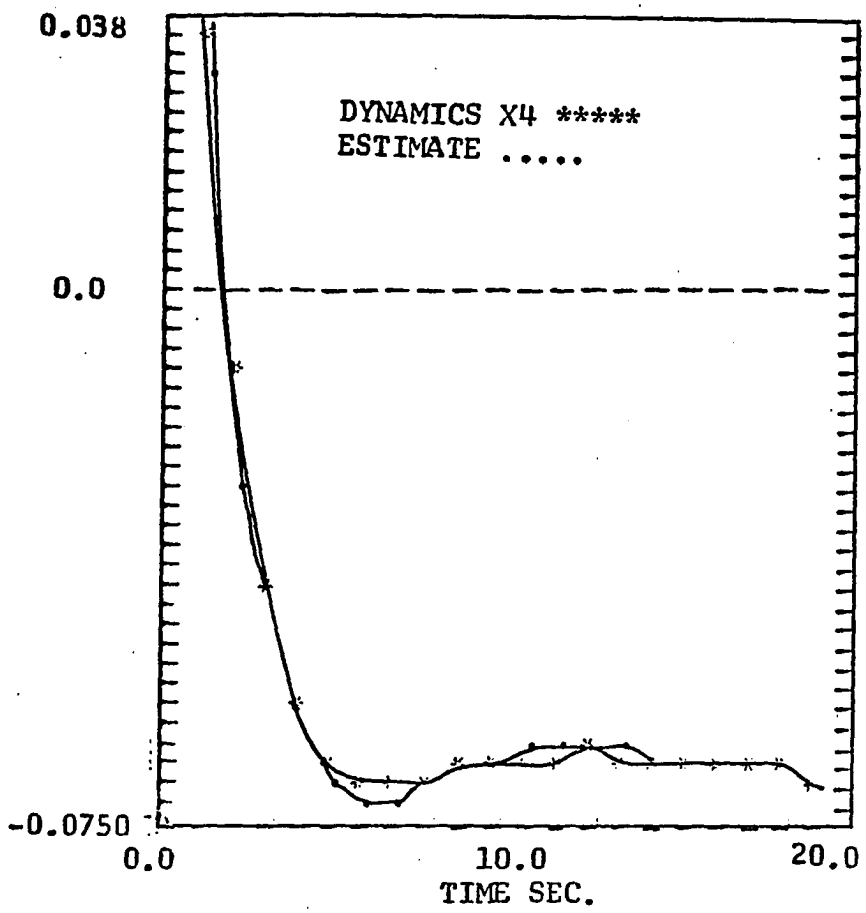


FIG. 2:14 DOUBLE THE SAMPLING PERIOD
SECOND ORDER NONLINEAR FILTER
CONSTANT NOISE

TIME SEC.	X 4	ESTIMATE
1.0	0.0380	0.0330
2.0	-0.0111	-0.0294
3.0	-0.0428	-0.0417
4.0	-0.0605	-0.0615
5.0	-0.0690	-0.0715
6.0	-0.0720	-0.0746
7.0	-0.0721	-0.0739
8.0	-0.0711	-0.0722
9.0	-0.0697	-0.0699
10.0	-0.0686	-0.0681
11.0	-0.0678	-0.0674
12.0	-0.0674	-0.0669
13.0	-0.0674	-0.0666
14.0	-0.0676	-0.0671
15.0	-0.0680	-0.0686
16.0	-0.0685	-0.0697
17.0	-0.0690	-0.0697
18.0	-0.0695	-0.0695
19.0	-0.0701	-0.0700
20.0	-0.0706	-0.0705

ACCUMULATED MEAN SQUARE ERROR = 0.1913E-04

CHAPTER III

COMBINED FILTERING AND CONTROL

III.1 Introduction

It is a universal desire to maximize the return on investment and this is especially true in engineering and business. This art of utilizing our capabilities most efficiently is generally termed 'Optimization'. From experience, optimization problems often appear paradoxical, in that it may appear impossible to find an optimal solution and yet we never give up searching for an optimal solution because of its importance.

A typical optimization problem is composed of the following parts:

A. Definition of the goal or the criteria.

B. Is the goal reachable? This includes the study of our environment; social factors, economic factors, and the problem itself.

C. Determination of the best policy based on A and B.

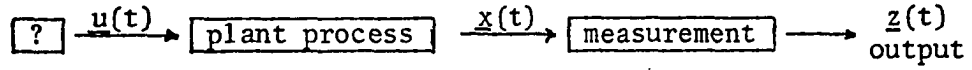
Mathematically speaking, we attack this problem in the following way:

1. Formulate the performance index.
2. Translate the physical problem descriptions and constraints into mathematical language.
3. Solve the problem mathematically.

The first two problems are very difficult since most physical process are too complicated to solve analytically or, due to our imprecise knowledge of the problem, we can never obtain an exact mathematical model for it. Applying the above approach to the real world can be termed 'optimal control'. Here, the term 'optimal control' may not be the real 'optimal' due to the deliberate simplification of the problem or imprecise knowledge of the process or the numerical approximations introduced in carrying out the solution. However, we settle for the 'next to optimal', or 'suboptimal' or simply 'optimal' result to make our way towards an ultimate solution.

Before we get involved in optimal stocastic control, let us introduce the five basic classes of control. We adopt Lee's [L3] pictorial approach:

A. Deterministic optimal control problem.

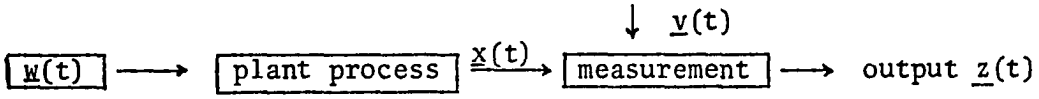


Given: The dynamic relationship between $\underline{x}$ and $\underline{u}$, $\underline{z}$ and $\underline{x}$.

Find: The control policy $\underline{u}(t)$ such that the output $\underline{z}(t)$ optimize our performance index.

B. Estimation problem:

Estimate the states of the system in a stocastic environment



Where $\underline{w}(t)$ is the driving process noise vector, and $\underline{v}(t)$ is the measurement noise vector.

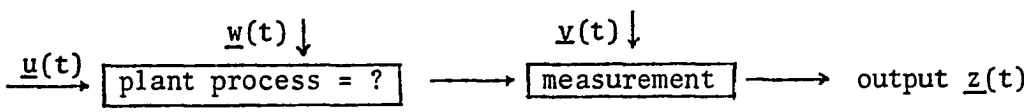
Given: 1. Relation between $\underline{x}$ and $\underline{w}$ (usually assume additive Gaussian white noise).

2. Relation between $\underline{z}$ and $\underline{x}$, $\underline{v}$ (also assume additive Gaussian white noise).

3. The statistical description of $\underline{w}$ and $\underline{v}$.

Find: The best estimate $\hat{x}(t/T)$ based on the measurement $\underline{z}(t)$ for t up to T. In the case t = T, we have the filtering problem; for t > T, we have the prediction problem and for t < T, the smoothing problem.

C. Identification problem.



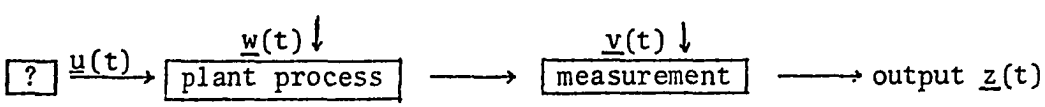
Given: 1. The statistical description of $\underline{w}$ and $\underline{v}$.

2. Physical relation between $\underline{z}$ and $\underline{x}$, $\underline{z}$.

3. The measurement $\underline{z}$ and input $\underline{u}$.

Find: The best estimate of the plant process.

D. Stocastic control problem.



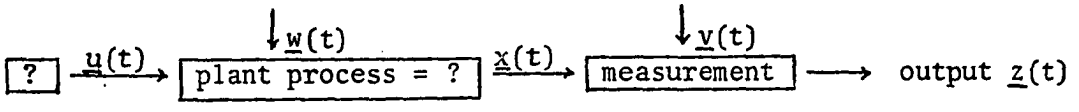
Given 1. Relation between $\underline{x}$ and $\underline{w}$, $\underline{u}$.

2. Relation between $\underline{z}$ and $\underline{z}$, $\underline{v}$.

3. The statistical description of $\underline{w}$ and $\underline{v}$.

Find: The optimal policy $\underline{u}(t)$ such that our best estimate of $\underline{x}(t)$ optimizes our performance index. If $\underline{u}(t)$ is a function of the output, we have the closed loop stocastic control problem.

E. Adaptive control problem.



Given: 1. The statistical description of $\underline{w}$, $\underline{v}$.

2. Relationship between $\underline{z}$ and $\underline{x}$, $\underline{v}$.

3. The output $\underline{z}(t)$ and the input $\underline{u}$ are measurable.

Find: $\underline{u}(t)$ such that the best estimates optimize the performance index.

Although we can classify the modern control problem into the above five areas, the actual problem we face may not belong to any single class. It may be some variation of the above categories or combinations among them.

1. A + B: Combined estimation and optimal control. In practice this cascades a filter to an optimal deterministic controller. If the process is linear with a quadratic performance index, the existence and uniqueness of the optimal stocastic control can be assured by the separation principle. The computational aspects of this problem can be carried out by calculating the filter estimate and the optimal controller gain separately. If the process is nonlinear, the separation principle no longer holds. Theoretically, there may exist several

local optimal solutions and we can no longer separate the filtering and control. Iterative solution of the split two point boundary value solution via quasilinearization and invariant embedding may be a way to approach a suboptimal solution. Lee and Sage [S2] have had some success in applying Bellman's original idea in this field.

2. D + E: This is a combined filtering, identification and optimal control problem. Here again, if the process is linear with a quadratic performance index, we can be assured of a unique optimal solution by the separation principle. This occurs because we can treat the uncertain model parameters as additional state variables and do the filtering for the augmented system. An alternative to this problem is to apply the model error compensation techniques or adaptive filtering to generate the best estimate and then cascade it to the optimal deterministic controller. Other variations in this category includes the so-called learning controller in which we apply on-line optimization using the plant's output, the controller iteratively applying a learning algorithm (refer to Fu) [F2] to learn the optimum control policy while minimizing the expected performance index (McLaren and Pan) [P4]. In the situation where the uncertain parameter only consists of a finite number of values, this is a typical filtering-detection-control problem where pattern recognition and statistical decision theory play an important role.

III.2 Combined Linear Filtering and Linear Control with a Quadratic Performance Index.

Problem Formulation and System Model

The System Model can be assumed to be described as below:

$$\underline{x}(k+1) = \Phi(k+1,k)\underline{x}(k) + \Gamma(k+1,k)\underline{w}(k) + \Psi(k+1,k)\underline{u}(k) \qquad \text{III-1}$$

with the observation:

$$\underline{z}(k+1) = H(k+1)\underline{x}(k+1) + \underline{v}(k+1) \qquad \text{III-2}$$

for

$$k = 0,1,2 \ldots$$

where

$\underline{x}$ = n - vector (state)

$\underline{u}$ = r - vector (control)

$\underline{w}$ = p - vector (system disturbance)

$\underline{v}$ = m - vector (measurement error)

Φ = nxn state transition matrix

$\Gamma(k+1,k)$ = nxp disturbance transition matrix

$\Psi(k+1,k)$ = nxr control transition matrix

$H(k+1)$ = mxn measurement matrix

$\underline{x}(0)$ = zero mean gaussian random vector with positive semidefinite covariance matrix P(0)

The performance index is defined

$$\text{Min } J_n = \text{MinE}\left[\sum_{i=1}^{N} \underline{x}'(i)A(i)\underline{x}(i) + \underline{u}'(i-1)B(i-1)\underline{u}(i-1)\right] \qquad \text{III-3}$$

Where A(i) and B(i-1) are positive semidefinite metrices which are nxn and rxr respectively, and E is the expected value operator. Optimal

control of the above stocastic system was first conjectured by Kalman and Koepcke [K3] with the separation principle stated as follows:

Separation Principle.

The optimal control system for the stocastic linear system described above consists of the optimal linear filter cascaded with the optimal feedback gain matrix of the deterministic regulator. The parameter for the two parts of the control system are determined separately. This principle was proved by Joseph [B8].

Implementation.

1. Estimation:

$$\underline{\hat{x}}(k|k-1) = \Phi(k,k-1)\underline{\hat{x}}(k-1|k-1) + \Psi(k,k-1)\underline{u}(k-1) \qquad \text{III-4}$$

$$\underline{\hat{x}}(k|k) = \underline{\hat{x}}(k|k-1) + K(k)[\underline{z}(k)-H(k)\underline{\hat{x}}(k|k-1)] \qquad \text{III-5}$$

Where K(k) is the Kalman filter gain defined exactly the same as in the linear Kalman filter.

2. Optimal Control:

$$\underline{u}(k) = S(k)\underline{x}(k) \qquad \text{III-6}$$

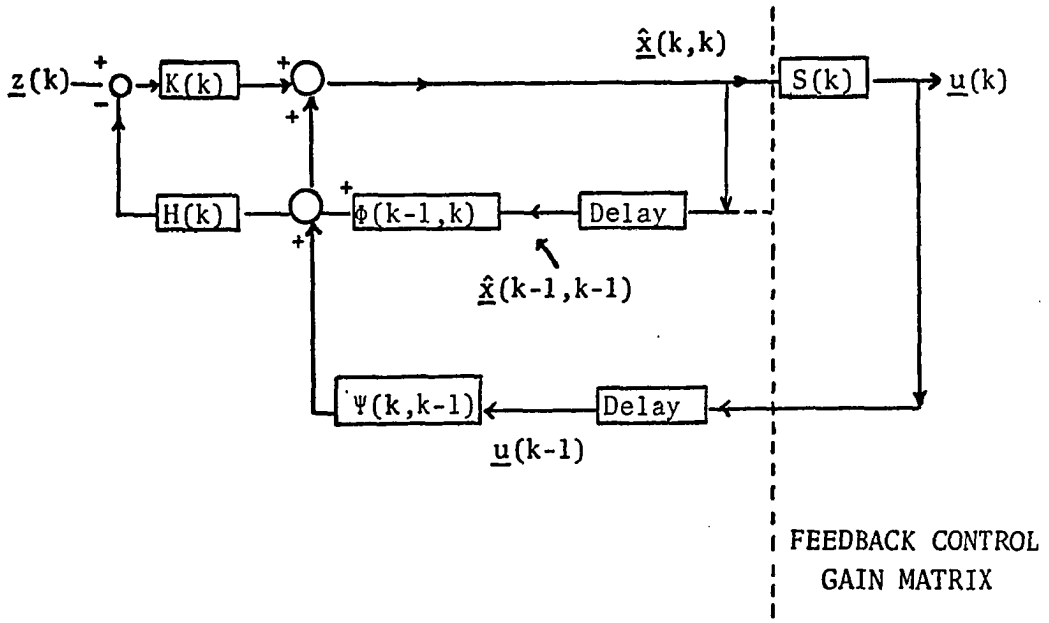
$$S(k) = -[\Psi'(k+1,k)W(k+1)\Psi(k+1,k)+ B(k)]^{-1} \Psi(k+1,k)W(k+1)\Phi(k+1,k) \qquad \text{III-7}$$

$$W(k) = \Phi'(k+1,k)W(k+1)\Phi(k+1,k)+ \Phi'(k+1,k)W(k+1)\Psi(k+1,k)\cdot S(k) + A(k) \qquad \text{III-8}$$

for

$k = N-1, N-2, \ldots 0$ and $W(N) = A(N)$.

A block diagram for this algorithm is shown below:



$k = 1,2, \ldots,N-1$

OPTIMAL FILTER

The above algorithm has already been developed and completely tested against the following example given by Meditch:

$$x(k+1) = x(k) + w(k) + 2u(k)$$

$$z(k+1) = x(k+1) + v(k+1)$$

$$J_3 = E[x^2(3) + \sum_{i=1}^{3} u^2(i-1)]$$

with

$$Q(k) = 25,\ R(k+1) = 15, P(0) = 100; \quad x(0) = 0$$

the filter equation is

$$\hat{x}(k|k) = \hat{x}(k-1|k-1) + 2u(k-1) + K(k)[z(k)-\hat{x}(k-1|k-1)-2u(k-1)]$$

where

$$u(k-1) = S(k-1)\hat{x}(k-1|k-1) \quad .$$

The following data has been checked out:

A. Filter:

k	P(k|k-1)	K(k)	P(k|k)
0	...	...	100
1	125	0.893	13.39
2	38.4	0.720	10.80
3	35.8	0.704	10.57

B. Optimal feedback controller:

k	S(k)	W(k)
3	...	1
2	-0.400	0.200
1	-0.222	0.111
0	-0.154	0.077

Therefore the optimal control at each state is:

$$u(0) = -0.154\ \hat{x}(0|0) = 0$$

$$u(1) = -0.222\ \hat{x}(1|1)$$

$$u(2) = -0.4\ \hat{x}(2|2)$$

The dual relation between filtering and control can be put into the following correspondence:

	Filtering	Control
transition matrix	F(t)	F'(-t)
time	(t_0,t)	$(-t,-t_0)$
weighting matrix	$\Gamma(t)$	K'(-t)

gain matrix	$H(t)$	$\Psi'(-t)$
performance index	$\frac{1}{2}\|\|\underline{x}-\hat{\underline{x}}\|\|^2_p$	J_n
	Controllability Unbiased Estimation	Observability Minimum Energy
weighting matrix	$Q(t)$	$A(-t)$
	$R(t)$	$B(-t)$

Notice that the control sequence are calculated backward (negative time sense). This duality correspondence allows us to develope the control policy simply by dualizing the theory of linear filter.

III.3 Suboptimal Nonlinear Stocastic Control

Consider a general nonlinear stocastic control problem. Let the system be described by n state vector $\underline{x}(t)$ with the observation m vector $\underline{y}(t)$ corrupted with additive noise v:

$$\text{System:} \quad \underline{\dot{x}} = \underline{f}(\underline{x},\underline{u},\underline{k},t) \qquad \text{III-9}$$

$$\text{Observation:} \quad \underline{y} = \underline{h}(x,t) + \underline{v} \qquad \text{III-10}$$

$$\text{Performance Index:} \quad J = E\int_0^{t_f} F(\underline{x},\underline{u},t)dt \ . \qquad \text{III-11}$$

The optimal stocastic control for the system can be stated as 'chose $\underline{u}(t)$ to minimize J'.

Solution of III-9 to III-11 is not possible even in the deterministic nonlinear optimal control problem and often involves solving a two point split boundary value problem. Payne [P1] presents a systematic computation approach to the solution of optimal control problem. Balakrisknan [B1] recently presented the 'Epsilon technique' to avoid

the integration of dynamic equation by converting it into a minimization problem it solves simultaneously the twin problems of integration and finding the optimal solution. The separation principle for the linear optimal stocastic control is a special case for the solution of III-9 to III-11; and assumes III-9 and III-10 are linear and III-11 is quadratic.

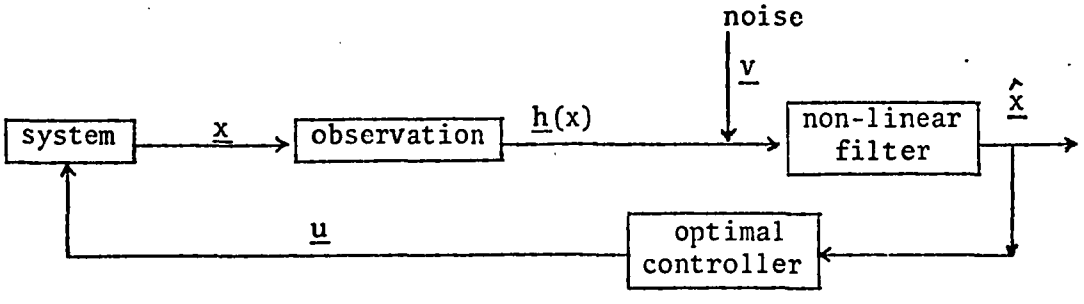
Although the structure of the optimal nonlinear solution is unknown, a suboptimal stocastic controller can be based on the idea of separating the filtering and control. Thus, cascading a nonlinear filter to the nonlinear optimal deterministic controller presents a reasonable approach. Assuming the non-linear filter is represented by non-linear extended Kalman filter, we then have the suboptimal estimator and its error covariance matrix:

$$\dot{\hat{\underline{x}}} = \underline{f}(\hat{\underline{x}},\underline{u},\underline{k},t) + Ph_x'H[\underline{y}-\underline{h}(\hat{\underline{x}},t)] \qquad \text{III-12}$$

$$\dot{P} = f_x P + Pf_x' + P[h_x'H(\underline{y}-\underline{h}(\hat{\underline{x}},t))]_x P \qquad \text{III-13}$$

where $\underline{k}$ is the model parameter vector. In the case of combined identification and control, we can augment the $\dot{\underline{k}} = 0$ into the filter state equation.

Seinfeld [S8] presents the following scheme as one possible approach to the solution of the problem:



The solution of the system equations III-9 to III-11 can be simplified by replacing the overall integral of III-11 by a single stage. This idea was successfully applied to nonlinear control by Lapidus, Lee [L1], [L3]. Applying the idea of Lee: Let the system be described by the discrete data form:

1. System $\underline{x}(k+1) = f[\underline{x}(k),\underline{u}(k),\underline{w}(k)]$ III-14

   Observation $\underline{z}(k) = \underline{h}[\underline{x}(k),\underline{v}(k)]$ III-15

2. Optimal filter. We assume that the nonlinear optimal filtering estimate $\hat{\underline{x}}(k,k)$ is given at all times.

3. Optimal prediction. $\hat{\underline{x}}(k+1,k) = \underline{f}[\hat{\underline{x}}(k,k),\underline{u}(k)]$ III-16

4. Design the controller by single stage optimization, such that:

$$\min_{\underline{u}(k)} J = ||\hat{\underline{x}}(k+1,k)||^2_{A(k)} + ||\underline{u}(k)||^2_{B(k)} \qquad \text{III-17}$$

subject to the constraints

$$\hat{\underline{x}}(k+1,k) = \underline{f}[\hat{\underline{x}}(k,k),\underline{u}(k)] \ . \qquad \text{III-18}$$

5. The solution to the above problem is

$$\underline{u}(k) = - B^{-1}(k) \frac{\partial f}{\partial u(k)} A(k) \quad f[\hat{\underline{x}}(k,k), \underline{u}(k)] \ . \qquad \text{III-19}$$

Therefore, we can solve for $\underline{u}(k)$ as a function of $\hat{x}(k,k)$ or $\hat{x}(k+1,k)$, this result is optimal for a single stage process. This technique can usually be applied to the closed loop optimal regulator where the error is not very large. We are able to approximate the behavior of a multistage process by properly choosing the weighting matrix functions $A(k)$, $B(k)$ such that they are time varying and proportional to the magnitude of the error.

A special case of nonlinear stocastic suboptimal control was treated by Seinfeld [S2] recently, where he assumes the control appears linearly in a nonlinear stocastic process, and thus uses the single stage process performance index

$$\min_{\underline{u}} J = (\underline{\hat{x}}-\underline{x}^d)'Q(\underline{\hat{x}}-\underline{x}^d) \qquad \text{III-20}$$

Thus by minimizing J at each instant of time, such that the rate of movement toward $\underline{x}^d$ is a maximum,

$$\begin{aligned}\dot{F} &= 2(\underline{\hat{x}}-\underline{x}^d)'Q\dot{\underline{\hat{x}}} \\ &= 2(\underline{\hat{x}}-\underline{x}^d)'Q[f(\underline{x},\underline{u},\underline{k},t)] + Ph_{\underline{x}}'H[\underline{y}-\underline{h}(\underline{x},t)]\end{aligned} \qquad \text{III-21}$$

Assuming

$$\dot{\underline{\hat{x}}} = \underline{w}(\underline{\hat{x}},\underline{k},\underline{t}) + S(\underline{\hat{x}},\underline{k},t)\underline{u} \qquad \text{III-22}$$

and substitute III-22 into III-21; we then have

$$\dot{F} = 2(\underline{\hat{x}}-\underline{x}^d)Q[w(\underline{\hat{x}},\underline{k},t)] + S(\underline{\hat{x}},\underline{k},t)\underline{u} + Ph_{\underline{x}}'H[\underline{y}-\underline{h}(\hat{x},t)] \; . \qquad \text{III-23}$$

Let

$$\beta = 2S'(\underline{\hat{x}},\underline{f},t)Q(\underline{\hat{x}}-\underline{x}^d) \; .$$

A minimum F is obtained if

$$\begin{aligned}u_i &= u_i^* \text{ for } \beta_i < 0 \\ &= u_{i*} \text{ for } \beta_i > 0\end{aligned} \qquad \text{III-24}$$

This is a bang-bang type of control with singular control when $\beta_i = 0$.

Seinfeld applied this algorithm to a CSTR with first order reaction, a two-state variable problem.

Fortunately, the control vector also appears linearly in our nonlinear reactor model. Rewriting our reactor model in the previous chapter:

$$\dot{x}_1 = -(c_1+c_4)x_1 + c_3(1+x_4)^2\exp[k_1x_1/(1+x_1)] + c_4x_2 + c_1u_1 \qquad \text{III-25}$$

$$\dot{x}_2 = -(c_5+c_6)x_2 + c_5x_1 + c_6x_3 \qquad \text{III-26}$$

$$\dot{x}_3 = -(c_7+c_8)x_3 + c_8x_2 + c_7u_2 \qquad \text{III-27}$$

$$\dot{x}_4 = -c_1x_4 - c_2(1+x_4)^2\exp[K_1x_1/(1+x_1)] + c_1u_3 \qquad \text{III-28}$$

where $u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ is the control vector and is defined as

$$u_1 = \frac{T_i - T_s}{T_s}\ ; \qquad u_2 = \frac{T_{c_i} - T_s}{T_s}\ ; \qquad u_3 = \frac{C_{A_i} - C_s}{C_s}\ .$$

It is obvious that $\underline{u}$ appears linearly in Equations III-25 to III-28, with the constant control gain matrix

$$S = \begin{vmatrix} c_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & c_7 & 0 \\ 0 & 0 & c_1 \end{vmatrix}\ .$$

Assuming in identity weighting matrix and $\underline{x}^d = 0$, since we are interested in returning to the steady state

$$\beta = \begin{vmatrix} c_1 & 0 & 0 & 0 \\ 0 & 0 & c_7 & 0 \\ 0 & 0 & 0 & c_1 \end{vmatrix} \begin{vmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{vmatrix}$$

with

$$u_i = u_i^* \quad \text{if} \quad \beta_i < 0$$

$$= u_{i*} \quad \text{if} \quad \beta_i > 0 \ .$$

Then, we have the bang-bang controller. In our case, the choice of u_i^* and u_{i*} is very important, with $u_i^* = 1.0$, and $u_{i*} = -1.0$ will cause model divergence. This means the heat removal rate is too small to quench the reaction heat, thus leads to runaway reaction. With too high a value of the bounds will lead to excessive oscillation of the dynamics and degrading of the product.

The best set of values seems to be:

$$u_1^* = 5.0;\ u_{1*} = -5.0; \quad u_2^* = 0.5,\ u_{2*} = -0.5; \quad u_3^* = 1.0,\ u_{3*} = -1.0$$

Using these set of bounds, the transients due to a 10 per cent upset in x4 disappeared almost instantaneously. Thus the dynamics are able to return to the desired steady state in a very short period without degrading the product quality.

Another type of control introduced by Wells, as we used it throughout our work for comparison purposes is steady state control.

This can easily be obtained by setting the left hand side of Equations III-25 to III-28 to zero; the resulting control will be:

$$u_1 = -c_3/c_1; \quad u_2 = 0.0; \quad u_3 = c_2/c_1 \quad .$$

Substituting these relations into Equation III-25 to III-28 thus arriving at our reactor model equations as we use it throughout studies. In our simulation, we found that bang-bang control is much better than Wells' steady state control in compensating the initial 10 per cent upset in concentration. Figure 3.1 shows that the bang-bang control can keep the reactor wall temperature (x2) almost at steady state while it increases steadily with a maximum offset of 22 per cent by steady state control. Figure 3.2 indicates that bang-bang control can compensate the concentration offset while steady state fails in 10 seconds. Figure 3.3 and Fig. 3.4 show bang-bang control can compensate the transient even with a reduced rate constant RK = 6.0 (corresponds to higher reaction rate) while Wells' steady state model diverges.

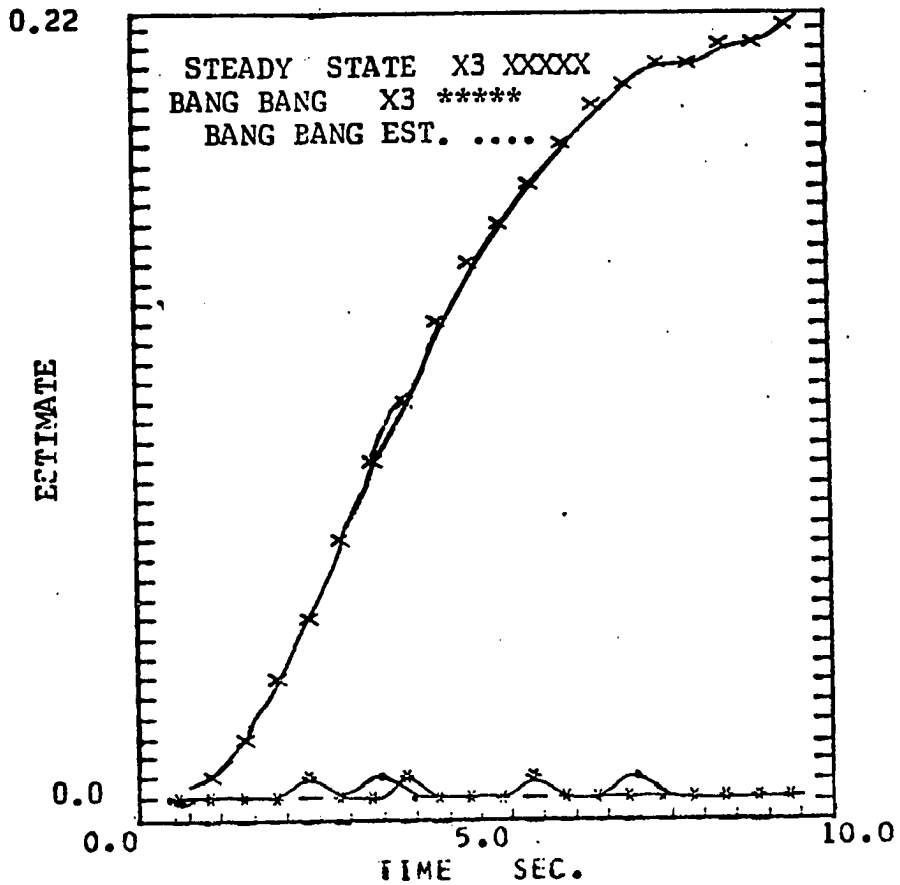


FIG. 3.1 BANG BANG CONTROL
SECOND ORDER NONLINEAR FILTER
HUANGS ADAPTIVE I

TIME SEC.	STEADY ST. X 3	BANG BANG X3	BANG BANG ESTIMATE
0.5	0.0007	0.0002	0.0002
1.0	0.0052	0.0001	0.0002
1.5	0.0150	-0.0003	-0.0004
2.0	0.0294	0.0004	0.0005
2.5	0.0472	0.0019	0.0018
3.0	0.0666	0.0009	0.0005
3.5	0.0863	0.0000	0.0016
4.0	0.1051	0.0013	0.0004
4.5	0.1225	-0.0002	0.0008
5.0	0.1379	0.0002	0.0006
5.5	0.1513	0.0001	0.0002
6.0	0.1625	0.0018	0.0016
6.5	0.1718	0.0002	-0.0004
7.0	0.1793	0.0006	-0.0004
7.5	0.1853	0.0003	0.0014
8.0	0.1900	-0.0007	0.0001
8.5	0.1936	0.0000	0.0007
9.0	0.1964	0.0002	0.0007
9.5	0.1986	-0.0002	0.0000
10.0	0.2002	0.0012	-0.0014

ACCUMULATED MEAN SQUARE ERROR = 0.7786E-06

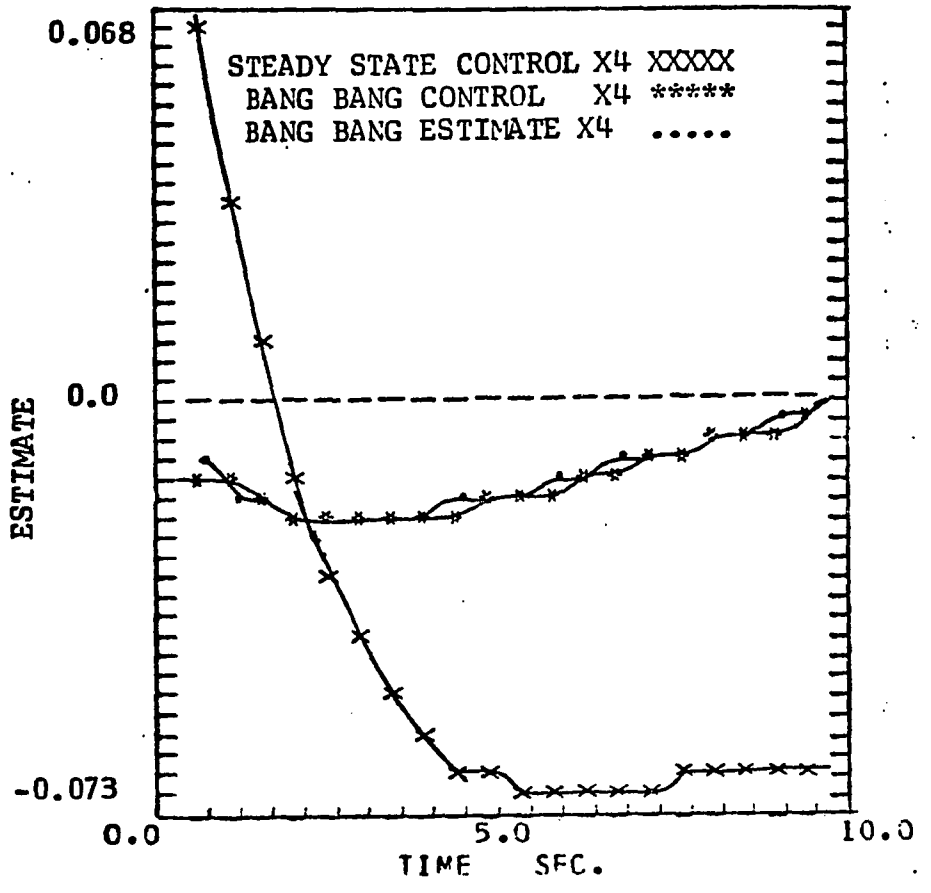


FIG. 3.2 BANG BANG CONTROL
SECOND ORDER NONLINEAR FILTER
HUANGS ADAPTIVE I

TIME SEC.	STEADY ST. X 4	BANG BANG X4	BANG BANG ESTIMATE
0.5	0.0683	-0.0134	-0.0132
1.0	0.0373	-0.0167	-0.0172
1.5	0.0097	-0.0191	-0.0194
2.0	-0.0135	-0.0207	-0.0208
2.5	-0.0319	-0.0216	-0.0214
3.0	-0.0459	-0.0220	-0.0214
3.5	-0.0562	-0.0220	-0.0211
4.0	-0.0634	-0.0216	-0.0208
4.5	-0.0682	-0.0209	-0.0199
5.0	-0.0712	-0.0200	-0.0187
5.5	-0.0727	-0.0188	-0.0176
6.0	-0.0733	-0.0175	-0.0163
6.5	-0.0732	-0.0160	-0.0148
7.0	-0.0727	-0.0145	-0.0132
7.5	-0.0719	-0.0129	-0.0117
8.0	-0.0710	-0.0113	-0.0102
8.5	-0.0702	-0.0097	-0.0087
9.0	-0.0693	-0.0081	-0.0072
9.5	-0.0686	-0.0065	-0.0057
10.0	-0.0680	-0.0050	-0.0034

ACCUMULATED MEAN SQUARE ERROR = 0.9340E-06

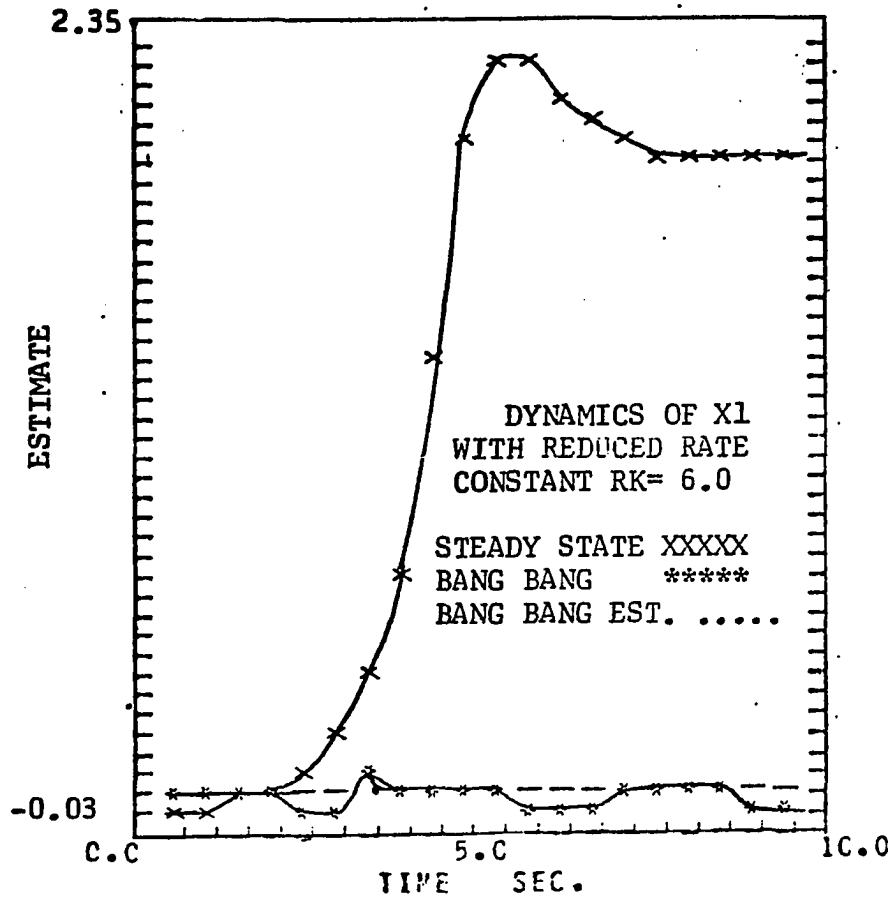


FIG. 3.3 BANG BANG CONTROL
NONLINEAR EXTENDED KALMAN FILTER
HUANGS ADAPTIVE I

TIME SEC.	STEADY STATE X 1	BANG BANG X1	BANG BANG ESTIMATE
C.5	C.C186	C.C422	C.0422
1.0	C.C41C	C.C274	C.C274
1.5	C.C698	0.C136	C.0136
2.C	C.1C93	-C.CCC9	-C.CCC8
2.5	0.1671	-C.C139	-C.C142
3.0	C.2598	-C.C275	-C.0275
3.5	C.4262	C.C585	C.C576
4.0	C.7673	C.C459	C.C445
4.5	1.4398	C.C328	C.0318
5.0	2.1331	C.C189	C.C18C
5.5	2.3749	0.CC61	0.0053
6.C	2.3520	-0.CC74	-C.0077
6.5	2.2590	-C.C213	-C.C22C
7.0	2.1682	-0.0342	-0.0351
7.5	2.C996	C.C5C6	C.C5C6
8.0	2.C559	0.C361	C.C362
8.5	2.C342	0.C225	0.0222
9.0	2.C3C2	C.CC98	C.CC95
9.5	2.C396	-C.CC47	-C.CC49
1C.0	2.C5E4	-C.C184	-C.0185

ACCUMULATED MEAN SQUARE ERROR = C.3461E-06

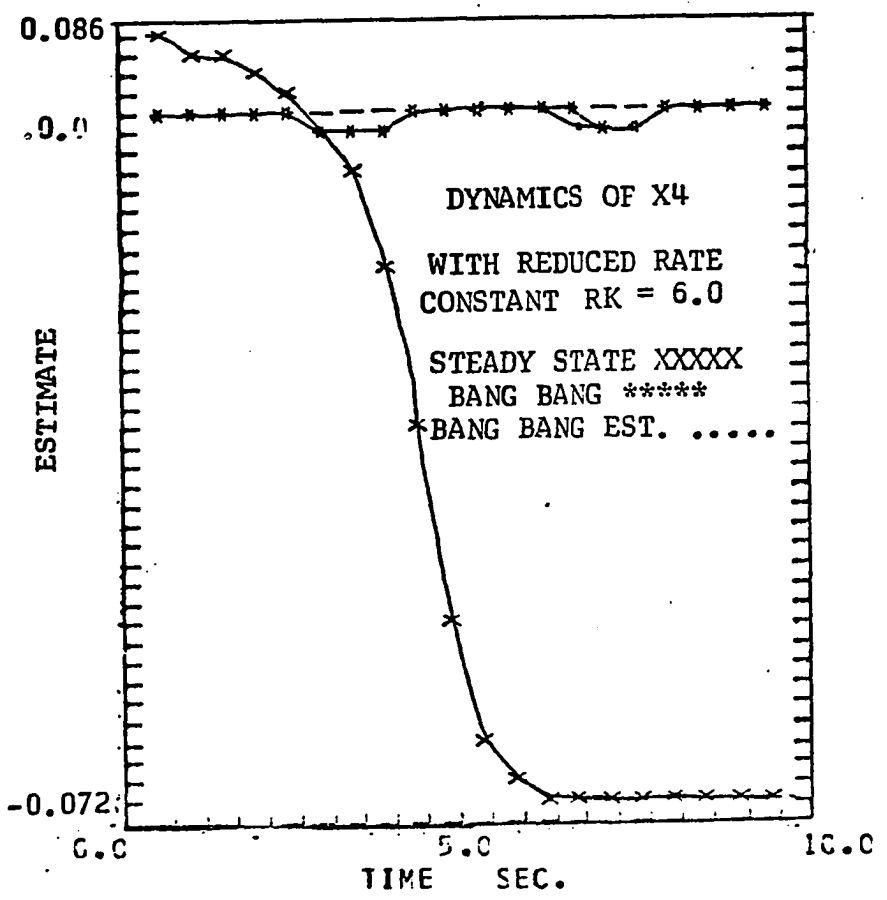


FIG. 3.4 BANG BANG CONTROL
NONLINEAR EXTENDED KALMAN FILTER
HUANGS ADAPTIVE I

TIME SEC.	STEADY STATE X 4	BANG BANG X4	BANG BANG ESTIMATE
0.5	0.0863	-0.0030	-0.0031
1.0	0.0721	-0.0002	-0.0002
1.5	0.0566	0.0020	0.0020
2.0	0.0388	0.0049	0.0048
2.5	0.0167	0.0070	0.0069
3.0	-0.0137	-0.0106	-0.0106
3.5	-0.0619	-0.0079	-0.0078
4.0	-0.1521	-0.0059	-0.0055
4.5	-0.3246	-0.0036	-0.0031
5.0	-0.5234	-0.0008	-0.0002
5.5	-0.6368	0.0010	0.0018
6.0	-0.6843	0.0035	0.0040
6.5	-0.7015	0.0062	0.0070
7.0	-0.7056	0.0082	-0.0110
7.5	-0.7045	-0.0085	-0.0083
8.0	-0.7018	-0.0056	-0.0054
8.5	-0.6991	-0.0033	-0.0028
9.0	-0.6972	-0.0015	-0.0010
9.5	-0.6964	0.0014	0.0021
10.0	-0.6966	0.0039	0.0045

ACCUMULATED MEAN SQUARE ERROR = 0.1859E-04

CHAPTER IV

MODELING ERROR COMPENSATION AND ADAPTIVE FILTERING

IV. Introduction

A perfect model does not exist which can represent exactly the physical problem. This can either be due to incomplete knowledge of the physical system, or the system is too complex to be modeled exactly. Therefore, it is normal to use an approximate mathematical model which can represent the system in the region of interest. A good model has to be accurate enough to provide useful numerical results and yet simple enough to fit the computational environment.

The restrictions on the applicability of filtering are:

1. The process model is described by a set of differential equations in the continuous case and difference equations in the discrete case.

2. The random noise in the system and the measurement are independent and uncorrelated (can be relaxed).

3. The system can be represented by a finite number of state variables (preferably as low as possible).

4. The process dynamic differential equations can be linearized successfully (or differentiable).

Often the model parameters have uncertainty associated with these, (in some cases the parameter is completely unknown) and only an estimated value can be used. Yet the model dynamic response can be very sensitive to the variations of this parameter. In filtering, we can treat the estimated parameter as an additional state variable and do the filtering for the augmented system. Processing of the measurement observations will adjust the parameter value and converges it to the actual parameter value, and thus gives a better model estimate. However, additional state variables not only increase the computational load, but also increases the computer memory storage requirement drastically. It is obvious that an alternative model error compensation technique is needed.

Various on-line and off-line model error compensation techniques have been tested on the reactor model in this work and can compensate several types of model errors. The on-line model error compensation (or adaptive) techniques developed in this work have been shown to be powerful in the sense that they do not require any additional computer storage and computation time yet the results are better than most of the existing error compensation techniques.

IV.2 Modeling Errors and Filter Divergence

There are several types of errors in modeling the system dynamics. Generally the most serious error is the uncertainty in modeling physical parameters. If the filter is constructed on the basis of an erroneous model, even if the error covariance matrix and filter gain are kept low, the filter learns the wrong model too well after processing

enough observations. Thus, the filter model and the actual model going in different directions may eventually lead to catastrophic results. The phenomenon is called 'filter divergence'.

This divergence often occurs when the noise inputs to the system are small and measurement noise is small. Then the filter gain and error covariance are kept very low so that the filter is capable of learning the incorrect model too well. The subsequent observations then have little effect on the estimate. However, the actual system model is different from the filter model so that the estimate and the actual state diverge.

For a nonlinear system, filter divergence is often related to the stability of the system. This is especially true for our reactor model since there exists three steady state solutions (Lapidus, McGuire) [L1), [M1]. Therefore, uncertainty in model parameters may trigger the filter estimate leading to an undesired steady state.

A typical filter divergence in the filtering of the reactor model is shown in Fig. 4.1, 4.2 where the actual model used a reduced rate constant of 1.0 while the filter model used 0.8. The calculated error covariances based on the filter model are too low, as is the gain matrix. Therefore, the filter is capable of learning the incorrect model (with a rate constant of 0.8) too well and gives too low an estimate for both the concentration (x4) and temperature (x1). Other types of errors such as deliberate simplification of a higher dimension system and replacing it by a lower order model may also cause model divergence. From the above it is clear that model error compensation and adaptive filtering techniques are needed in filtering practice.

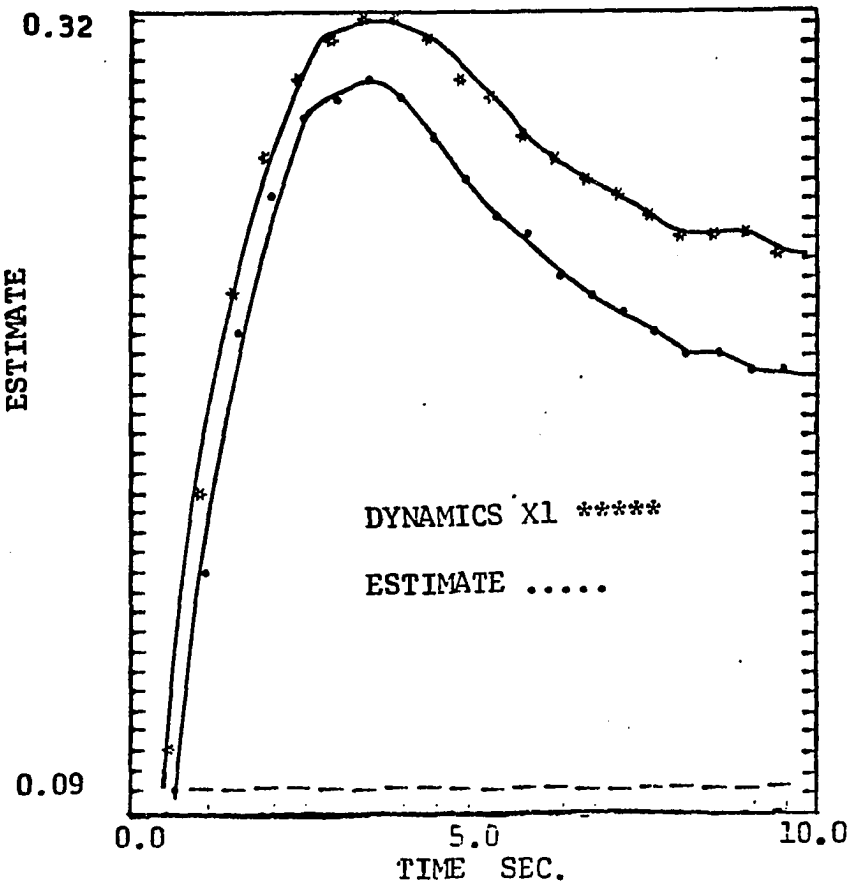


FIG. 4.1 FILTER DIVERGENCE
NONLINEAR EXTENDED KALMAN FILTER WITH
20 % rate constant error (WITHOUT COMPENSATION)

TIME SEC.	X 1	ESTIMATE
0.5	0.0988	0.1026
1.0	0.1810	0.1654
1.5	0.2427	0.2262
2.0	0.2845	0.2529
2.5	0.3097	0.2761
3.0	0.3222	0.2863
3.5	0.3255	0.2891
4.0	0.3228	0.2852
4.5	0.3165	0.2788
5.0	0.3083	0.2695
5.5	0.2995	0.2598
6.0	0.2907	0.2517
6.5	0.2827	0.2434
7.0	0.2755	0.2372
7.5	0.2695	0.2325
8.0	0.2645	0.2277
8.5	0.2607	0.2249
9.0	0.2578	0.2223
9.5	0.2559	0.2215
10.0	0.2547	0.2207

ACCUMULATED MEAN SQUARE ERROR = 0.1166E-02

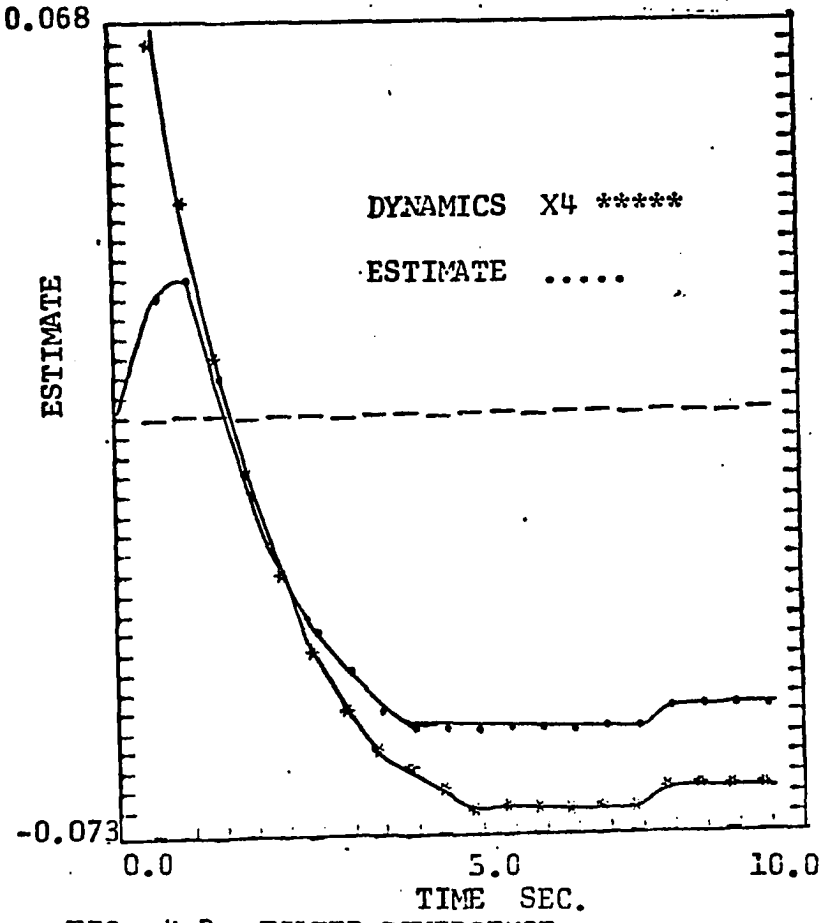


FIG. 4.2 FILTER DIVERGENCE
NONLINEAR EXTENDED KALMAN FILTER WITH
20 % RATE CONSTANT ERROR (WITHOUT COMPENSATION)

TIME SEC.	X 4	ESTIMATE
0.5	0.0683	0.0237
1.0	0.0377	0.0158
1.5	0.0107	-0.0006
2.0	-0.0120	-0.0180
2.5	-0.0301	-0.0317
3.0	-0.0440	-0.0420
3.5	-0.0543	-0.0495
4.0	-0.0616	-0.0544
4.5	-0.0666	-0.0575
5.0	-0.0698	-0.0589
5.5	-0.0716	-0.0593
6.0	-0.0725	-0.0593
6.5	-0.0727	-0.0585
7.0	-0.0724	-0.0578
7.5	-0.0718	-0.0570
8.0	-0.0711	-0.0559
8.5	-0.0703	-0.0551
9.0	-0.0696	-0.0542
9.5	-0.0689	-0.0537
10.0	-0.0684	-0.0531

ACCUMULATED MEAN SQUARE ERROR = 0.2523E-03

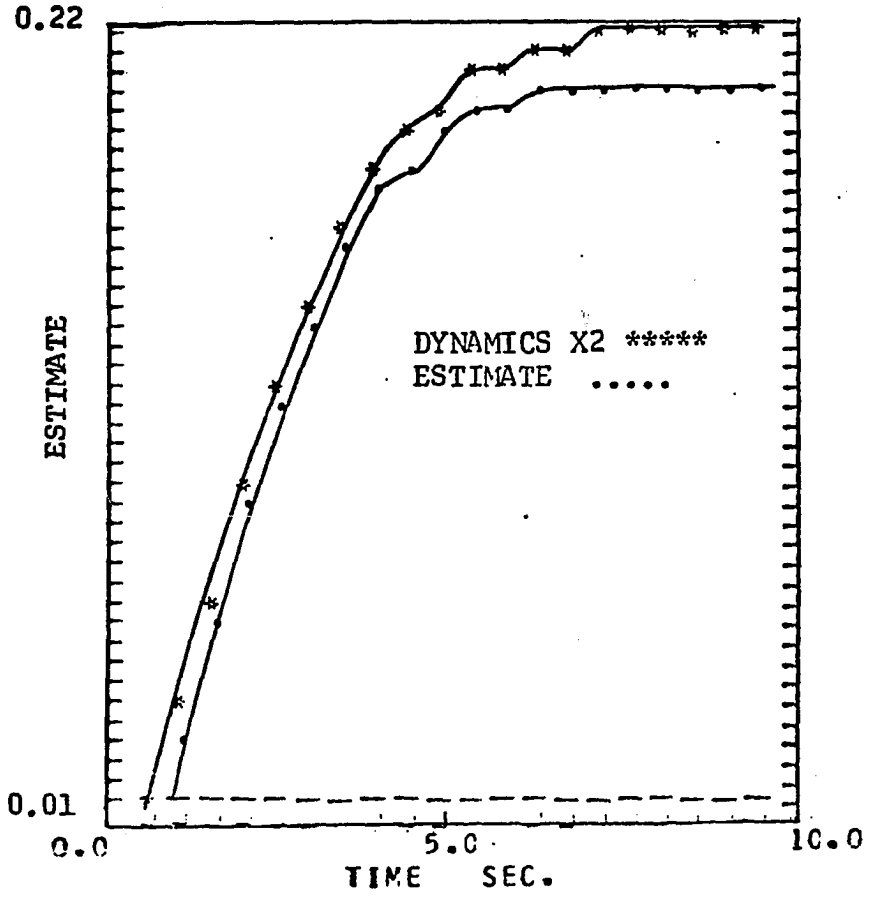


FIG.4.3 FILTER DIVERGENCE
SECCND CRDER NCNLINEAR FILTER
NO NOISE CCMPENSATICN

TIME SEC.	X 2	ESTIMATE
C.5	C.C118	C.C118
1.0	C.0369	0.C331
1.5	0.C673	0.0636
2.C	C.0980	0.C910
2.5	0.1261	0.1191
3.0	C.1503	0.1433
3.5	C.17C0	0.1637
4.0	0.1857	0.1790
4.5	C.1977	0.19C9
5.0	0.2067	0.1984
5.5	0.2132	0.2034
6.0	C.2178	C.2072
6.5	0.2210	0.2C88
7.0	0.2231	0.2103
7.5	0.2244	0.2113
8.0	0.2252	C.2110
8.5	0.2257	0.2112
9.0	0.2261	C.21C6
9.5	0.2264	C.21C9
10.0	0.2267	0.2107

ACCUMULATED MEAN SQUARE ERRCR = 0.11C6E-03

IV.3 Modeling Error Compensation Techniques

1. State Augmentation (Schmidt Filter [S4]).

Most of the errors in filtering can be traced to the uncertainty in model parameters. Instead of treating the parameters as state variables and filtering the augmented system, an alternative is to use the Schmidt filter. The Schmidt filter includes the model uncertainty in the calculation of the augmented error covariance matrix, and thus raises the error covariance matrix and therefore raises the Kalman filter gain matrix and makes the proper compensation. The systematic approach to compensate model error by this filter also makes it suitable for on-line adaptive filtering.

The numerical experience of applying this technique to off-line error compensation due to 20 per cent error in the rate constant K_1 can be summarized as follows:

A. Computation of the augmented error covariance still requires a significant amount of computer storage and computation time.

B. Precise knowledge of the uncertainty in the parameters is required to start reliable filtering. Too low an initial parameter error covariance gives too small a compensation such that the accumulation of uncertainty in modeling parameters will still give poor filter performance. On the other hand, too high an initial parameter error covariance matrix will drive the filter outside the observability and controllability region (the error covariance matrix of the estimate becoming negative) and therefore cause filter divergence.

C. Even with the above disadvantages acting against this filter, the Schmidt filter offers a stable, consistent and smooth

filter estimate while almost all of the rest of error compensation techniques suffer from oscillations.

2. Wolf's Error Compensation Technique 1.

Wolf [W6] introduced a systematic approach to determine a fictitious noise input to the system such that it will compensate for the model errors.

He uses the relation:

$$Q(k) = \alpha^2 \Delta t^2 \ .$$

Where

$Q(k)$ = system fictitious error noise covariance input at time k.

α = a parameter determined by simulation experience.

Δt = sample period for the sample data system.

The experience of applying this technique to the reactor model can be summarized as:

A. The parameter α is unknown, it changes for different system models. It requires prior knowledge of the filter experience to determine a proper value of α to give a good filter performance. Otherwise, too high a fictitious noise will cause filter divergence while too low a noise will not be enough to compensate the model errors.

B. Since we are introducing a constant amount of noise into the system, the filter estimate suffers serious oscillations, and this may not be acceptable in a very sensitive system. The oscillations in reactor temperature are observed as in Fig. 4.13 and Fig. 4.14.

3. Wolf's Error Compensation Technique II.

This technique was also reported by Wolf [W6]. Instead of using a constant amount of fictitious noise, he inputs the noise according to:

$$Q(k) = \Delta\underline{x}(k)\Delta'\underline{x}(k) \quad .$$

Here $\Delta\underline{x}(k)$ is computed by picking a likely parameter error and using the system dynamics equations to actually compute the difference $\Delta\underline{x}(k)$. This method avoids oscillations in the estimate as we expected, since it uses an accurate fictitious noise input.

We use this approach to treat the integration error using a simple integration routine (Euler's integration) and simulate a rigorous integration routine (Runge-Kutta Merson). Numerical experience on our reactor model can be summarized as follows:

A. This method gives a smooth estimate, but the performance of the filter again depends on the prior knowledge of the amount of errors in the parameters. Therefore, this method can only be applied to the situation where $\Delta\underline{x}(k)$ can be evaluated confidently.

B. The filter stability is very sensitive to this type of noise input since Q(k) is usually very small. The filter controllability region is also small and this method therefore, suffers divergence.

4. Pine's Computational Round-off Error Model.

Pine [P2] developed a noise input model according to

$$Q(k) = 10^{-2p} \begin{vmatrix} x_1^2(k+1,k) & & \\ & \ddots & \\ & & x_n^2(K+1,k) \end{vmatrix} \quad .$$

This model was reported by Schlee [P2]. He applied this technique successfully to correct the round off error due to too short a word length where p = 8 for single precision word length. However, p = 6.5 gives better performance.

Round off errors in matrix operations and inversion have been reduced in this work by using the Shur relation [G1] for matrix inversion together with matrix iteration. Pine's method does not show any significant improvement.

5. Overweighting the most recent observation.

Fagin [F1] introduced the technique of exponentially age-weighting the old observation. He replaced the measurement noise covariance

$$R(1),\ R(2),\ldots R(k)$$

by

$$e^{(t_k-t_1)/\tau_{R(1)}},\ e^{(t_k-t_2)/\tau_{R(2)}},\ldots,\ e^{(t_k-t_k)/\tau_{R(k)}}\ ,$$

and this leads to the new gain matrix calculation

$$K(k+1) = P(k+1,k)H'(k+1)[H(k+1)P(k+1,k)H'(k+1) + e^{-(t_{k+1}-t_k)/\tau_{R(k+1)}}]^{-1}\ ,$$

and the recursive relation for error covariance matrix is

$$P(k+1,k+1) = e^{(t_{k+1}-t_k)/\tau}\ [I-K(k+1)H(k+1)]P(k+1,k)\ .$$

Note that the smaller τ is, the faster the old observations are forgotten.

This technique has been applied successfully to the linear filter example where only a scalar observation is used. The state estimate can be adjusted as close as possible to the measurement by using a proper τ. However, vector observations have to be processed on one by one in their components. Schmidt treats the estimate as a linear combination of the estimate based on the current and past observation. The estimate based on the current observation along sounds also promising for systematic overweighting of the current observations, but all of these involve the proper weighting coefficient for the scalar observation and weighting coefficient *vector* for vector observations.

6. Error Compensation Techniques Introduced by This Work.

A. By using the formula found in this work relating the fictitious noise to the estimate error covariance matrix:

$$Q(k) = \Delta t^2 P(k) \quad .$$

Where

Q(k) = fictitious noise error covariance matrix (input

Δt = sample period for the system

P(k) = estimate error covariance matrix.

Very satisfactory filtering performances were observed in compensating various modeling errors:

a. Modeling error due to integration round off errors.

b. Modeling errors due to model parameter uncertainty (as much as 20 per cent deviation in rate constant).

c. Approximating a higher dimensional system by a lower dimensional model.

All of the filtering results show better performance than the other techniques introduced in this section. Yet it does not require any additional computer storage and/or computation time. Another advantage of this method is that it does not require any prior knowledge of the amount of uncertainty of the model errors or constants which are based on filtering experience. Numerical experience with this technique on our reactor model show this method always operates the filter within the controllability and observability region and give consistently improved smooth estimates. The results are shown in Fig. 8.1 to Fig. 8.5.

B. Fixing the Error Covariance Matrix (P matrix) at A Proper Constant Value.

As is observed in this work, the calculated error covariance matrix based on the incorrect model always gives too low an estimate error covariance matrix which in turn will make the Kalman filter gain matrix too low. This action will then lead to insufficient correction of the filter estimate. In other words, the filter is too optimistic and thinks the modeling error is very small, and the filter correction is not enough to compensate the model errors. One way to remedy this drawback is to fix the estimate error covariance matrix at a constant higher value thus raising the gain matrix. By fixing the error covariance at the value which occurs at stage number 3 it shows satisfactory results on different types of model errors.

Utilizing the above idea the new error covariance calculation will be:

$$P(k+1,k+1) = P(3,3) \text{ for } k > 3$$

and substitute this fixed P value into the gain matrix calculation.

C. By Stocastic Approximation.

Another systematic scheme for changing the error covariance matrix, and thus raising the Kalman filter gain is using the Harmonic sequence stocastic approximation.

The harmonic sequence 1, $\frac{1}{2}$, 1/3, $\frac{1}{4}$,... etc. is of central importance in stocastic approximation schemes where operating conditions are continually being adjusted. If the step size is decreased according to the harmonic sequence, the procedure will eventually reach the sought after value, no matter how far away the problem is started. This is because the harmonic sequence is divergent in the sense that the sum of all of its terms is infinite, i.e.

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty \ .$$

The harmonic sequence is the fastest shrinking series of the type n^{-p} that is divergent, that is, the harmonic sequence offers unlimited correction effort if necessary.

Another desirable feature of harmonic sequence is that the sum of its squares is convergent:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty \ .$$

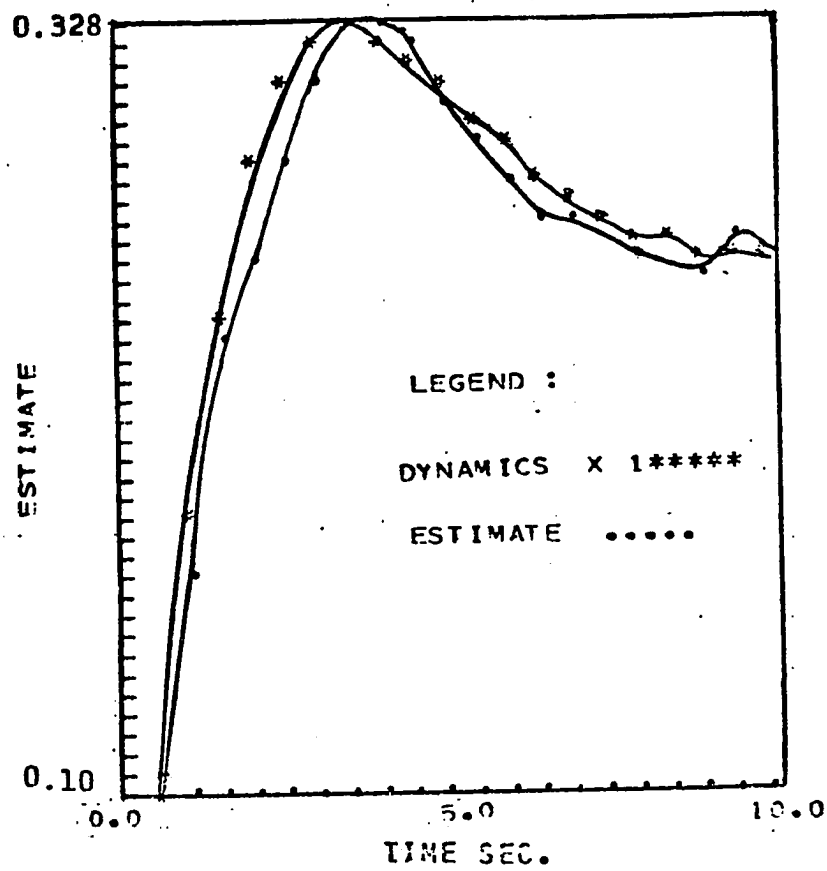
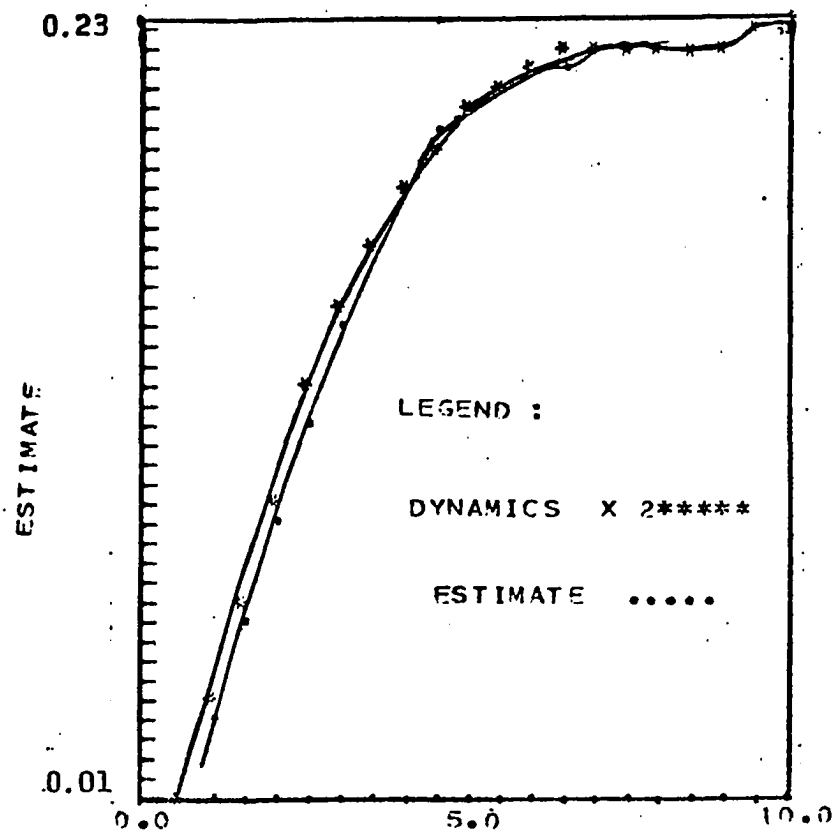


FIG. 4.4 HUANGS ADAPTIVE III
SECOND ORDER NONLINEAR FILTER IS USED
MODEL RATE CONSTANT ERROR = 20.00 PERCENT

TIME SEC.	X 1	ESTIMATE
0.50	0.0988	0.1026
1.00	0.1810	0.1661
1.50	0.2427	0.2347
2.00	0.2845	0.2547
2.50	0.3097	0.2876
3.00	0.3222	0.3099
3.50	0.3255	0.3280
4.00	0.3229	0.3266
4.50	0.3165	0.3243
5.00	0.3083	0.3073
5.50	0.2995	0.2910
6.00	0.2907	0.2837
6.50	0.2827	0.2716
7.00	0.2755	0.2680
7.50	0.2695	0.2682
8.00	0.2645	0.2587
8.50	0.2607	0.2606
9.00	0.2579	0.2525
9.50	0.2559	0.2614
10.00	0.2547	0.2582



20 % rate constant error

FIG. 4.5 HUANGS ADAPTIVE III
SECOND ORDER NONLINEAR FILTER IS USED

TIME SEC.	X 2	ESTIMATE
0.50	0.0118	0.0117
1.00	0.0369	0.0330
1.50	0.0673	0.0631
2.00	0.0980	0.0911
2.50	0.1261	0.1186
3.00	0.1503	0.1441
3.50	0.1700	0.1676
4.00	0.1857	0.1855
4.50	0.1977	0.2001
5.00	0.2067	0.2086
5.50	0.2132	0.2135
6.00	0.2178	0.2178
6.50	0.2210	0.2195
7.00	0.2231	0.2217
7.50	0.2244	0.2243
8.00	0.2252	0.2243
8.50	0.2257	0.2261
9.00	0.2261	0.2255
9.50	0.2264	0.2283
10.00	0.2267	0.2290

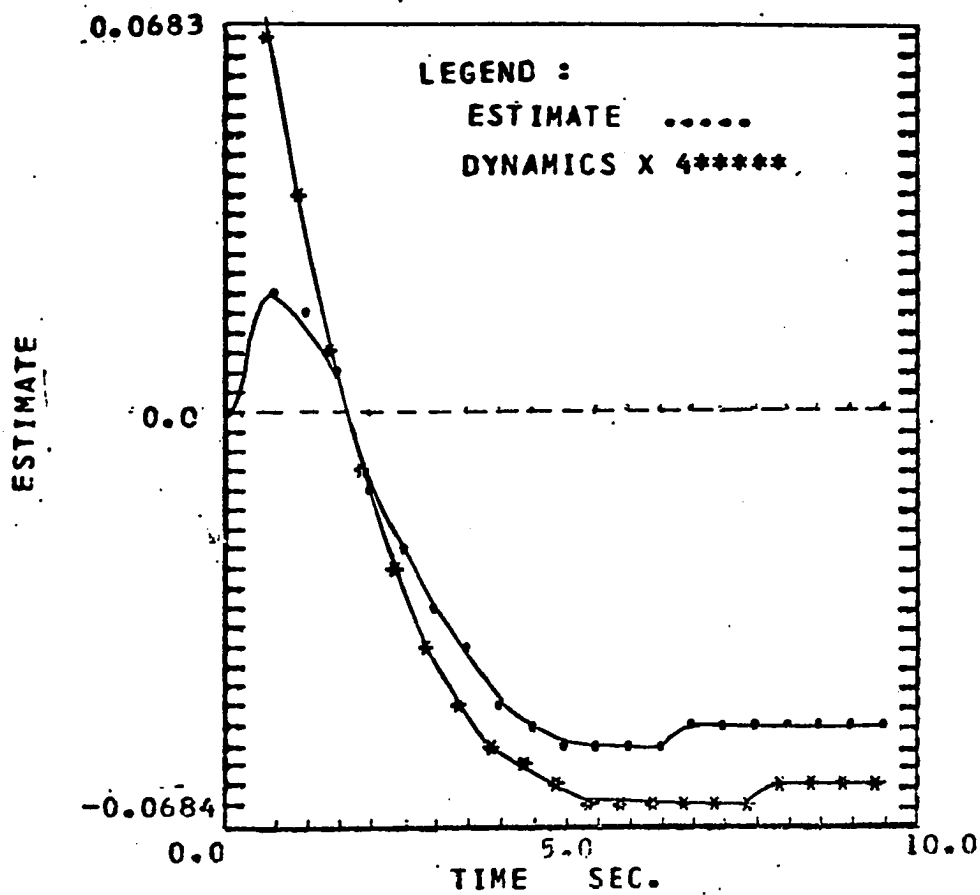


FIG. 4.6 HLANGS ADAPTIVE III
SECCND CRDER NCNLINEAR FILTER
MODEL RATE CCNSTANT ERROR = 20.CO PERCENT

TIME SEC.	X 4	ESTIMATE
0.5	0.0683	0.0217
1.0	C.0377	0.C174
1.5	C.C1C7	0.CC83
2.0	-0.C120	-0.0135
2.5	-C.0301	-0.C255
3.0	-C.C440	-0.0358
3.5	-0.C543	-0.0443
4.0	-C.C616	-C.C533
4.5	-0.0666	-0.0574
5.0	-C.C698	-0.0612
5.5	-C.C716	-0.C635
6.0	-0.0725	-0.C609
6.5	-C.C727	-0.C610
7.0	-0.0724	-0.C586
7.5	-0.0718	-0.0575
8.0	-C.0711	-0.C58C
8.5	-0.0703	-0.C577
9.0	-0.C696	-0.0584
9.5	-C.0689	-0.C567
10.0	-0.0684	-0.0577

This implies that the individual random errors will tend to cancel each other out in the long run. Thus the harmonic sequence offers a general guide to weighting the new and old data subject to random errors. See Wilde [W4], Gardner [G2] for a detailed discussion on harmonic sequences.

By applying properties of harmonic sequence to systematically raising the estimate error covariance matrix in this work:

$$P_{k+1} = P_k + \frac{a_k}{k} (P_k - P_{k+1}) \quad . \qquad \text{IV-1}$$

Where k is the sample period number, a_k is a weighting coefficient, P_k is the error covariance matrix at sample period number k. Satisfactory results were obtained in applying this technique to compensate various model errors in our reactor system.

IV.4 On-line Adaptive Filtering

All of the filtering algorithms assume a complete apriori knowledge of the process and measurement noise statistics to start the filter. However, in practice, these statistics are usually inexactly known. The inexact knowledge of the prior statistics in the design of filter can lead to large estimation error or even filter divergence. This problem becomes more serious in non-linear filtering since the nonlinear model is more sensitive to errors in the prior statistics. The controlability and observability region is much smaller than that of in the linear case, therefore the filter is able to operate only in a very small region. Inexact knowledge of the prior statistics can easily drive the filter outside of the region and cause filter

divergence. A good example of this type of filter divergence can be found in our reactor model. With a process noise covariance matrix Q(k) = 0.0001, this gives very good filtering result. However, if Q(k) is raised to 0.01, this will cause filter divergence. This behavior is due to the sensitive nonlinear reactor model. Another type of filter divergence is usually caused by uncertainty in model physical parameters as described in the previous section.

The purpose of an adaptive filter is to provide on-line reliable estimates of the states and parameters. Therefore, the filter must be able to adjust itself to the uncertain environment based on the measurement observations and making proper correction to the errors due to all of these uncertainties. Most of the work in adaptive filtering have been restricted as to the prior statistics, and try to use various learning algorithms to identify the true process noise and measurement noise covariance matrix. Sage-Mehra [S2] [M3] wrote several reviews of adaptive filteirng and introduced their own adaptive filters. All of their techniques are restricted to the identification of the Q and R matrices. These techniques all suffer the disadvantages of many computations and require much additional storage. Sage [S3] used a fixed interval smoothing algorithm as an adaptive technique, which requires an additional matrix inverse and many matrix operations. Mehra [M3] used a correlation method from the time series correlation (Box) [B4] by applying the 'innovation properties of the observation residue'. His method also requires additional storage and a matrix inverse and can only be applied to linear time invariant systems.

In this section we only introduce those adaptive techniques suited for real time applications, i.e. limited computer memory and speed. We are not only aimed at solving the problem of errors in the prior statistics but also study the case with uncertain parameters.

Schmidt's State Augmentation

The behavior of a dynamical system depends on certain parameters which may be constant or vary with time. Very often, some of these parameters are unknown or imprecisely known. Such parameters may be considered as random variables with known apriori statistics.

Given a linear system

$$\underline{x}(k+1) = \Phi(k+1,k)\underline{x}(k) + \Psi(k+1,k)\underline{u} + \Gamma(k)\underline{w}(k+1)$$
$$\underline{y}(k) = M(k)\underline{x}(k) + N(k)\underline{p} + \underline{v}(k) \quad . \qquad \text{IV-2}$$

Where

$\underline{u}$ = dynamical parameter vectors

and

$\underline{p}$ = measurement parameter vectors

$M(k)$ = mxn measurement matrix

$N(k)$ = measurement parameter gain matrix

$\Psi(k+1,k)$ = dynamics parameter gain matrix

$y(k)$ = mxl observation vector

The following assumptions are necessary:

$$E(\underline{u}) = 0; \quad E(\underline{p}) = 0; \quad E(\underline{u}\underline{u}') = Uo; \quad E(\underline{p}\underline{p}') = Wo \quad .$$

Further,

$$\underline{p},\ \underline{u},\ \underline{x},\quad \underline{w}(k),\ \underline{v}(k) \text{ are uncorrelated.}$$

Define the augmented state space as follows:

$$\bar{\underline{x}}(k) = \begin{vmatrix} \underline{x}(k) \\ \underline{u}(k) \\ \underline{p}(k) \end{vmatrix} \qquad \text{where} \qquad \begin{aligned} \underline{u}(o) &= \underline{u} \\ \underline{p}(o) &= \underline{p} \end{aligned}$$

The augmented system can be written as:

$$\bar{\underline{x}}(k+1) = \begin{vmatrix} \Phi(k+1,k) & \Psi(k+1,k) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \bar{\underline{x}}(k) + \begin{vmatrix} \Gamma(k) \\ 0 \\ 0 \end{vmatrix} \underline{w}(k+1) \qquad \text{IV-4}$$

with observations:

$$y(k) = [M(k)\ 0\ N(k)]\bar{x}(k) + v(k) \quad . \qquad \text{IV-5}$$

The linear filter is directly applicable to this augmented system with the estimate:

$$\hat{\bar{\underline{x}}}(k,k) = \begin{vmatrix} \hat{\underline{x}}(k,k) \\ \hat{\underline{u}}(k,k) \\ \hat{\underline{p}}(k,k) \end{vmatrix} .$$

Therefore, we can estimate the original states together with the uncertain parameters in a recursive manner. However, it often creates computer storage and computational load problems. This is especially serious for a large system and it is not justified or feasible to treat the problem in this way. In the meantime, we are aware of the fact that the ignorance of the uncertainty of these parameters often causes filter divergence. This is especially true in the reactor

system which is very sensitive to parameter variations (chemical kinetics). Different parameters can lead to different steady states or even to runaway reactions.

An alternative, suggested by Schmidt, is to take into account the effect of the uncertain parameters in degrading the state estimate without really estimating the parameters. This can be done by modifying the error covariance and filter gain matrix in the following manner and will save computer storage in the filter estimate of the parameters.

The error covariance matrix for the augmented system is

$$\bar{P} = \begin{vmatrix} P & C_u & C_p \\ C_u' & U_o & 0 \\ C_p' & 0 & W_o \end{vmatrix}$$

Where P: error covariance of original state

$$C_u = E[(\underline{x}-\hat{\underline{x}})\underline{u}']$$

$$C_p = E[(\underline{x}-\hat{\underline{x}})\underline{p}']$$

For the prediction stage we only estimate the original state:

$$\hat{\underline{x}}(k+1,k) = \Phi(k+1,k)\hat{\underline{x}}(k,k) \quad , \qquad \text{IV-6}$$

and $\underline{u}$, $\underline{p}$ are invariant in prediction. The modified error covariance matrix for the augmented system is:

$$\bar{P} = \begin{vmatrix} \Phi(k+1,k) & \Psi(k+1,k) & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{vmatrix} \begin{vmatrix} P(k,k) & C_u(k,k) & C_p(k,k) \\ C_u'(k,k) & U_o & 0 \\ C_p'(k,k) & 0 & W_o \end{vmatrix}$$

$$\begin{vmatrix} \Phi(k+1,k) & 0 & 0 \\ \Psi'(k+1,k) & I & 0 \\ 0 & 0 & I \end{vmatrix} \begin{vmatrix} \Gamma(k)Q(k+1)\Gamma'(k) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} \qquad \text{IV-7}$$

We can then arrive at the following recursive algorithm by carrying the matrix product component wise:

$$P(k+1,k) = \Phi(k+1,k)P(k,k)\Phi'(k+1,k) + \Phi(k+1,k)C_u(k,k)$$
$$\Psi'(k+1,k) + \Psi(k+1,k)C_u'(k,k)\Phi'(k+1,k) + \Psi(k+1,k)$$
$$U_o\Psi'(k+1,k) + \Gamma(k)Q(k+1)\Gamma'(k) \qquad \text{IV-8}$$

$$C_u(k+1,k) = \Phi(k+1,k)C_u(k,k) + \Psi(k+1,k)U_o \qquad \text{IV-9}$$

$$C_p(k+1,k) = \Phi(k+1,k)C_p(k,k) \ . \qquad \text{IV-10}$$

Let $\bar{M}(k)$ be the augmented measurement matrix:

$$\bar{M}(k) = [M(k)\ 0\ N(k)] \ , \qquad \text{IV-11}$$

and

$$Y(k+1,k) = \bar{M}(k+1)\bar{P}(k+1,k)\bar{M}'(k+1) + R(k+1)$$
$$= M(k+1)P(k+1,k)M'(k+1) + M(k+1)C_p(k+1,k)N'(k+1)$$
$$+ N(k+1)C_p(k+1,k)M'(k+1) + N(k+1)W_oN'(k+1) + R(k+1) \qquad \text{IV-12}$$

Process the observation and corrections:

$$\underline{\hat{x}}(k+1,k+1) = \underline{\hat{x}}(k+1,k) + \bar{K}(k+1)[\underline{y}(k+1) - M(k+1)\underline{\hat{x}}(k+1,k)] \ . \qquad \text{IV-13}$$

Where $\bar{K}(k+1)$ is the modified Kalman filter gain

$$\bar{K}(k+1) = [P(K+1,k)M'(k+1) + C_p(k+1,k)N'(k+1)]Y^{-1}(k+1,k) \ . \qquad \text{IV-14}$$

The new error covariance matrix for the augmented system is:

$$P(k+1,k+1) = P(k+1,k) - \bar{K}(k+1)[M(k+1)P(k+1,k) + N(k+1)C_p'(k+1,k)] \qquad \text{IV-15}$$

$$C_u(k+1,k+1) = C(k+1,k) - \bar{K}(k+1)M(k+1)C_u(k+1,k) \qquad \text{IV-16}$$

$$C_p(k+1,k+1) = C_p(k+1,k) - \bar{K}(k+1)[M(k+1)C_p(k+1,k)+N(k+1)W_o]. \qquad \text{IV-17}$$

With the initial conditions:

$$C_u(0,0) = C_p(0,0) = 0 \quad .$$

It is obvious that the computational load and storage is considerably less than that required the augmented state filter.

In applying the Schmidt filter to our reactor system assuming we are having a 20 per cent error in the reduced rate constant K_1 we have: the linearized system as given by

$$\underline{\hat{x}}(k+1) = \Phi(k+1,k)\underline{\hat{x}}(k) + \Psi(k+1,k)\underline{u} + \Gamma(k+1)\underline{w}(k+1) \quad .$$

$\Phi(k+1,k)$ is the transition matrix which is defined as the Jacobian in Chapter II.

$\Psi(k+1,k)$ is the Jacobian of the parameter matrix

$$\Psi(k+1,k) = \frac{\partial f}{\partial u} \quad .$$

In this particular example with error in the rate constant,

$$\Psi(k+1,k) = \frac{\partial \underline{f}}{\partial K_1} \begin{vmatrix} \partial f_1/\partial K_1 \\ 0 \\ 0 \\ \partial f_4/\partial K_1 \end{vmatrix}$$

Where

$$\partial f_1/\partial K_1 = - c_3K_1B^2E - K_1c_3 + c_3B^2Ex_1/(1+x_1)$$

$$\partial f_4/\partial K_1 = - c_2B^2Ex_1/(1+x_1) + c_2B^2EK_1$$

and the parameter covariance matrix is $U_o = 0.04$.

Use the same prior statistics as in Chapter II, we are able to carry out the sequential filtering on both the extended Kalman filter and the second order filter. The results on Fig. 4.9 show very good temperature estimate with a slight oscillation.

Fixed interval smoothing and fixed point smoothing.

The smoothing problem deals with the system's states which are of the form

$$\underline{x}(k|j) = \underline{f}_k[\underline{z}(i), \quad i = 1,\ldots,j]$$

where $j > k$; that is, the time at which it is desired to estimate the state lies to the left of the time of the last measurement, this is equivalent to interpolation of the estimate based on the present and past filtered estimate. We know that the optimal smoothed estimate is unique since it is related to the filter estimate. Our purpose in this section is to develop a recursive smoothing algorithm which permits us to apply it to real time environment.

Smoothing can be used to improve the filtering estimate at the cost of additional computer storage and compution time; it is still useful in the post experiment analysis. With slight manipulation shown by Meditch [M2], this technique can be used for on-line adaptive filtering. Jazwinski's [J2] limited memory filter is closely related to the idea of smoothing. Sage [S2] used fixed interval smoothing algorithm as an adaptive filtering technique.

It has been shown by Meditch [M2] that the following three different types of smoothing algorithms are equivalent.

1. Fixed interval smoothing.

Consider the operation of a chemical reactor system; since it is impossible to design a perfect model, we cannot expect the reactor dynamics to follow exactly the desired profile during the operating period. Sometimes, the discrepancy between the design and operation is intolerable, which may cause a runaway reaction or may lead to an undesirable product. Therefore, it is desirable to make the correction later during the reactor operation. Let us suppose that during the reactor operation, we obtain measurement data at N time points $K = 1,2,...N$ using an optimal filtering technique discussed in Chapter II. We then have the optimal filtering estimate $x(k,k)$, $k = 1,2,...N$. The idea of fixed interval smoothing is to refine this set of filtered estimate utilizing additional available measurement data. For example, with $N = 200$, $K = 51$, the question is whether the optimal smoothered estimate of $x(51)$ based on 200 measurements a better estimate at that time, than the filtered estimate $x(51,51)$. Intuitively, it should be the case. Therefore, with the experiment completed in the interval $(0,N)$, for each time k within the interval, we wish to obtain the optimal estimate based on all the data in the interval. Fixed interval smoothing cannot be carried out online during the experiment, but it is used for post experiment refinement of the estimate.

2. Fixed point smoothing.

Returning to our chemical reactor system where we are interested in the operating cycle to keep the dynamics at a desired state at a particular time. Let us assume that the reactor has been operating

for some time. We wish to use additional data to obtain and improve the estimate $\hat{x}(N)$ where N is the time of particular interest. Up to the time N we have the filtered estimate $\hat{x}(N,N)$ and what we like to do now is to determine $\hat{x}(N,N+1)$, $\hat{x}(N,N+2)$, etc. Here, we are concerned with an estimate at a fixed point in time which is based not only on measurements up to that time, but the additional measurement taken beyond it.

The algorithm developed by Meditch is suitable for online applications. This algorithm can be easily applied to our reactor system in order to determine a better estimate of our initial reactant concentration and temperatures. The same algorithm is also used as an adaptive technique as shown in Fig. 8.8. We present the algorithm as follows:

a. The optimal fixed point smoothed estimate:

$$\underline{\hat{x}}(k,j) = \underline{\hat{x}}(k,j-1) + W(j)H'(j)R^{-1}(j)[\underline{z}(j)-H(j)\Phi(j,j-1)\underline{\hat{x}}(j-1,j-1)] \qquad \text{IV-18}$$

for

$$j = k+1,k+2...$$

with the initial condition $\underline{\hat{x}}(k,k)$ given.

b. W(j) is the nxn smoothing filter gain matrix which is calculated by the recursive relation:

$$W(j) = W(j-1)\Phi'(j,j-1)[I-S(j)P(j,j)] \quad , \qquad \text{IV-19}$$

for

$$j = k+1,k+2...$$

with the initial condition $W(k) = P(k,k)$ and $S(j) = H'(j)R^{-1}(j)H(j)$.

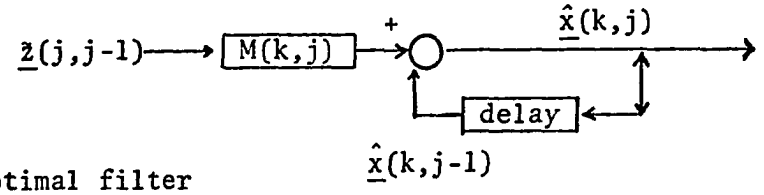
c. The estimate error covariance matrix satisfies:

$$P(k,j) = P(k,j-1) - W(j)[S(j)P(j,j-1)S(j)+S(j)]W'(j) \qquad \text{IV-20}$$

for

$$j = k+1,k+2,...$$

with the initial condition P(k,k)

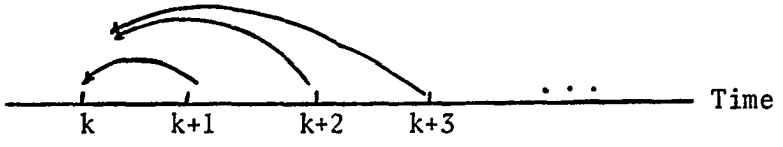


from optimal filter

$$j = k+1,k+2,...$$

$$\tilde{\underline{z}}(j,j-1) = \underline{z}(j) - H\hat{\underline{x}}(j,j-1) \qquad \text{IV-21}$$

$$M(k,j) = W(j)H'(j)R^{-1}(j)$$



Limited memory filter.

The filter developed in Chapter II utilizes all the available observations. It is optimal only if the filter is operated with complete knowledge of the dynamics, measurement function and the statistics of the dynamic system. If the dynamics are imprecisely known, then the filter might learn the wrong model too well and cause filter divergence as described in the previous section. A reasonable approach to avoid filter divergence is to limit the filter memory so that the

estimate does not become 'too good', or to keep the Kalman filter gain from getting too small thus making too small a correction. By limiting the filter memory, we mean computing an estimate based on observations from only the recent past and discarding the estimate from the distant past. Jazwinski [J2] developed the theory of limited memory filter and applied it successfully to the rectilinear orbit problem.

The limited memory filter estimate generates the optimal estimate based on a 'moving window' of the most recent N observations. Two Kalman filters and a predictor are required in the implementation and N observations have to be stored in memory. Three matrix inversions are required. Jazwinski applied the idea of discarding the old data in batches of N, therefore, he filters the N observations and predicts over the same time arc, thus reducing some of the computer storage and computation time.

Even with Jazwinski's modification, the limited memory filter has not been attractive enough to encourage wide application. On the other hand, the smoothing algorithms presented in the previous section are more promising.

This work.

All of the three types of error compensation techniques developed in this work can be used as adaptive filtering, since they can be used in an on-line filtering environment. The filter can adjust by itself to compensate for any model errors, or incomplete knowledge of statistical parameters. A clear advantage offered by these adaptive offered by these adaptive filters is that it does not require any

additional storage or computational time and yet are reliable enough to be applied to adaptive situation.

Two additional adaptive filtering techniques developed by this work as described follows are even more powerful in removing model and prior estimate uncertainties and play an important role in model approximation and model decomposition as in the filtering studies for large systems.

A. Iterated Fixed Point Smoothing.

This algorithm has been developed in this work and successfully applied to approximate the 4th order reactor model by a second order. In this algorithm, fixed point smoothing is applied whenever the filter estimate residue (the difference between the filter estimate and the measurement data) is greater than one standard deviation, thus the model uncertainties or inaccurate initial estimate is removed by smoothing through processing additional measurement data. The sequential nature of the fixed point smoothing algorithm allows us applying it easily online and only requires slightly additional storage and computational time.

The results shown on the next page indicates the fast rate of disappearance of the 10 per cent error in x4 in the estimate, applying one smoothing can remove most of the uncertainty in x4, while two smoothing can put the estimate coincide with the true dynamics in the approximation of our 4th order reactor model by a second order model in the presence of 10 per cent error in initial estimate of x4.

Refer to the appendix for detailed filtering results for this particular example.

Table IV-1. Approximate the 4th order model by a second order using Huang's II (iterated fixed point smoothing) sample output for smoothing the initial condition uncertanties.

True Dynamics	Without Smoothing Estimate	First Smoothing Iteration Estimate	Second Smoothing Iteration Estimate
Initial Conditions:			
X1 : 0.0	0.0	0.0244	-0.0017
X2 : 0.1	0.0	0.0733	0.104
Period No. 1			
X1 : 0.10026	0.0764	0.1064	0.107
X2 : 0.06832	0.0192	0.0483	0.072
Period No. 2			
X1 : 0.1842	0.1627	0.1822	0.038
X2 : 0.0373	0.0172	0.0237	0.038
Period No. 3			
X1 : 0.2475	0.2192	0.229	0.2449
X2 :: 0.0096	-0.00065	-0.0004	0.0064
Period No. 4			
X1 : 0.2901	0.2674	0.2707	0.2825
X2 : -0.0135	-0.01675	-0.0184	-0.0164

B. Overweight the most recent measurement by stochastic approximation.

Fagin's error compensation technique by exponentially age-weighting the old data requires a prior knowledge of the weighting factor which varies in every filtering in the time varying system. Improper values of the weight factor can easily lead to filter divergence in non-linear filter and often requires extensive simulation and filtering experience before a suitable factor can be found.

Schmidt presents the following systematic overweighting the data by computing an estimate as a linear combination of the estimate based on the current observation and past data (Kalman estimate) and the estimate based on the current observation alone. Let:

$$\Delta\underline{x} = \underline{x}(k+1) - \hat{\underline{x}}(k+1,k)$$

$$\Delta\hat{\underline{x}} = \hat{\underline{x}}(k+1,k+1) - \hat{\underline{x}}(k+1,k)$$

$$\Delta\underline{z} = \underline{z}(k+1) - H(k+1)\hat{\underline{x}}(k+1,k) = H(k+1)\Delta\underline{x} + \underline{v}(k+1) .$$

The estimate based on the current observation and past data is the usual Kalman estimate:

$$\Delta\hat{\underline{x}} = P(k+1,k)H'(k+1)[H(k+1)P(k+1,k)H'(k+1) + R(k+1)]^{-1}\Delta\underline{z} .$$

If the observation is scalar, then, the quantity in the bracket is also scalar, then the least square estimate is

$$\Delta\hat{\underline{x}} = [H'(k+1)R^{-1}(k+1)H(k+1)]\# H'(k+1)R^{-1}(k+1)\Delta z . \qquad \text{IV-22}$$

Where

A# is the pseudo inverse of A.

For scalar observation, H is a row matrix therefore,

$$H^{\#} = H'/H\,H' \quad .$$

So that IV-22 becomes:

$$\Delta\underline{\hat{x}} = \frac{H'(k+1)}{H(k+1)H'(k+1}\,\Delta\underline{z}$$

or

$$\Delta\underline{\hat{x}} = cH'(k+1)R(k+1)[H(k+1)H'(k+1)]^{-1} X[H(k+1)P(k+1,k)H'(k+1)+R(k+1)]^{-1}\Delta\underline{z} \qquad \text{IV-23}$$

with the appropriate constant c, taking our estimate to be a linear combination of the Kalman estimate and Equation IV-23 we have the estimator:

$$\begin{aligned}\underline{\hat{x}}(k+1,k+1) = \underline{\hat{x}}(k+1,k) + [&P(k+1,k)H'(k+1)\\ &+cH'(k+1)R(k+1)[H(k+1)H'(k+1)]^{-1}\\ &X[H(k+1)P(k+1,k)H'(k+1)+R(k+1)]^{-1}\\ &X[z(k+1)-H(k+1)\underline{\hat{x}}(k+1,k)] \qquad \text{IV-24}\end{aligned}$$

or

$$H(k+1)\underline{\hat{x}}(k+1,k+1) = H(k+1)\underline{\hat{x}}(k+1,k) + \frac{[H(k+1)P(k+1,k)H'(k+1)+cR(k+1)]}{[H(k+1)P(k+1,k)H'(k+1)+R(k+1)]}\,\Delta\underline{z} \quad .$$

Thus with c = 1 we have:

$$H(k+1)\underline{\hat{x}}(k+1,k+1) = H(k+1)\underline{\hat{x}}(k+1,k) + \Delta\underline{z} = \underline{z}(k+1)$$

thus the estimate is the observation itself. With c = 0, Equation IV-24 reduces to the Kalman filter estimate.

The recursive relation for the error covariance matrix is:

$$P(k+1,k+1) = P(k+1,k) - P(k+1,k)H'(k+1)$$
$$X[H(k+1)P(k+1,k)H'(k+1)+R(k+1)]^{-1}H(k+1)P(k+1,k)$$
$$+ \frac{c^2R^2(k+1)[H'(k+1)H(k+1)]}{[H(k+1)P(k+1,k)H'(k+1)+R(k+1)][H(k+1)H'(k+1)]^2} .$$

Note that the term involving c^2 produces a desirable increase in P. However, the appropriate value of c is not known and must be determined by simulation experience and different values of c is required for each observation type.

In our algorithm, we apply stochastic approximation to generate the sequence of value of c, thus c is equal to the values in harmonic sequence i.e. $c_k = 1/k$ and the sequence is restarted whenever the prediction residue is greater than that of its standard deviation. Thus providing overweight on the most recent data only the prediction is bad.

Successful application of this algorithm to the approximation of our 4th order reactor model by a second order model proved to be a powerful adaptive filtering technique as shown in Fig. 4.7 and Fig. 4.8.

IV.5 Numerical Examples and Discussions

The filtering result in compensating 20 per cent rate constant error can be found in Fig. 4.1 to Fig. 4.15. Figures 4.1, 4.2, and 4.3 show the filter divergence which occurs in X1, X2 and X4 without model error compensation while Fig. 4.4 shows that performance given by the compensation techniques introduced in this work is better than

that given by the Schmidt filter shown in Fig. 4.9. Figure 4.10 shows Wolf's II used in the estimation of reactor temperature. Figure 4.6 shows the improved estimate in reactant concentration given by this work. Slight offset after the fifth period is observed in this work III, as shown in Fig. 4.6. Figure 4.12 shows offset by constant noise compensation, while Fig. 4.14 was Wolf's II. Even with Schmidt's filter, state augumentation fails to eliminate the offset. This is probably due to the nongaussian property of X4 under these conditions (20 per cent model error) which leads to the undesired steady state solution of the Riccati equation of the error covariance matrix. This offset disappears as the model error is reduced to less than 10 per cent.

However, using the vector overweight both the reactor temperature and concentration measurement data by stocastic approximation introduced by this work show excellent filtering performance and eliminate the offset.

After an extensive study on model error compensation and adaptive filtering techniques and with applications to our reactor model, we are more confident in implementing either off-line filtering analysis or on-line filtering for any general non-linear filtering problem. A table of summary of our numerical experiences on the reactor will provide a useful guide for future filtering practice.

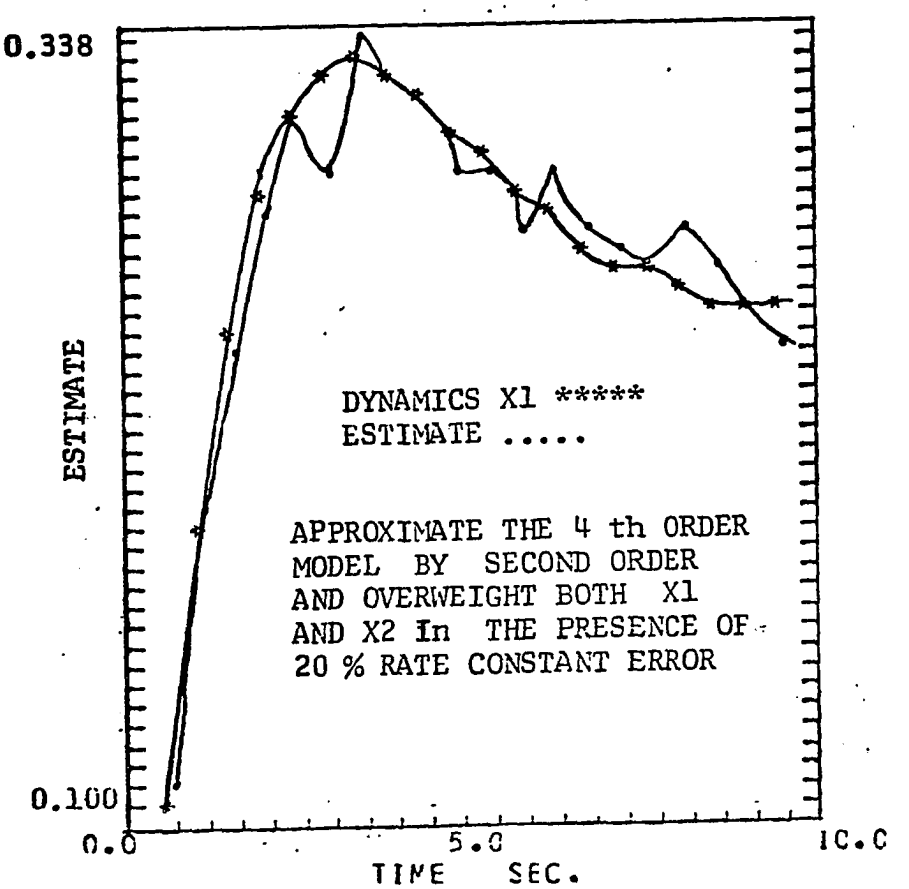


FIG. 4.7 OVERWEIGHT BOTH X1 AND X2
SECCND CRDER NCNLINEAR FILTER AND HUANG'S IV

TIME SEC.	X 1	ESTIMATE
C.5	C.1CC2	C.1C7C
1.0	C.1842	C.1898
1.5	C.2475	C.2434
2.C	C.29C2	C.2836
2.5	0.3153	C.3140
3.0	C.327C	C.2970
3.5	C.3293	C.3381
4.0	0.3254	C.3266
4.5	C.3179	C.3183
5.C	C.3C87	C.2946
5.5	C.2989	C.2979
6.0	C.2895	C.2745
6.5	C.28C9	C.2930
7.0	0.2734	C.2761
7.5	C.2671	C.2685
8.0	C.2621	C.2644
8.5	0.2583	C.2743
9.C	C.2556	C.2660
9.5	C.2538	C.2555
1C.0	0.2528	C.2424

ACCUMULATED MEAN SQUARE ERRCR = C.1C90E-03

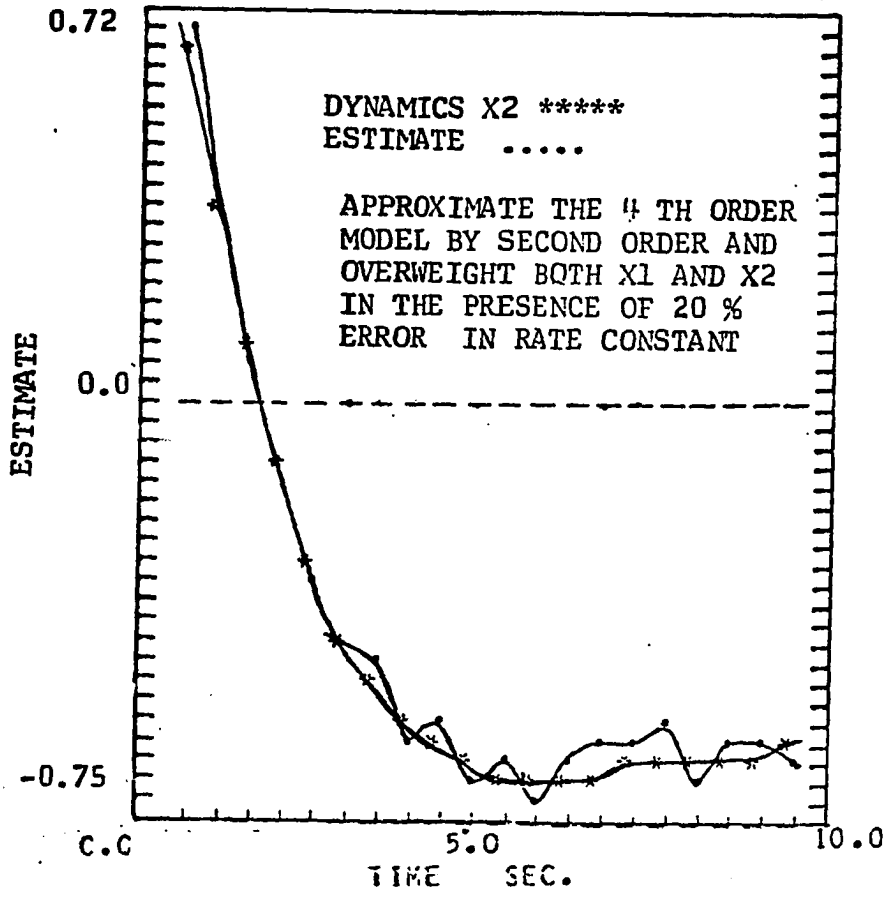


FIG. 4.8 OVERWEIGHT BOTH X1 AND X2
SECCND CRDER NCNLINEAR FILTER AND HUANG'S IV

TIME SEC.	X 2	ESTIMATE
C.5	C.C683	C.C727
1.C	C.C373	C.C376
1.5	0.CC97	C.CC99
2.0	-C.C135	-C.C126
2.5	-C.C319	-C.C34C
3.C	-C.C459	-C.C46C
3.5	-C.C562	-C.C528
4.C	-C.C634	-C.C655
4.5	-C.C682	-C.C63C
5.0	-C.C712	-0.C755
5.5	-C.C727	-C.C697
6.0	-C.0733	-C.C78C
6.5	-C.C732	-C.C694
7.0	-C.C727	-C.C683
7.5	-C.C719	-0.C666
8.C	-C.C71C	-C.C643
8.5	-C.C7C2	-C.C730
9.C	-0.C693	-C.C651
9.5	-C.C686	-C.C677
10.0	-C.C68C	-C.C7C6

ACCUMULATED MEAN SQUARE ERRCR = 0.1290E-04

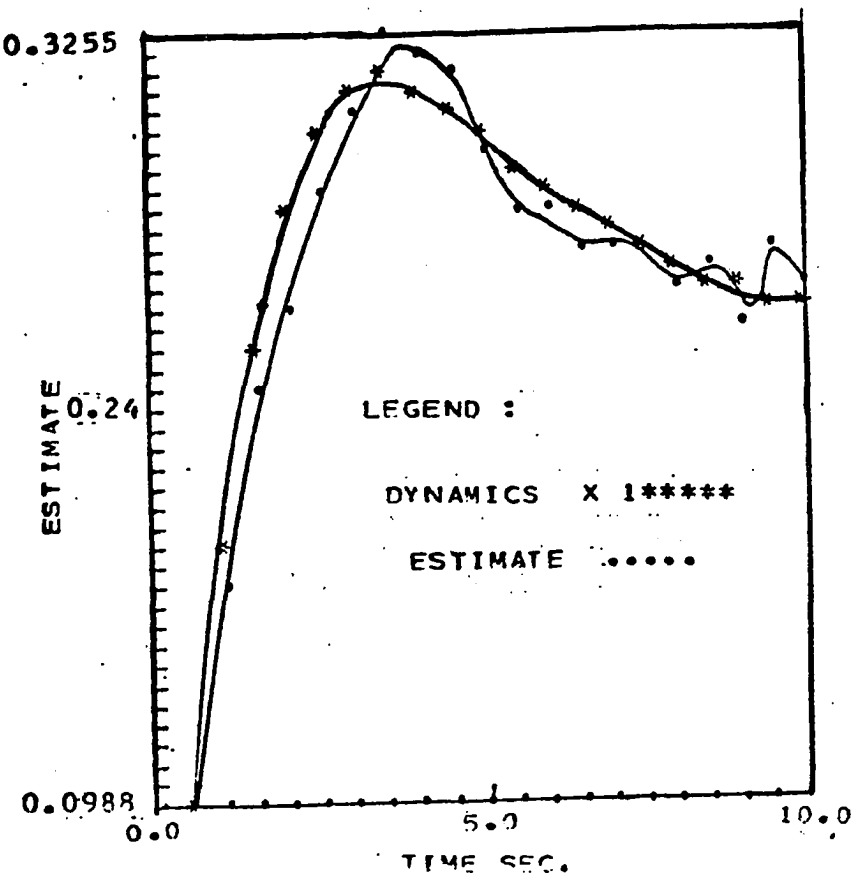


FIG. 4.9 SCHMIDT FILTER
SECOND ORDER NONLINEAR FILTER IS USED
MODEL RATE CONSTANT ERROR = 20.00 PERCENT

TIME SEC.	X 1	ESTIMATE
0.50	0.0988	0.1026
1.00	0.1810	0.1659
1.50	0.2427	0.2319
2.00	0.2845	0.2541
2.50	0.3097	0.2900
3.00	0.3222	0.3164
3.50	0.3255	0.3402
4.00	0.3228	0.3324
4.50	0.3165	0.3308
5.00	0.3083	0.3044
5.50	0.2995	0.2858
6.00	0.2907	0.2831
6.50	0.2827	0.2719
7.00	0.2755	0.2709
7.50	0.2695	0.2741
8.00	0.2645	0.2590
8.50	0.2607	0.2653
9.00	0.2578	0.2494
9.50	0.2559	0.2700
10.00	0.2547	0.2585

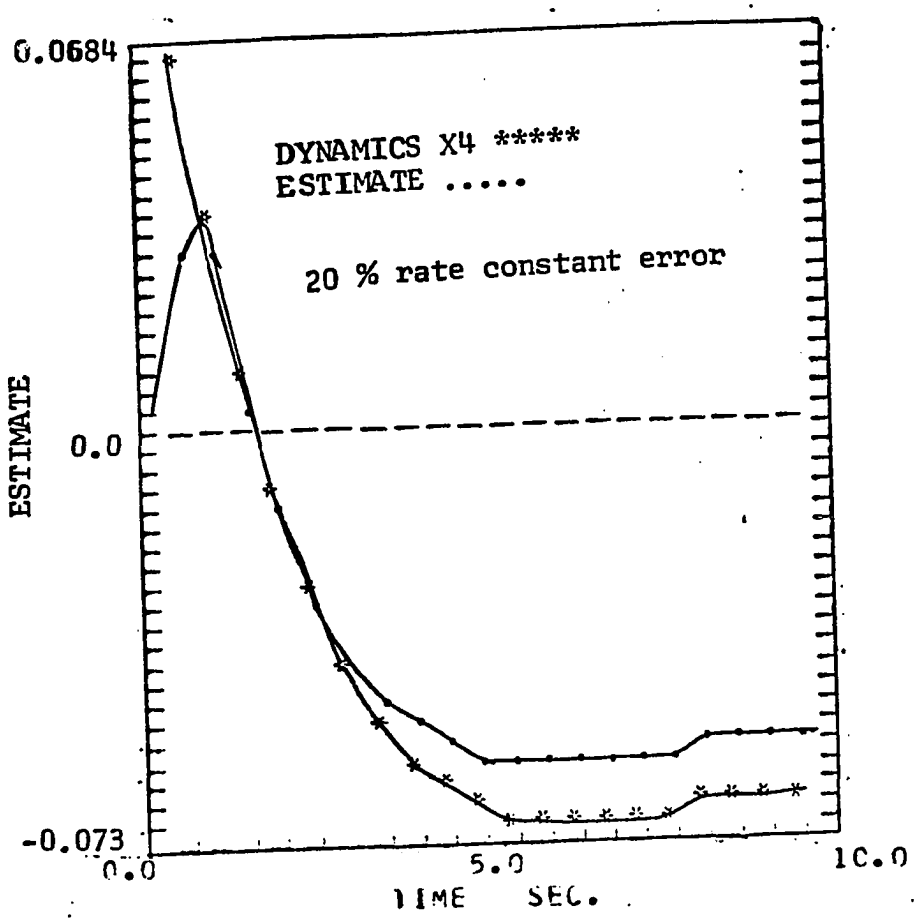


FIG. 4.10 MODEL ERROR COMPENSATION
SECUND CRDER NCNLINEAR FILTER
WCLFS ACAPTIVE II

TIME SEC.	X 4	ESTIMATE
0.5	0.0683	0.0392
1.0	0.0377	0.0196
1.5	0.0107	0.0318
2.0	-0.0120	-0.0036
2.5	-0.0301	-0.0230
3.0	-0.0440	-0.0376
3.5	-0.0543	-0.0477
4.0	-0.0616	-0.0545
4.5	-0.0666	-0.0589
5.0	-0.0698	-0.0613
5.5	-0.0716	-0.0624
6.0	-0.0725	-0.0632
6.5	-0.0727	-0.0626
7.0	-0.0724	-0.0624
7.5	-0.0718	-0.0621
8.0	-0.0711	-0.0606
8.5	-0.0703	-0.0602
9.0	-0.0696	-0.0591
9.5	-0.0689	-0.0593
10.0	-0.0684	-0.0586

ACCUMULATED MEAN SQUARE ERROR = 0.1493E-03

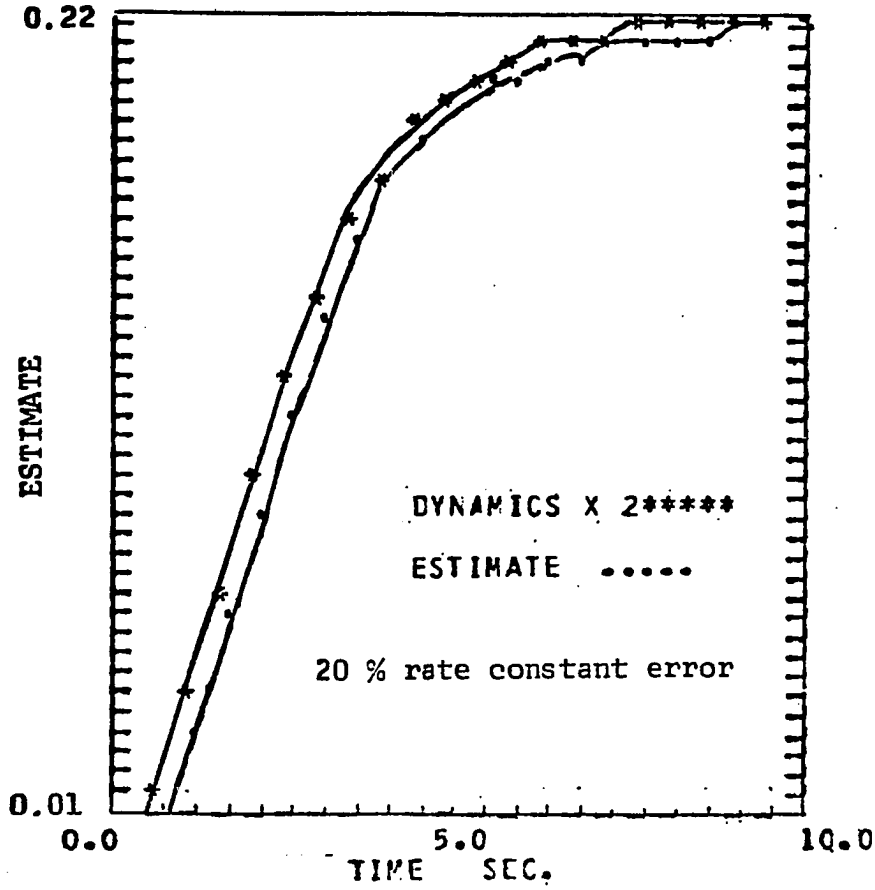


FIG. 4.11 FILTER DIVERGENCE
SECCND CRDER NCNLINEAR FILTER
CONSTANT NOISE

TIME SEC.	X 2	ESTIMATE
C.5	C.C118	0.C115
1.0	0.C369	C.C3C8
1.5	0.C673	0.C596
2.C	0.C980	0.C882
2.5	C.1261	0.1168
3.0	0.15C3	0.1428
3.5	C.17C0	0.1654
4.0	0.1857	0.1820
4.5	0.1977	0.1957
5.C	C.2067	C.2C45
5.5	0.2132	C.2C85
6.0	0.2178	0.2135
6.5	C.2210	0.2148
7.0	0.2231	C.2176
7.5	0.2244	C.2205
8.0	C.2252	0.2218
8.5	0.2257	0.2238
9.0	0.2261	0.2239
9.5	C.2264	C.2269
10.0	0.2267	C.2278

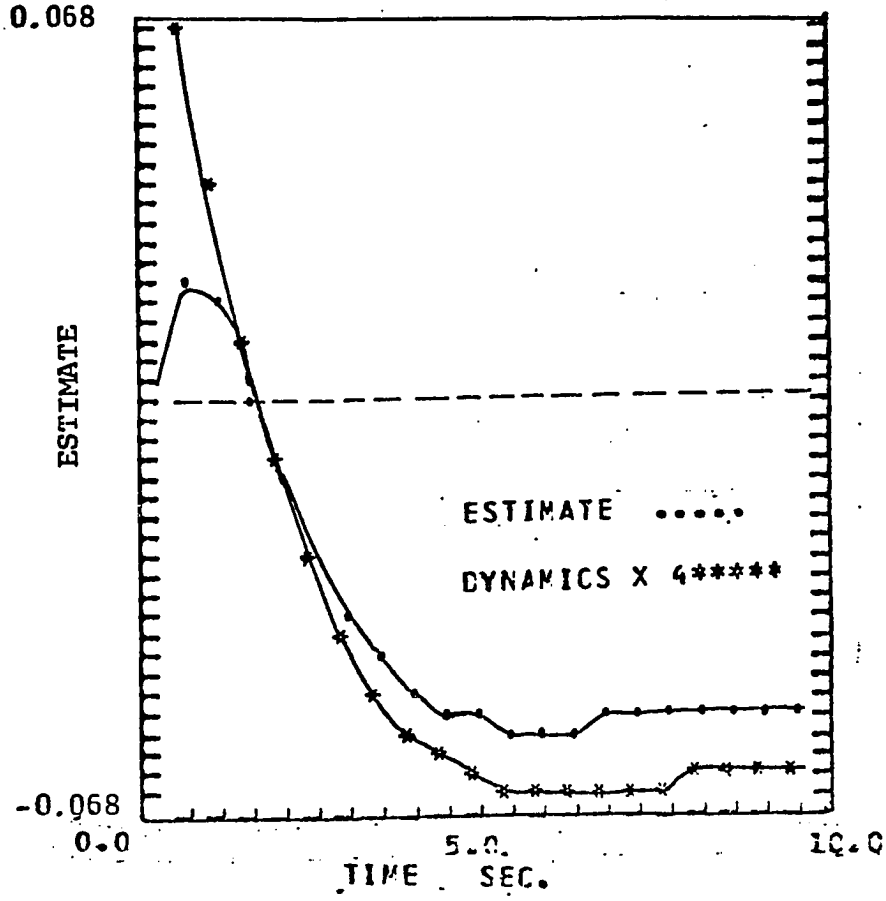


FIG. 4.12 FILTER DIVERGENCE
SECOND ORDER NONLINEAR FILTER
CONSTANT NOISE 20 % rate constant error

TIME SEC.	X 4	ESTIMATE
0.5	0.0683	0.0217
1.0	0.0377	0.0174
1.5	0.0107	0.0027
2.0	-0.0120	-0.0149
2.5	-0.0301	-0.0282
3.0	-0.0440	-0.0389
3.5	-0.0543	-0.0472
4.0	-0.0616	-0.0534
4.5	-0.0666	-0.0573
5.0	-0.0698	-0.0598
5.5	-0.0716	-0.0609
6.0	-0.0725	-0.0607
6.5	-0.0727	-0.0604
7.0	-0.0724	-0.0595
7.5	-0.0718	-0.0587
8.0	-0.0711	-0.0583
8.5	-0.0703	-0.0578
9.0	-0.0696	-0.0575
9.5	-0.0689	-0.0570
10.0	-0.0684	-0.0571

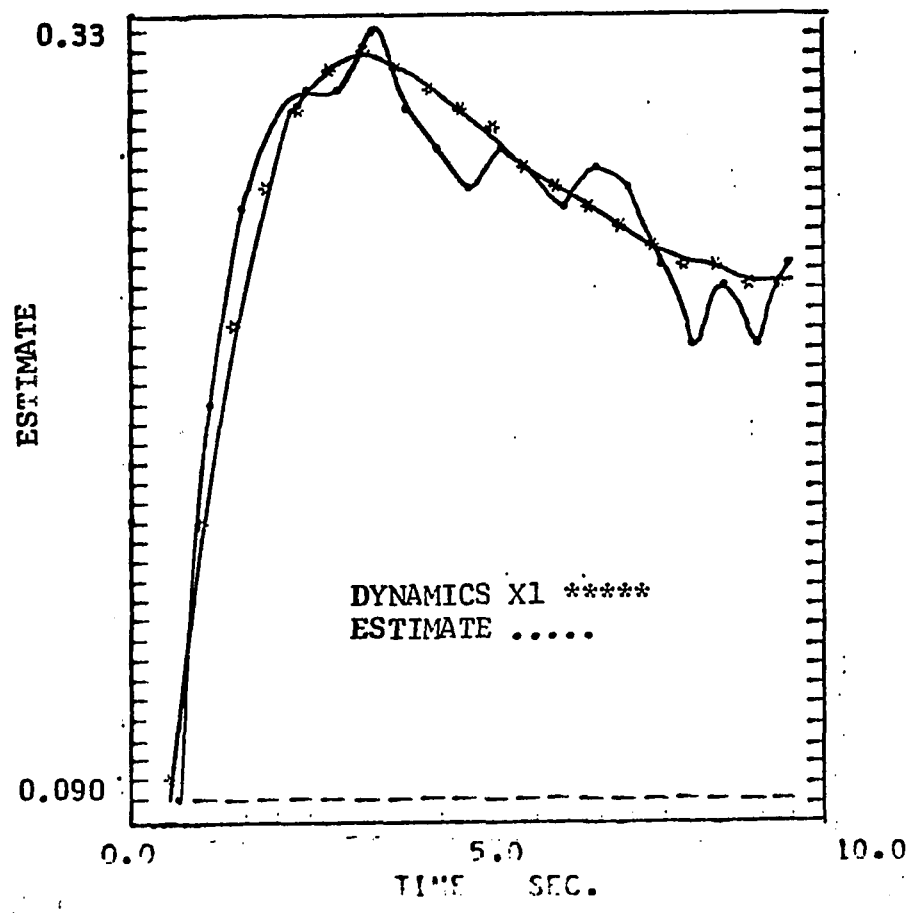


FIG. 4.13 20 % rate constant error
NONLINEAR EXTENDED KALMAN FILTER
WOLFS ADAPTIVE I

TIME SEC.	X 1	ESTIMATE
0.5	0.0288	0.0924
1.0	0.1310	0.2145
1.5	0.2427	0.2776
2.0	0.2845	0.2834
2.5	0.3097	0.3127
3.0	0.3222	0.3144
3.5	0.3255	0.3331
4.0	0.3228	0.3055
4.5	0.3165	0.2983
5.0	0.3083	0.2834
5.5	0.2995	0.2982
6.0	0.2907	0.2897
6.5	0.2827	0.2777
7.0	0.2755	0.2885
7.5	0.2695	0.2843
8.0	0.2645	0.2581
8.5	0.2607	0.2365
9.0	0.2578	0.2550
9.5	0.2559	0.2356
10.0	0.2547	0.2596

ACCUMULATED MEAN SQUARE ERROR = 0.5246E-02

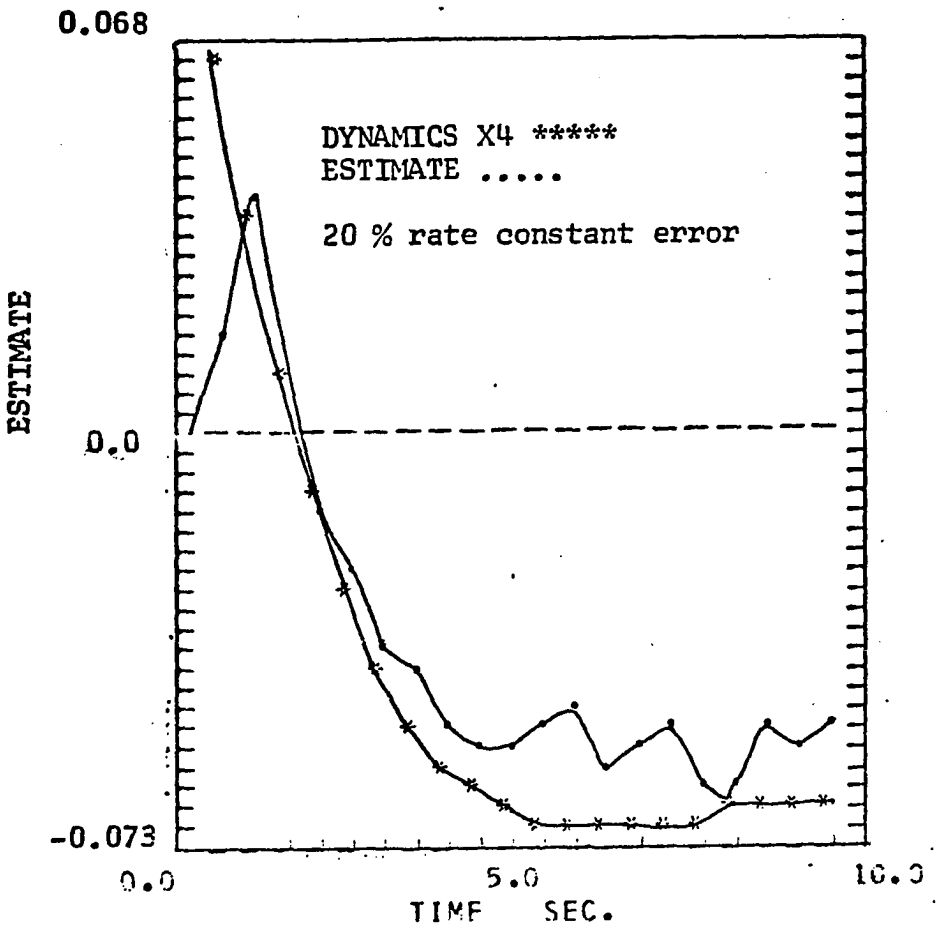


FIG. 4.14 MODEL ERROR COMPENSATION
NONLINEAR EXTENDED KALMAN FILTER
WOLFS ADAPTIVE I

TIME SEC.	X 4	ESTIMATE
0.5	0.0683	0.0165
1.0	0.0377	0.0419
1.5	0.0107	0.0115
2.0	-0.0120	-0.0152
2.5	-0.0301	-0.0264
3.0	-0.0440	-0.0407
3.5	-0.0543	-0.0445
4.0	-0.0616	-0.0550
4.5	-0.0666	-0.0593
5.0	-0.0698	-0.0597
5.5	-0.0716	-0.0531
6.0	-0.0725	-0.0500
6.5	-0.0727	-0.0614
7.0	-0.0724	-0.0566
7.5	-0.0718	-0.0550
8.0	-0.0711	-0.0661
8.5	-0.0703	-0.0671
9.0	-0.0696	-0.0559
9.5	-0.0689	-0.0565
10.0	-0.0684	-0.0530

ACCUMULATED MEAN SQUARE ERROR = 0.2574E-03

Table IV-2. Model error compensation technique and adaptive filtering.

Technique	Model Error Type					Adaptivity
	Integration error	Machine round off	Small model parameter error	Large Model parameter error	Statistics error	
Wolf I	C	C	B	C	C	D
Wolf II	B	C	A	C	B	B
Overweight	D	C	C	D	B	D
This Work I	A	B	A	C	A	A
This Work II	NJ	NJ	A	A	A	A
This Work III	C	C	B	A	B	B
Smoothing	NJ	NJ	B	A	B	A
Schmidt	NJ	NJ	A	A	C	A
Pine	B	A	C	C	C	B
This Work IV	NJ	NJ	A	A	A	A

NJ: Not justified
A: Excellent
B: Good
C: Fair
D: Bad

CHAPTER V

OBSERVABILITY, CONTROLABILITY AND INFORMATION MATRIX

V.1 Introduction

In estimating the state and parameters of a dynamic system from its measurement observations, it is important to ask the question: What, if anything, can be gained from filtering the observation data? How much information about the state of the system can be extracted from the Data? Can the state be determined from the measurement observation? This chapter tries to answer all of these questions by relating the state estimation to its measurement system.

The amount of information that can be gained from filtering the observation is related to the system model itself. In order to design a better filter, i.e. to gain more information from the system, it is usually desirable to reformulate the model of the system or take additional data or use an alternative measurement sensor system. This can be done by relating the information matrix to the process transition matrix and measurement matrix and consequently developing the upper and lower bounds of the estimation error covariance matrix in terms of the controlability and observability matrix.

A given dynamic system with its measurement sensor system is observable if its initial state $x(0)$ can be determined from the set

of measurements [z(1), z(2),...z(n)] from some finite N. If this is true for any initial time t_o, then the system is said to be completely observable.

A given dynamic system is controllable at time t_o if there exists a control policy sequence u(t), depending on $x(t_o)$ and for which $x(t_1) = 0$ over some interval $t_o \leq t \leq t_1$, that is, we can steer the system from x(t) to $x(t_o)$ by the control sequence u(t). If this is true for any t_o, then the system is completely controllable.

V.2 Information Matrix and Observabiltiy

The information matrix of a dynamic system can be defined as:

$$In \quad (k,1) = \sum_{i=1}^{k} \Phi'(i,k)H'(i)R_i^{-1}H(i)\Phi(i,k) \quad . \qquad \text{V-1}$$

Where

$\Phi(i,k)$ = process transition matrix as defined in Chapter II.

H(i) = measurement matrix

R = measurement error covariance matrix

k = sample period number.

Refer to Chapter II for system descriptions.

This information matrix is positive definite, since by assumptions R_i is positive definite and therefore its inverse. The information matrix is also nonsingular, otherwise, there will be no information about them in the observations since certain linear combinations of the elements of x(k) cannot be determined. This can be visualized in the determination of upper bound of error covariance matrix

in the next section. A singular information matrix will lead to infinity as the upper bound of the error covariance matrix.

Also, it should be noticed that the information matrix is independent of the observations, this nice property makes it possible to precompute the information matrix before the filtering begins.

Kalman [K3] first related this information matrix to the system observability by utilizing the following definitions:

The dynamic system of Chapter II is completely observable if and only if the information matrix is positively definite:

1. $$I_n(k+1,0) = \Phi'(k,k+1)I_n(k,0)\Phi(k,k+1)+H'(k+1)R^{-1}(k+1)H(k+1) \qquad \text{V-2}$$

2. It is related to error covariance matrix by:

$$P^{-1}(k,k) = \Phi'(0,k)P_0^{-1}\Phi(o,k) + I_n(k,0) \ . \qquad \text{V-3}$$

V.3 Controlability.

The similarities between the controlability and observability are quite evident from the definitions.

Define the controlability matrix

$$C(k,0) = \sum_{i=0}^{k} \Phi(k,i+1)\Gamma(i)\ Q(i+1)\ \Gamma'(i)\ \Phi'(k,i+1). \qquad \text{V-4}$$

Where

Q(i+1) is the process noise covariance matrix

$\Gamma(i)$ is the process noise gain matrix.

A discrete dynamic system is completely controlable if and only if its controlability matrix is positive definite.

The duality relation between controlability and observability was first found by Kalman [K3] in a linear noise free system. This duality is also true for stochastic system:

Observable	Controlable
$I_n > 0$	$C > 0$
(t,t_o)	(t_o,t)
$H(t)$	$\Gamma(t)$
$R^{-1}(t)$	$Q(t)$

All of the theory regarding controlability and observability only apply to linear systems. For non-linear systems there is no global theory relating to controlability and observability, nor does the duality relation exist. In such problems, it is convenient to linearize the system equations about some assumed set of nominal conditions and apply the linear system theory to the linearized system.

Some additional definitions regarding controlability and observability will be useful for developing the upper and lower bounds of the error covariance matrix in the next section.

A stochastic linear system is uniformly completely controlable if there exists a positive integer N and positive constants α, β such that

$$0<\alpha I \leq C(k,i-N) \leq \beta I \quad . \qquad \text{V-5}$$

A stochastic linear system is uniformly completely observable if there exists a positive integer N and positive constants α, β such that

$$o<\alpha I \leq I_n(k,k-N) \leq \beta I \quad .$$

However, some progress has been made in the areas of non-linear controlability and observability for non-linear process with control appearing linearly. Roitenberg relates the observability of a non-linear system to its system Liapunov function. Thus one will be able to investigate the observability for the system:

$$\dot{\underline{x}} = A(t)\underline{x} + \underline{f}(x) + \underline{q}(t) \quad . \qquad \text{V-7}$$

With the observations:

$$\underline{y} = c\underline{x} \quad .$$

Where

A(t): matrix function of time

$\underline{f}(x)$: non-linear function of x

$\underline{q}(t)$: forcing function.

Haynes and Hermes [H1] in a separate paper, introduced non-linear controlability via Lie theory. Their theory is restricted to the following system:

$$\dot{x} = B[x(t)]u(t) \quad ,$$

where

$B[\underline{x}(t)]$ is a nonlinear matrix function of x(t)

and

$\underline{u}(t)$ is the control vector.

Lobry also uses Lie groups and arrived at the same conclusion.

All these approaches are confined to qualitative studies only. Further work is definitely needed in order to apply these techniques to real system evaluation.

V.4 Bounds for the Estimation Error Covariance Matrix

By using the concept of stochastic observability and controllability we can develope some qualitative properties of the linear filter in terms of an upper and lower bound of their estimation error covariance matrix. The following development is based on the work of Kalman, Deyst [K3] and Price and Sorenson [D2].

Let us define our dynamic stochastic system again:

$$\underline{x}(k+1) = \Phi(k+1,k)\underline{x}(k) + \Gamma(k)\underline{w}(k+1) \quad . \qquad \text{V-8}$$

With the discrete linear observations

$$z(k) = H(k)x(k) + v(k) \quad . \qquad \text{V-9}$$

Refer to Chapter II for nomenclature.

1. If the dynamic system V-8 is uniformly completely observable and uniformly completely controlable, and if $P_o \geq 0$ then P(k,k) is uniformly bounded from above for all $k \geq N$.

$$P(k,k) \leq I_n^{-1}(k,k-N) + C(k,k-N) \leq \left(\frac{1+\alpha\beta}{\alpha}\right) I \text{ for } k \geq N \quad . \qquad \text{V-10}$$

Refer to Jazwinski [J2] for a detailed proof.

2. If the dynamic system V-8 is uniformly, completely observable and uniformly, completely controlable and if $P_o \geq 0$ then P(k,k) is uniformly bounded from below for all k>N by

$$P(k,k) \geq [I_n(k,k-N) + C^{-1}(k,k-N)]^{-1} \geq \left(\frac{\alpha}{1+\alpha\beta}\right) I \text{ for } k > N \quad . \qquad \text{V-11}$$

3. If the dynamic system V-8 is uniformly completely controllable and $P_0 \geq 0$, then $P(k,k) \geq 0$ for all $k > N$.

Properties 1 and 2 give us qualitative upper and lower bounds for the estimate error covariance. The uniformity property will provide us some idea of the best and worst we can obtain for a given filter model. Property 3 provides us with the necessary condition for a stable filter by keeping the error covariance matrix inside the positive definite cone. Another advantage of the above development can be visualized since it is independent of the observation data, which makes it possible for prefiltering analysis and allows the design of a better filter model that will give acceptance error bounds since the bounds are only dependent on the system model.

V.5 Numerical Examples and Discussions

The upper and lower bounds of the estimate error covariance matrix in linear filtering theory can be extended to the linearized system of a non-linear problem. This is applied successfully to our reactor model as shown in Fig. 5.1 to Fig. 5.8. Observe that the filter estimation error lies within the bounds after the filter reaches the steady state (k>N) even though it started outside the bound.

It is essential that the filter has to be kept within the bound when it is approaching the steady state, since the bounds qualitatively define a region of controlability and observability or equivalently a region of the existence of a stable filter. This is shown in Fig. 5.1 through Fig. 5.8.

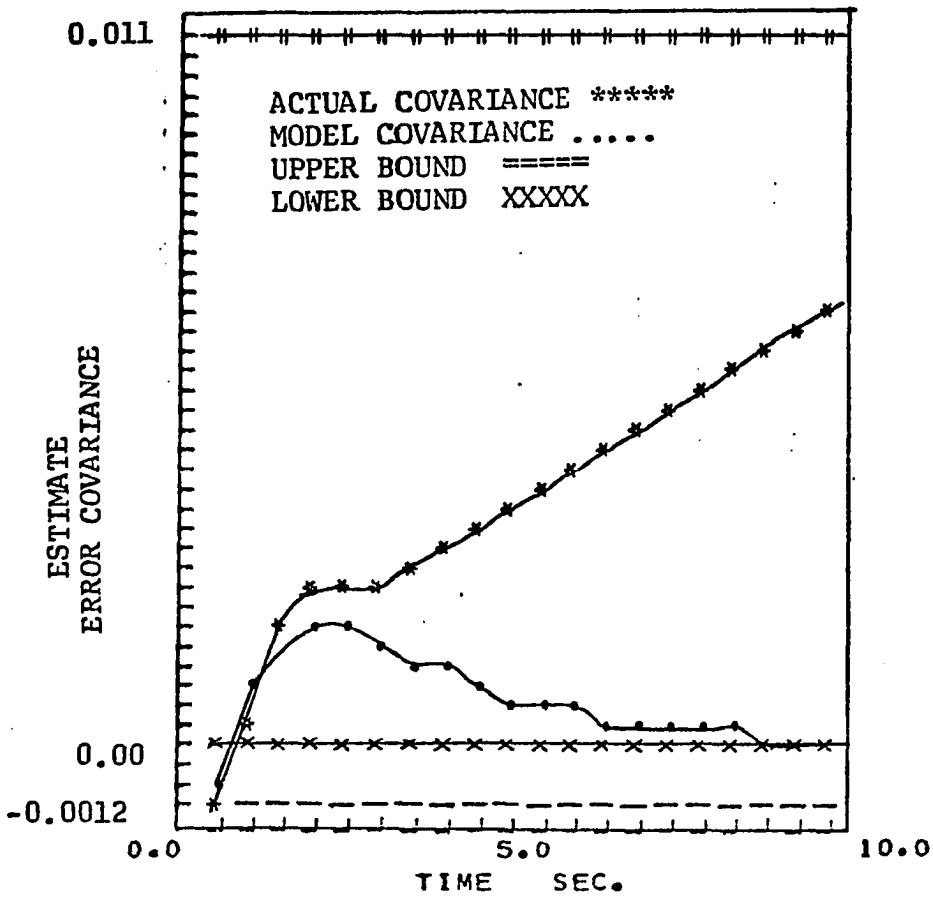


FIG. 5.1 FILTER DIVERGENCE
SECOND ORDER NONLINEAR FILTER
SENSITIVITY ANALYSIS DUE TO MODEL ERROR
NO NOISE COMPENSATION

TIME SEC.	ACTUAL P	MODEL P
0.5	-0.0012	-0.0007
1.0	0.0005	0.0011
1.5	0.0025	0.0024
2.0	0.0031	0.0025
2.5	0.0032	0.0023
3.0	0.0034	0.0020
3.5	0.0036	0.0017
4.0	0.0040	0.0014
4.5	0.0044	0.0012
5.0	0.0048	0.0010
5.5	0.0052	0.0008
6.0	0.0057	0.0007
6.5	0.0061	0.0005
7.0	0.0065	0.0004
7.5	0.0070	0.0004
8.0	0.0074	0.0003
8.5	0.0078	0.0003
9.0	0.0082	0.0002
9.5	0.0085	0.0002
10.0	0.0089	0.0002

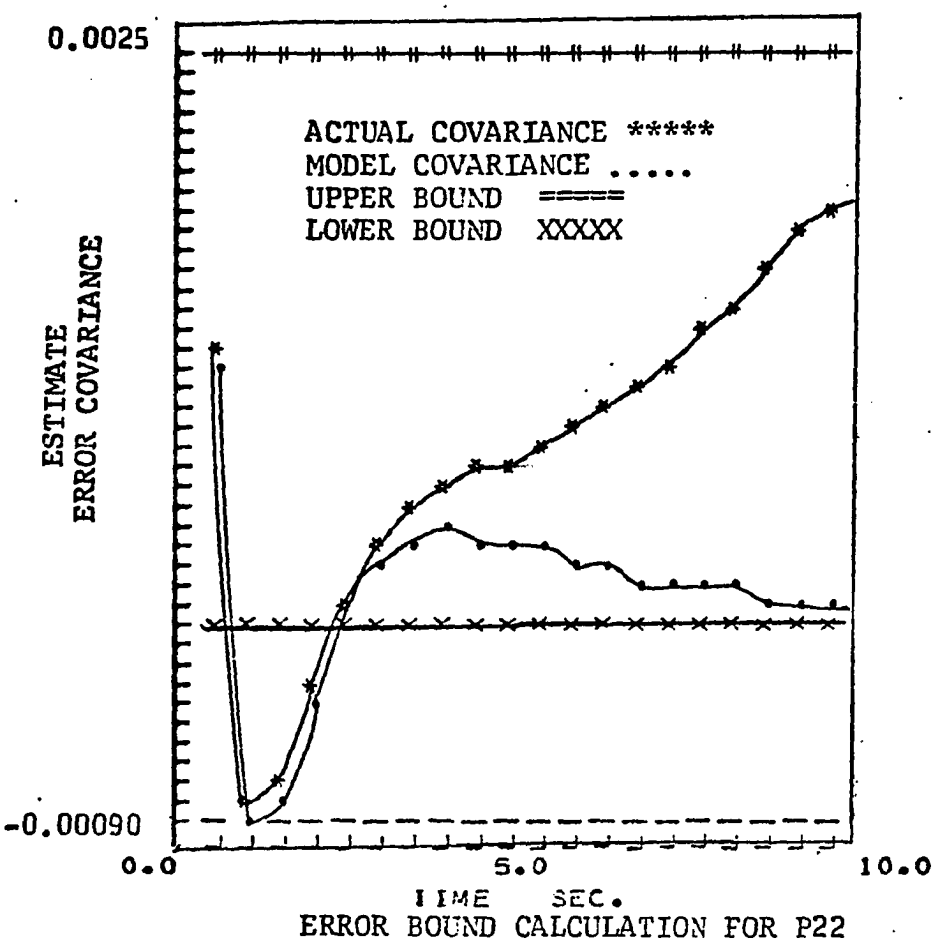
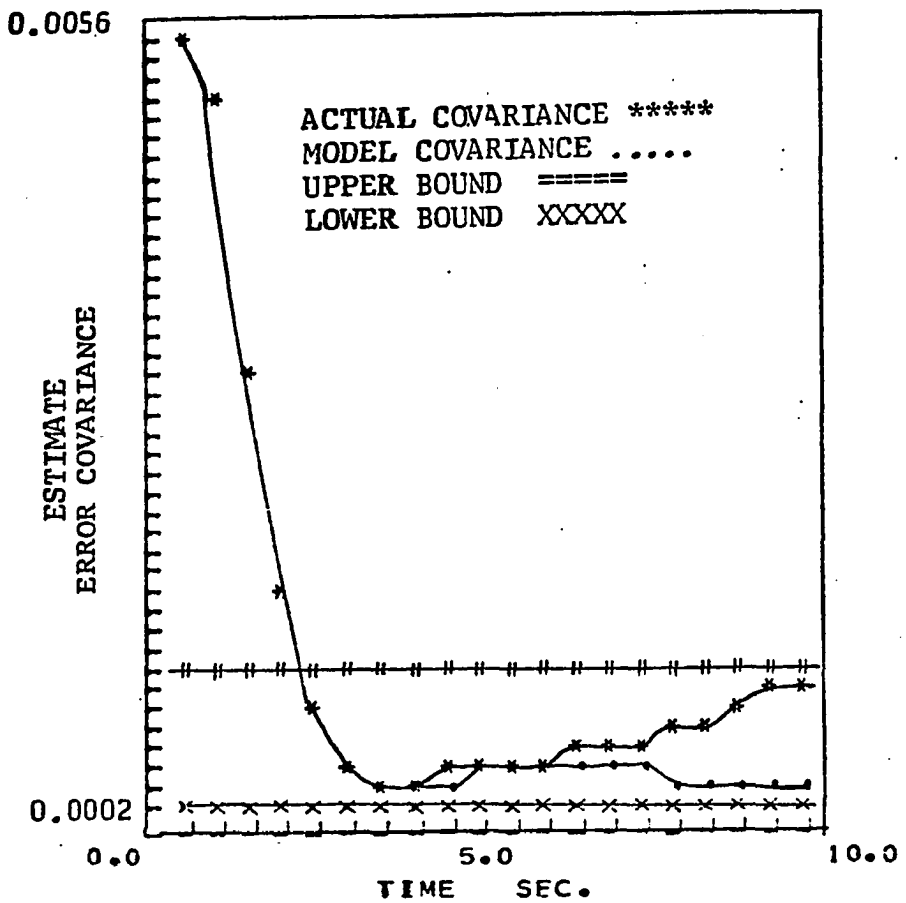


FIG. 5.2 FILTER DIVERGENCE
SECOND ORDER NONLINEAR FILTER
SENSITIVITY ANALYSIS DUE TO MODEL ERROR
NO NOISE COMPENSATION

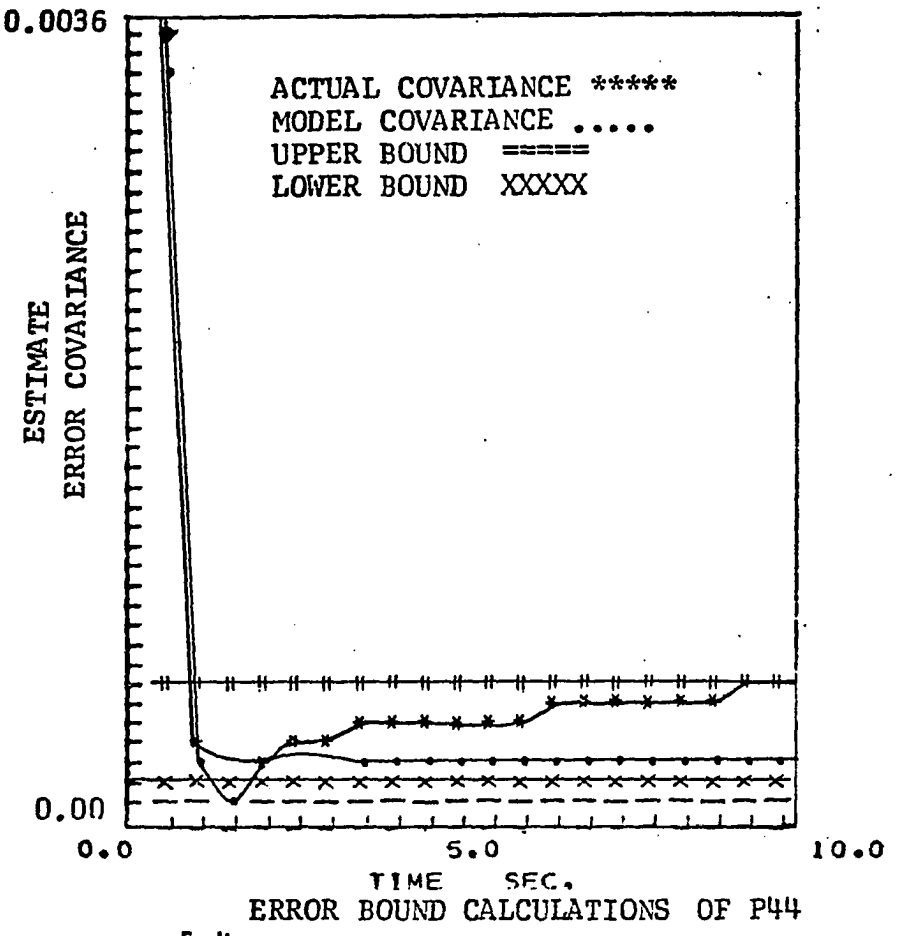
TIME SEC.	ACTUAL P	MODEL P
0.5	0.0016	0.0015
1.0	-0.0010	-0.0011
1.5	-0.0009	-0.0010
2.0	-0.0003	-0.0004
2.5	0.0002	0.0001
3.0	0.0006	0.0004
3.5	0.0008	0.0005
4.0	0.0009	0.0006
4.5	0.0010	0.0006
5.0	0.0010	0.0005
5.5	0.0011	0.0005
6.0	0.0012	0.0004
6.5	0.0013	0.0004
7.0	0.0015	0.0003
7.5	0.0016	0.0003
8.0	0.0018	0.0003
8.5	0.0020	0.0002
9.0	0.0021	0.0002
9.5	0.0023	0.0002
10.0	0.0025	0.0002



ERROR BOUND CALCULATION FOR P33

FIG. 5.3 FILTER DIVERGENCE
SECOND ORDER NONLINEAR FILTER
SENSITIVITY ANALYSIS DUE TO MODEL ERROR
NO NOISE COMPENSATION

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0056	0.0056
1.0	0.0052	0.0052
1.5	0.0031	0.0031
2.0	0.0016	0.0015
2.5	0.0007	0.0007
3.0	0.0003	0.0003
3.5	0.0002	0.0002
4.0	0.0002	0.0002
4.5	0.0002	0.0002
5.0	0.0003	0.0002
5.5	0.0003	0.0002
6.0	0.0003	0.0002
6.5	0.0004	0.0002
7.0	0.0004	0.0002
7.5	0.0005	0.0002
8.0	0.0005	0.0002
8.5	0.0006	0.0002
9.0	0.0007	0.0002
9.5	0.0008	0.0002
10.0	0.0009	0.0002



ERROR BOUND CALCULATIONS OF P44

FIG. 5.4 FILTER DIVERGENCE
SECCND ORDER NONLINEAR FILTER
SENSITIVITY ANALYSIS DUE TO MODEL ERROR
NO NOISE COMPENSATION

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0036	0.0034
1.0	0.0001	0.0000
1.5	-0.0000	-0.0001
2.0	0.0000	-0.0001
2.5	0.0001	-0.0000
3.0	0.0002	-0.0000
3.5	0.0002	0.0000
4.0	0.0002	0.0000
4.5	0.0002	0.0000
5.0	0.0003	0.0000
5.5	0.0003	0.0000
6.0	0.0003	0.0000
6.5	0.0003	0.0000
7.0	0.0003	0.0000
7.5	0.0003	0.0000
8.0	0.0003	0.0000
8.5	0.0004	0.0000
9.0	0.0004	0.0000
9.5	0.0004	0.0000
10.0	0.0004	0.0000

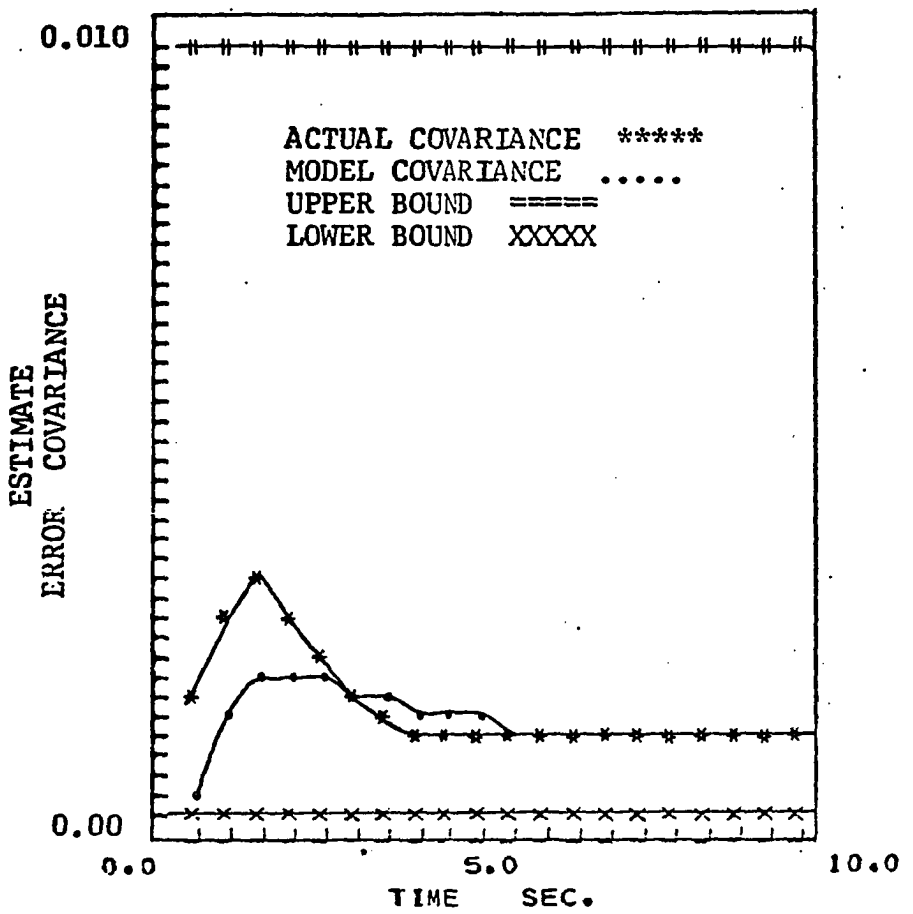


FIG. 5.5 ERROR BOUND PLOT FOR P11
SECCND ORDER NONLINEAR FILTER
SENSITIVITY ANALYSIS DUE TO PRIOR STATISTICS
CONSTANT NOISE

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0021	0.0005
1.0	0.0036	0.0017
1.5	0.0044	0.0026
2.0	0.0037	0.0026
2.5	0.0029	0.0025
3.0	0.0023	0.0022
3.5	0.0019	0.0021
4.0	0.0016	0.0019
4.5	0.0015	0.0018
5.0	0.0015	0.0017
5.5	0.0015	0.0017
6.0	0.0015	0.0016
6.5	0.0015	0.0016
7.0	0.0016	0.0016
7.5	0.0016	0.0016
8.0	0.0016	0.0016
8.5	0.0016	0.0016
9.0	0.0016	0.0016
9.5	0.0016	0.0016
10.0	0.0016	0.0016

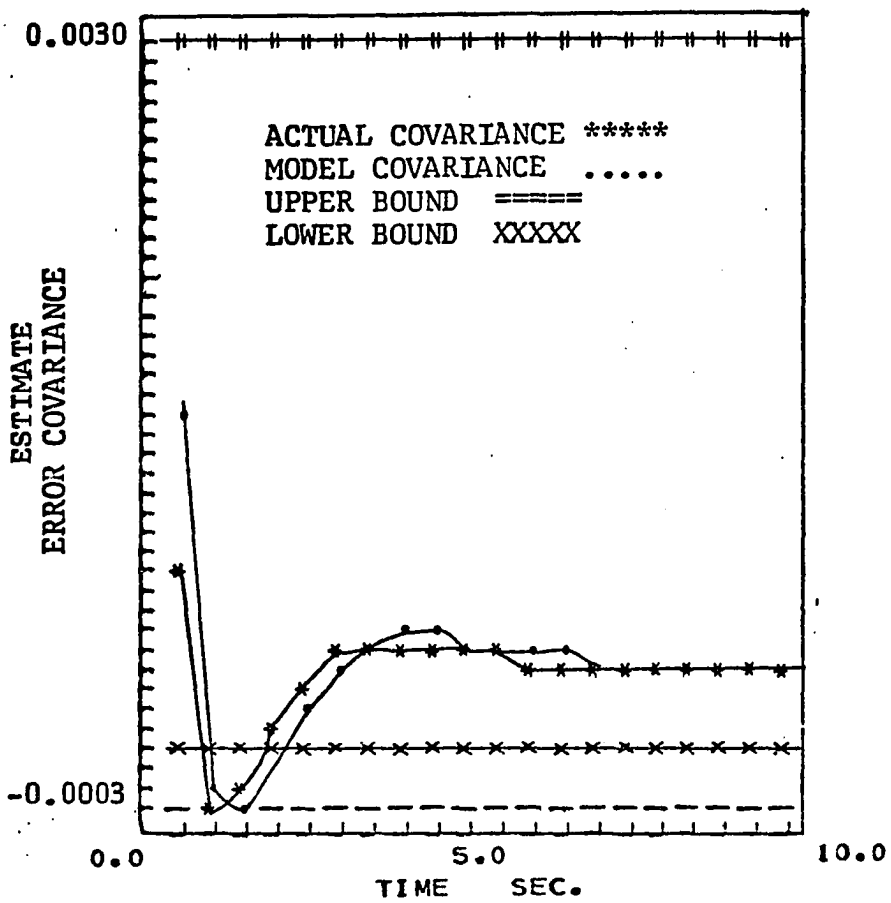


FIG. 5.6 ERROR BOUND PLOT FOR P22
SECOND ORDER NONLINEAR FILTER
SENSITIVITY ANALYSIS DUE TO PRIOR STATISTICS
CONSTANT NOISE

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0008	0.0016
1.0	-0.0003	-0.0002
1.5	-0.0002	-0.0003
2.0	0.0001	-0.0001
2.5	0.0003	0.0002
3.0	0.0005	0.0004
3.5	0.0005	0.0005
4.0	0.0005	0.0005
4.5	0.0005	0.0005
5.0	0.0004	0.0005
5.5	0.0004	0.0005
6.0	0.0004	0.0004
6.5	0.0004	0.0004
7.0	0.0004	0.0004
7.5	0.0004	0.0004
8.0	0.0004	0.0004
8.5	0.0004	0.0004
9.0	0.0004	0.0004
9.5	0.0004	0.0004
10.0	0.0004	0.0004

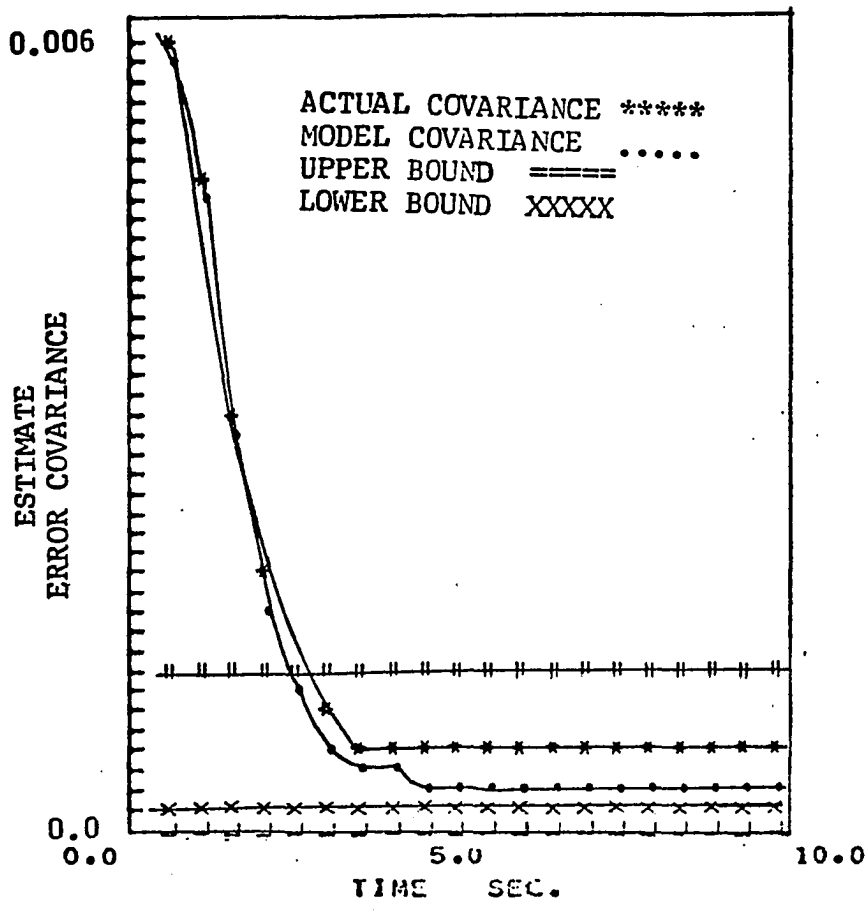
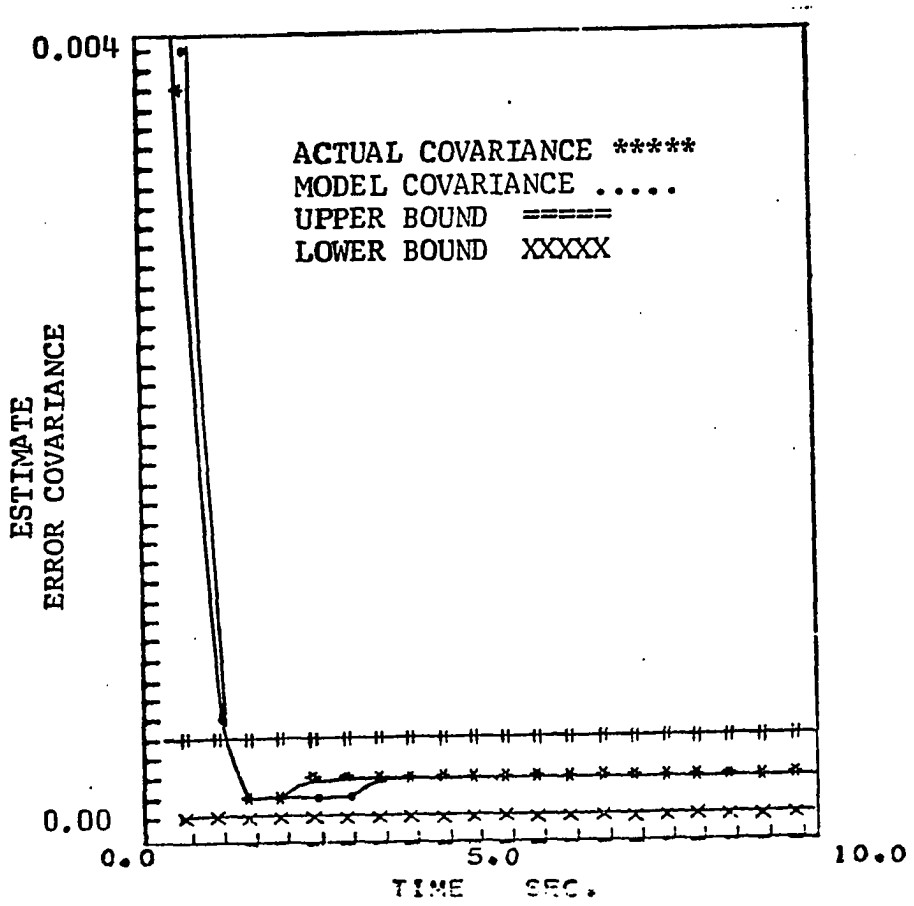


FIG. 5.7 ERROR BOUND PLOT FOR P33
SECOND ORDER NONLINEAR FILTER
SENSITIVITY ANALYSIS DUE TO PRIOR STATISTICS
CONSTANT NOISE

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0057	0.0056
1.0	0.0046	0.0044
1.5	0.0029	0.0027
2.0	0.0017	0.0015
2.5	0.0010	0.0008
3.0	0.0007	0.0005
3.5	0.0005	0.0003
4.0	0.0004	0.0002
4.5	0.0004	0.0002
5.0	0.0004	0.0002
5.5	0.0004	0.0002
6.0	0.0004	0.0002
6.5	0.0004	0.0002
7.0	0.0004	0.0002
7.5	0.0004	0.0002
8.0	0.0004	0.0002
8.5	0.0004	0.0002
9.0	0.0004	0.0002
9.5	0.0004	0.0002
10.0	0.0004	0.0002



ERROR BOUND CALCULATION FOR P44

FIG. 5.8 ERROR BOUND PLOT FOR P44
SECOND ORDER NONLINEAR FILTER
SENSITIVITY ANALYSIS DUE TO PRIOR STATISTICS
CONSTANT NOISE

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0042	0.0044
1.0	0.0005	0.0006
1.5	0.0001	0.0001
2.0	0.0001	0.0001
2.5	0.0002	0.0001
3.0	0.0002	0.0002
3.5	0.0002	0.0002
4.0	0.0002	0.0002
4.5	0.0002	0.0002
5.0	0.0002	0.0002
5.5	0.0002	0.0002
6.0	0.0002	0.0002
6.5	0.0002	0.0002
7.0	0.0002	0.0002
7.5	0.0002	0.0002
8.0	0.0002	0.0002
8.5	0.0002	0.0002
9.0	0.0002	0.0002
9.5	0.0002	0.0002
10.0	0.0002	0.0002

CHAPTER VI

STOCHASTIC STABILITY

VI.1 Introduction.

Deterministic Stability is a branch of the qualitative theory of dynamic systems. Most current results are oriented to the qualitative properties of differential equations (to avoid solving the actual differential equation explicity). Lefschetz, Krasovskii [B8] have made very important contributions in this field.

Related questions arise in the analysis of stochastic processes. The first paper that suggested the existence of a stochastic Liapunov function as in the deterministic case was that by Bertram and Sarachik [B8], and Krasovskii [B8]. It is desirable to define several technical terms for the analysis of the stability of stochastic processes before we continue to introduce the stability theorems.

Definition 1.

Let the dynamic system be

$$\begin{aligned} \underline{\dot{x}} &= \underline{f}(\underline{x},t) \\ \underline{x}(t_o) &= \underline{c} \quad . \end{aligned} \tag{VI-1}$$

Then an equilibrium point of VI-1 is stable if for every $\varepsilon>0$. There exists a $\delta(\varepsilon,t_o)$ such that if $||x_o-x_e||<\delta$ then the dynamics $\Phi(t,x_o,t_o)$

lie within a certain bounded region of x_e (x_e is said to be uniformly stable if it is stable and δ may be chosen independent of t_o).

Definition 2.

Let x_e be an equilibrium point of VI-1, then x_e is asyptotically stable if it is stable and every motion of VI-1 starting sufficient close to x_e converges to x_e as $t \to \infty$.

Definition 3.

A sequence of random variables $(x_n), n = 1,2,\ldots n$ such that the conditional expectation

$$E(x_n | x_{n-1}, \ldots x_1) \leq x_{n-1}$$

and

$$E|x_1| < \infty$$

is called a supermartingale. There are some useful properties of supermartingale. Let (x_n) be a positive supermartingale and $\lambda > 0$ then

$$P(\sup x_i \geq \lambda) \quad \frac{E(x_1^+)}{\lambda} \quad .$$

This means the probability of that the maximum value of the random sequence will be greater than λ is less then $E(x_1^+)/\lambda$.

Definition 4.

Let the discrete stochastic system be

$$\underline{x}_n = \underline{f}(\underline{x}_{n-1}, \underline{r}_{n-1}) \qquad \text{VI-2}$$

$$\underline{x}_o = \underline{C} \quad .$$

Where $f(0,r_n)$ are a sequence of random vectors and c is also a random vector.

The dynamic system VI-2 is stochastically stable about 0 if and only if every $0<\rho<1$ and ε positive, there exists $0<\sigma<1$ and δ positive so that $P(\sup_n ||x_n|| \geq \varepsilon) < \rho$ for all c such that

$$P(||c|| \geq \delta) < \sigma.$$

A sufficient condition for stability can be found by a stochastic Liapunov function.

Theorem: Suppose there exists a continuous positive scalar function V(x) satisfying:

a) $V(0) = 0$;

b) $V(x) \to \infty$ as $||x|| \to \infty$;

c) $V(x_n)$ is a supermartingale along the dynamics of VI-2 then VI-2 is said to be stochastically stable.

Theorem: If the dynamic system VI-2 is stochastically stable and there exists a positive real function $\gamma(||x_{n-1}||)$ such that

$$E[V(x_n)|x_{n-1},\ldots x_1) - V(x_{n-1}) \leq - \gamma(x_{n-1})$$

is satisfied along VI-2, then VI-2 is stochastically asymptotically stable. Refer to Bucy [B8] for a detailed proof.

VI.2 Some Forms of Stochastic Functions.

Since Bucy's [B8] introduction of the supermartingale property of a stochastic Liapunov function, Kusher [K4] made several successful

attempts to introduce various forms of a stochastic Liapunov function for continuous random differential equations. However, there still are few studies about stochastic Liapunov functions for discrete stochastic systems.

1. Let V be the Liapunov function of the filter system

$$\begin{aligned} \underline{x}_n &= \underline{f}(\underline{x}_{n-1}, \underline{r}_{n-1}) \\ \underline{x}_o &= \underline{c} \quad . \end{aligned} \tag{VI-3}$$

Then $V(x_n) = x_n' P(x) x_n$ for $n > N$ (approaching steady state) where $P(n)$ is the error covariance matrix of the estimate at time (n).

Proof:

$V(x_n)$ satisfies the requirement for a Liapunov function since

a) $V(0) = 0$

b) $V(x_n) \to \infty$ as $x_n \to \infty$

c) From the properties of P introduced in the last chapter

$$(\frac{\alpha}{1+\alpha\beta}) I \leq P(n) \leq (\frac{1+\alpha\beta}{\alpha}) I \quad , \tag{VI-4}$$

so that

$$\frac{\alpha}{1+\alpha\beta} ||x||^2 \leq x' P(n) x \leq \frac{1+\alpha\beta}{\alpha} ||x||^2 I \quad . \tag{VI-5}$$

2. The second form of Liapunov function for the non-linear discrete filter introduced by this work is the analogy of the deterministic Krasovakii form:

$$V(x_n) = f'f$$

since this also satisfies:

$$V(0) = 0$$

$$V(x_n) \to \infty \quad \text{as} \quad x_n \to \infty$$

$V(x_n)$ is a positive scalar.

3. The Liapunov function introduced by Jazwinski [J2]

$$V(x_n) = x' P^{-1}(n) x \ . \qquad \text{VI-6}$$

Refer to Jazwinski for a detailed proof.

All of the above Liapunov functions were tested on our reactor system. The following conclusions were reached:

a) The first form introduced by this work gives the largest region of asymptotic stability (RAS), provided the error covariance P(n) stays in the positive semidefinite cone (i.e. the filter stays with the observability and controlability region).

b) The Krasovskii form is too conservative (gives a smaller RAS).

c) The third form does not apply to our reactor case since $P^{-1}(n)$ is a monotonic increasing function of n.

VI.3 Properties of the Riccati Equation

The asymptotic behavior is closely related to the solution of the estimate error covariance matrix equation which satisfies the Riccati matrix equation.

$$\dot{P} = F(t)P + PF'(t) - PH'(t)R^{-1}(t)H(t)P + \Gamma(t)Q(t)\Gamma(t), \qquad \text{VI-7}$$

where P is the error covariance matrix for the dynamic system.

$$d\underline{x} = F(t)\underline{x}(t)dt + G(t)d\underline{w}$$

$$\underline{x}(t_o) = \underline{c}$$

$$d\underline{z} = H(t)\underline{x}(t)dt + d\underline{v}$$

$$P(t_o) = P_o \geq 0 \ .$$

The solution of VI-7 has to stay inside the cone of positive semi-definite matrices. The solution of VI-7 satisfies

$$0 \leq P \leq \Phi(t,t_o)P\Phi'(t,t_o) + \int_{t_o}^{t} \Phi(t,\tau)\Gamma(\tau)Q(\tau)\Gamma'(\tau)\Phi(t,\tau)d\tau, \qquad \text{VI-8}$$

where

$\Phi(t,t_o)$ is the fundamental matrix of F(t).

Note that the second term on the right hand side of Equation VI-8 is the controlability matrix. The existence and uniqueness of the solution of the Equation VI-7 is given by Bucy [B8], since the upper bound of Equation VI-8 provides a Lipschitz constant for Equation VI-7. From Equation VI-8 it is obvious that the existence of a unique positive semidefinite solution of Equation VI-7 depends on:

1. The boundedness of the asymptotic stability properties of the system transition matrix.

2. P_o has to be positive semidefinite.

3. The controlability matrix of the filter has to be positive semidefinite.

Numerical solution of the Riccati equation.

Bucy's ASP program [B8] (automatic Synthesis Program) provides

an efficient procedure for the numerical solution of the Riccati matrix equation via quasilinearization. We will discuss two types of numerical schemes solving the Riccati matrix equation with subsequent successful application to our reactor system.

Let the solution of Riccati equation be S(P), then

$$\dot{P} = S(P) = FP + PF' - PH'R^{-1}HP + \Gamma Q \Gamma .$$

1. The first order method: (this is a simple first-order iteration)

$$P_n = P_{n-1} + \varepsilon S(P_{n-1}) \quad . \qquad \text{VI-9}$$

2. Second-order iteration:

$$P_n = P_{n-1} + \varepsilon S(P_{n-1}) + \tfrac{1}{2}\varepsilon^2 \frac{\partial S(P_{n-1})}{\partial P} S(P_{n-1}) \quad . \qquad \text{VI-10}$$

When both of these schemes were applied to our reactor system, the second order iteration scheme shows great improvement over the first order during the initial stages of the calcualtion.

The requirement that the P matrix symmetric can be seriously violated during the iteration process due to round off in the matrix operations but this can be avoided by symmetrizing the iterates. The problem of the iterates leaving the cone of positive semidefinite and thus leading to an unbounded filter can be avoided by adding a larger fictitious noise covariance term, the Q matrix. This increases the controlability region.

The computation time in carrying out the converged solution of the Riccati equation depends on the step size ε, but too small a ε leads to excessive iteration and computing time. Therefore, should be chosen as large as possible. Generally speaking, the computation time varies as the cube of the dimension of the Riccati matrix equation. However, all of the trouble in symmetrizing and staying positive definite are partially avoided in our work by using the special matrix calculational scheme, thus keeping the round off errors very low.

The solution of matrix Riccati equation governs the evolution of the estimation error covariance matrix for the discrete time filter. It is important to know the conditions for the uniqueness of the Riccati equation which therefore ensures the optimal filter performance. A good table made by Bucy [B8] is very useful for quick reference.

Let us define the following nomenclature:

A_1: Uniform R observability

A_2: Uniform Q controlability

A_3: Uniform Q controlability and uniform observability

P_1: Existence of a positive semi-definite equilibrium solution of the Riccati equation

P_2: Uniform asymptotic stability of the optimal filter

P_3: Uniform asymptotic stability of the Riccati equation

P_4: A priori uniform upper bound $(I_n^{-1} + C)$ for the solution of Riccati equation

P_5: A priori uniform lower bound $(I_n+C\ ^1)^{-1}$ for the solution of the Riccati equation

P_6: Unique positive semidefinite equilibrium solution of the Riccati equation

P_7: Monotonic convergence of the error covariance matrix.

For the time varying system:

With

$$R(t)>cI>0, \quad P(0)\geq 0$$

and

$$F(t),\ \Gamma(t),\ H(t),\ Q(t)$$

uniformly bounded. We have the following property table:

	P_1	P_2	P_3	P_4	P_5	P_6	P_7
A_1	X	-	-	X	-	-	X
A_2	-	-	-	-	X	-	X
A_3	X	X	X	X	X	X	X

An entry (X) in the matrix i, j indicates that under assumption i, the property holds while (-) indicates it is insufficient.

VI.4 Numerical Examples.

A sample computer output shows the Liapunov function suggested by this work and that of the Krasovski form. The Liapunov functions all start at the origin with a value of zero, then increases steadily until the filter approaches the steady state, then decrease monotonically. The Liapunov function of the form $\hat{x}'P\hat{x}$ shows a clear indication

of the stability property of the filter model, since the positive definiteness of the P matrix can assure a unique solution of the Riccati equation which in turn will assume a stable filter. The fact that P is an monotonic decreasing function of time after the filter reaches the steady state directly reflects the location of the steady state. For the Krasovski's form $\hat{\dot{x}}'\hat{\dot{x}}$, the trend of monotonic decreases start a little earlier where the filter may not be stable.

SAMPLE COMPUTER OUTPUT OF LYAPUNOV FUNCTIONS

TIME SEC.	THIS WORK $x'Px$	KRASOVSKI $x'x$
0.5	0.0	0.000000
1.0	0.000000	0.006903
1.5	0.000022	0.015543
2.0	0.000108	0.014413
2.5	0.000188	0.008476
3.0	0.000274	0.005843
3.5	0.000337	0.004166
4.0	0.000370	0.003170
4.5	0.000364	0.002413
5.0	0.000347	0.001921
5.5	0.000316	0.001465
6.0	0.000277	0.001091
6.5	0.000252	0.000799
7.0	0.000235	0.000556
7.5	0.000227	0.000364
8.0	0.000223	0.000230
8.5	0.000217	0.000148
9.0	0.000217	0.000095
9.5	0.000212	0.000056
10.0	0.000221	0.000033

CHAPTER VII

ERROR SENSITIVITY ANALYSIS AND FILTER ADAPTIVITY

VII.1 Introduction

In the implementation of a filter, we often face two difficult problems; the choice of a proper initial estimate of the state, the initial estimate error covariance matrix, system noise covariance matrix, measurement covariance matrix and the choice of a proper mathematical model for the system. The mathematical model must be simple in order to make it feasible for on-line calculations but it should be realistic enough to describe the system in the operating range; on the other hand, the prior statistics and the initial conditions are not easily available under normal conditions. Therefore, these often produce erroneous filtering results or filter divergence as shown in previous chapters.

The choice of different initial estimate error covariance does not effect the filtering results when the filter is approaching the steady state. This Gaussian-Markov properties, as proved in the last chapter, may not hold for non-linear systems or it may take longer to approach to the steady state, or there may exist several steady state solutions as discussed in the previous chapters. Different

choices of Q(k), (the process noise covariance matrix), and R(k), (the measurement noise covariance matrix) may lead to the erroneous filtering results as they may lead to an undesired steady state solution of the matrix Riccati equation.

It is clear that an error sensitivity analysis is urgently needed, especially in non-linear filtering, because of the uncertainty in both the model parameters and prior statistics. This uncertainty gives severe problems for non-linear filtering while it is less sensitive in the linear filtering case. The most complete development of an error sensitivity analysis has been done by Griffin and Sage [G4]. They consider both differential errors and large scale errors and also apply the results to smoothing. However, they only are useful for a simple time invariant model.

This work extends the results of Griffin and Sage and Jazwinski [J2] to a more general algorithm. One can analyze the error sensitivity due to model error (uncertainty in model parameters or integration truncation error), prior statistics, measurement errors, initial conditions or any combination of the above categories. This sensitivity analysis calculation can be utilized by the adaptive filters discussed in the previous chapters. Therefore, we are able to do a sensitivity analysis for various adaptive filters, a second order non-linear filter, an extended Kalman filter and the smoothing. Thus, we can relate the sensitivity to adaptivity.

An extensive error sensitivity analysis was applied to the chemical reactor model in this work. Some encouraging conclusive

results are obtained which will be valuable for non-linear filtering implementations. A relation between error densitivity and filter adaptivity found in this study also provides guides to adaptive filtering development.

VII.2 Development of General Sensitivity Analysis Algorithm

In order to apply the filtering to a system, the model parameters Φ(process transition matrix), H(process measurement matrix), noise statistics (Q and R matrix) and the initial conditions [x(0) and P(0)] must be specified.

Following the approach of Sage and Jazwinski [G4], [J2]; let the real system be represented by

$$\underline{x}(k+1) = \Phi(k+1,k)\underline{x}(k) + \Psi(k) + \Gamma(k)\underline{w}(k+1) \quad , \qquad \text{VII-1}$$

with the measurement

$$\underline{z}(k) = H(k)\underline{x}(k) + \underline{v}(k) \quad , \qquad \text{VII-2}$$

where $\underline{w}(k+1)$ is a zero mean Gaussian process noise with covariance Q(k+1). $\underline{v}(k)$ is a zero mean Gaussian measurement noise with covariance matrix R(k+1). The initial estimate is $\underline{x}(0)$ and the error covariance is P(0) while the filter model is:

$$\underline{x}_c(k+1) = \Phi_c(k+1,k)\underline{x}_c(k) + \Psi_c(k) + \Gamma_c(k)\underline{w}_c(k+1) \qquad \text{VII-3}$$

$$\underline{z}_c(k) = H_c(k)\underline{x}_c(k) + \underline{v}_c(k) \qquad \text{VII-4}$$

where $\underline{w}_c(k+1)$ is a zero mean process noise with covariance $Q_c(k)$, $\underline{v}_c(k)$ is a zero mean measurement noise with covariance $R_c(k)$, initial

estimate $\underline{x}_c(0)$, $P_c(0)$; $\Psi(k)$ and $\Psi_c(k)$ are in the dynamics to account for any approximations due to nonlinearities, or reduction in system dimension.

In order to operate a filter, it is necessary to assume uniform boundedness of the process transition matrix, measurement matrix and the estimate error covariance matrix and the positive definiteness of the error covariance matrix as stated in previous chapters.

Our filter design for the above system model is:

A. Prediction stage.

$$\hat{\underline{x}}(k+1,k) = \Phi_c(k+1,k)\hat{\underline{x}}(k,k) + \underline{\Psi}_c(k) \ . \qquad \text{VII-5}$$

B. Correction stage.

$$\hat{\underline{x}}(k+1,k+1) = [I-K_c(k)H_c(k)]\hat{\underline{x}}(k+1,k) + K_c(k+1)\underline{z}(k+1) \ , \qquad \text{VII-6}$$

where the Kalman filter gain is

$$K_c(k+1) = P_c(k+1,k)H_c'(k+1)[H_c(k+1)P_c(k+1,k)H_c'(k+1) + R_c(k+1)]^{-1} \qquad \text{VII-7}$$

and

$$P_c(k+1,k) = \Phi_c(k+1,k)P_c(k,k)\Phi_c'(k+1,k) + \Gamma_c(k)Q_c(k+1)\Gamma_c'(k) \qquad \text{VII-8}$$

$$\begin{aligned} P_c(k+1,k+1) &= [I-K_c(k+1)H_c(k+1)]P_c(k+1,k) \\ &= [I-K_c(k+1)H_c(k+1)]P_c(k+1,k)[I-K_c(k+1)H_c(k+1)]' \\ &\quad + K_c(k+1)R_c(k+1)K_c'(k+1) \ , \end{aligned} \qquad \text{VII-9}$$

with initial estimate $x(0)$ and error covariance $P_c(0,0) = P(0)$ given.

We note that the filter operates on the real data $\underline{z}(k)$.

It is obvious that the computed matrix $P_c(k+1,k+1)$ is not the actual estimation error covariance matrix of the system. This is due to the difference from the real model. Neither does this filter satisfy the minimum variance criterion for the actual system. The actual estimation errors are

$$\tilde{\underline{x}}(k+1,k) = \underline{x}(k+1 - \hat{\underline{x}}(k+1,k) = \Phi_c(k+1,k)\underline{x}(k,k) + \Delta\Phi(k+1,k)$$

$$+ \Gamma(k)\underline{w}(k+1) + \Delta\underline{\Gamma}(k) \qquad \text{VII-10}$$

$$\tilde{\underline{x}}(k+1,k+1) = \underline{x}(k+1) - \hat{\underline{x}}(k+1,k+1) = [I-K_c(k+1)H_c(k+1)]$$

$$\hat{\underline{x}}(k+1,k) - K_c(k+1)\Delta H(k+1)\underline{x}(k+1) - K_c(k+1)\underline{v}(k+1). \qquad \text{VII-11}$$

Where

$$\Delta\Phi(k+1,k) = \Phi(k+1,k) - \Phi_c(k+1,k)$$

$$\Delta\underline{\Psi}(k) = \underline{\Psi}(k) - \underline{\Psi}_c(k)$$

$$\Delta H(k) = H(k) - H_c(k) \ .$$

The performance measure of the actual estimation error covariance matrix is

$$P(k+1,k+1) = E[\underline{\tilde{x}}(k+1,k+1)\underline{\tilde{x}}(k+1,k+1)'] \ ,$$

and

$$P(k+1,k) = E[\underline{\tilde{x}}(k+1,k)\underline{\tilde{x}}(k+1,k)'] \ .$$

Then, the recursive relation of error covariance is

$$P(k+1,k) = \Phi_c(k+1,k)P(k,k)\Phi_c'(k+1,k) + \Gamma(k)Q(k+1)\Gamma'(k)$$

$$+ \Delta\Phi(k+1,k)X(k)\Delta\Phi'(k+1,k) + \Delta\Phi(k+1,k)C(k,k)\Phi_c'(k+1,k)$$

$$+ \Phi_c(k+1,k)C'(k,k)\Delta\Phi'(k+1,k) + \Delta\Phi(k+1,k)m(k)\Delta\Psi'(k)$$

$$+ \Delta\Psi(k)m'(k)\Delta\Phi'(k+1,k) + \Delta\Psi(k)\Delta\Psi'(k)$$

$$+ \Phi_c(k+1,k)\Delta m(k,k)\Delta\Psi'(k) + \Delta\Psi(k)\Delta m'(k,k)\Phi_c(k+1,k) \qquad \text{VII-12}$$

and

$$P(k+1,k+1) = [I-K_c(k+1)H_c(k+1)]P(k+1,k)[I-K_c(k+1)H_c(k+1)]'$$

$$+ K_c(k+1)R(k+1)K_c'(k+1) - [I-K_c(k+1)H_c(k+1)]$$

$$C'(k+1,k)\Delta H'(k+1)K_c'(k+1) - K_c(k+1)\Delta H(k+1)C(k+1,k)$$

$$[I-K_c(k+1)H_c(k+1)]' + K_c\Delta H(k+1)X(k+1)\Delta H'(k+1)K_c'(k+1) \qquad \text{VII-13}$$

where

$$X(k+1) = E[\underline{x}(k+1)\underline{x}(k+1)']$$

$$C(k,k) = E[\underline{x}(k)\underline{\tilde{x}}(k,k)']$$

$$C(k+1,k) = E[\underline{x}(k+1)\underline{\tilde{x}}(k+1,k)']$$

$$m(k) = E[\underline{x}(k)]$$

$$\Delta m(k+1,k+1) = E[\underline{\tilde{x}}(k+1,k+1)] \ .$$

Together with their recursive relations:

$$X(k+1) = \Phi(k+1,k)X(k)\Phi'(k+1,k) + \Psi(k)\Psi'(k) + \Gamma(k)Q(k+1)\Gamma'(k)$$

$$+ \Phi(k+1,k)m(k)\Psi'(k) + \Psi(k)m'(k)\Phi'(k+1,k) \qquad \text{VII-14}$$

$$C(k+1,k) = \Phi(k+k,k)C(k,k)\Phi_c'(k+1,k) + \Phi(k+1,k)X(k)\Delta\Phi'(k+1,k)$$
$$+ \Gamma(k)Q(k+1)\Gamma'(k) + \Phi(k+1,k)m(k)\Delta\Psi'(k) + \Psi(k)m'(k)$$
$$\Delta\Phi'(k+1,k) + \Psi(k)\Delta m'(k,k)\Phi_c'(k+1,k) + \Psi(k)\Delta\Psi'(k) \ , \qquad \text{VII-15}$$

$$C(k+1,k+1) = C(k+1,k)[I-K_c(k+1)H_c(k+1)]'-X(k)\Delta H'(k+1)K_c'(k+1) \qquad \text{VII-16}$$

$$\underline{m}(k+1) = \Phi(k+1,k)\underline{m}(k) + \Psi(k) \ , \qquad \text{VII-17}$$

$$\Delta m(k+1,k+1) = [I-K_c(k+1)H_c(k+1)][\Phi_c(k+1,k)\Delta m(k,k)$$
$$+ \Delta\Phi'(k+1,k)m(k) + \Delta\Psi(k)] - K_c(k+1)\Delta H(k+1)m(k+1) \qquad \text{VII-18}$$

With the initial conditions:

$$P(0,0) = P(0); \quad X(0) = E(\underline{x}_o\underline{x}_o') = P(0) + \underline{\hat{x}}(0)\underline{\hat{x}}'(0)$$

$$C(0) = E[\underline{x}_o(\underline{x}_o-\underline{\hat{x}}(0))'] = P(0)$$

$$m(0) = E(x_o) = \hat{x}(0)$$

$$m(0,0) = 0 \ .$$

This general sensitivity analysis algorithm has been developed into a subroutine and tested on the reactor model to evaluate various effects on the filter performance due to different error categories. Various special cases of errors treated herein are described by Jazwinski [J2], however, we are trying to cover more completely and present a detailed discussion and its relation to adaptivity.

VII. 3 Error Sensitivity Due to Prior Statistics

In this section we consider error to be existant only in P(0), Q(k) and R(k). Substituting the relations:

$$\Delta\Phi(k+1,k) = 0; \quad \Delta\Psi(k) = 0$$

$$\Delta H(k) = 0 \text{ into the general algorithm.}$$

i.e. assume that the process transition and measurement matrix is identical to that of the filter model. We then have the following recursive relations for the error covariance.

For the actual error covariance:

$$P(k+1,k) = \Phi(k+1,k)P(k,k)\,\Phi'(k+1,k) + \Gamma(k)Q(k+1)\Gamma'(k)$$

$$P(k+1,k+1) = [I-K_c(k+1)H(k+1)]P(k+1,k)[I-K_c(k+1)H(k+1))'$$

$$+ K_c(k+1)R_c(k)K_c'(k+1) \quad .$$

The error covariance for the filter model is:

$$P_c(k+1,k) = \Phi_c(k+1,k)P_c(k,k)\Phi_c'(k+1,k) + \Gamma(k)Q_c(k+1)\Gamma'(k)$$

$$P_c(k+1,k+1) = [I-K_c(k+1)H(k+1)]P_c(k+1,k)[I-K_c(k+1)H(k+1)]'$$

$$+ K_c(k+1)R_c(k+1)K_c'(k+1)$$

Let

$$E(k+1,k+1) = P(k+1,k+1) - P_c(k+1,k+1)$$

$$E(k+1,k+1) = [I-K_c(k+1)H(k+1)]E(k+1,k)[I-K_c(k+1)H(k+1)]'$$

$$+ K_c(k+1)[R(k+1)-R_c(k+1)]K_c'(k+1) \quad .$$

Case I:

Suppose $Q(k+1) \leq Q_c(k+1)$ and $R(k+1) \leq R_c(k+1)$ for all k then, if $E(k+1,k) \leq 0$; it is clear that $E(k+1,k+1) \leq 0$.

Case II:

Suppose $Q(k+1) \geq Q_c(k+1)$ and $R(k+1) \geq R_c(k+1)$ for all k then, if $E(k,k-1) \geq 0$; it is clear that $E(k+1,k) \geq 0$ and $E(k+1,k+1) \geq 0$.

Application of the general algorithm to the chemical reactor model has shown good results and we can arrive at the following conclusions:

1. If $P(0)<P_c(0)$ and $Q(k)<Q_c(k)$, $R(k)<R_c(k)$ for all k then, $P(k+1,k)<P_c(k+1,k)$ and $P(k+1,k+1)<P_c(k+1,k+1)$ for all k. Therefore the actual error covariance matrix is bounded by the error covariance matrix based on the filter model as in the first case.

2. On the other hand, if the prior statistics lie in the second category (define), the filter error covariance matrix is bounded by the actual model beforehand, and we cannot gain any information in this case.

Therefore, it is preferable to set the prior statistics according to case I, since we may precompute $P_c(k+1,k+1)$ (satisfying the Riccati Equation) to determine conservative estimates of the actual P(0), Q(k) and R(k) which will give satisfactory filter performance.

Applying the error bound calculations in the previous chapters for the filter error covariance, we can predict the estimation error of an uncertain actual model.

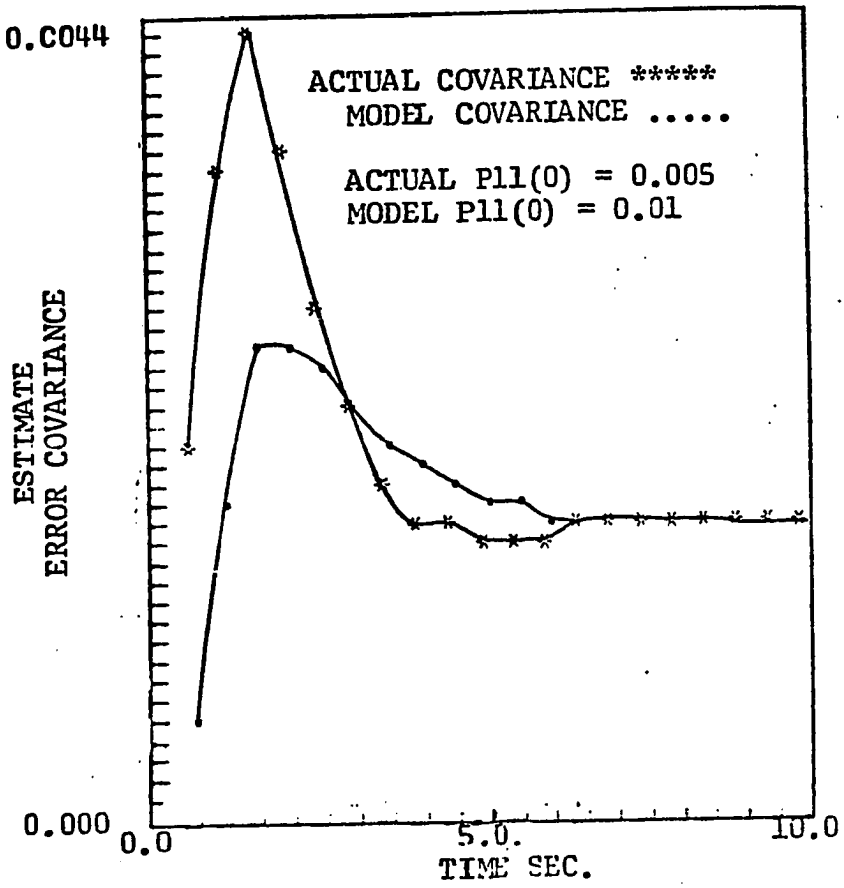


FIG. 7.1 Sensitivity Analysis
SECCND CRDER NCNLINEAR FILTER
SENSITIVITY ANALYSIS DUE TC PRICR STATISTICS
CONSTANT NCISE

TIME SEC.	ACTUAL P	MCDEL P
0.5	0.0015	-0.0001
1.0	0.0041	0.0018
1.5	C.CC46	C.CC29
2.0	0.0036	0.0028
2.5	0.0027	0.0024
3.0	C.0020	C.CC21
3.5	0.0017	0.0019
4.0	0.0015	0.0018
4.5	C.C015	0.CC18
5.0	0.0016	0.CC18
5.5	C.0016	0.0017
6.0	C.0016	0.CC17
6.5	0.0017	0.0017
7.C	C.CC17	0.CC17
7.5	0.0017	0.CC17
8.0	C.0017	0.0017
8.5	0.0017	0.CC17
9.0	C.0017	0.0017
9.5	0.0017	0.0017
10.0	0.0017	0.CC17

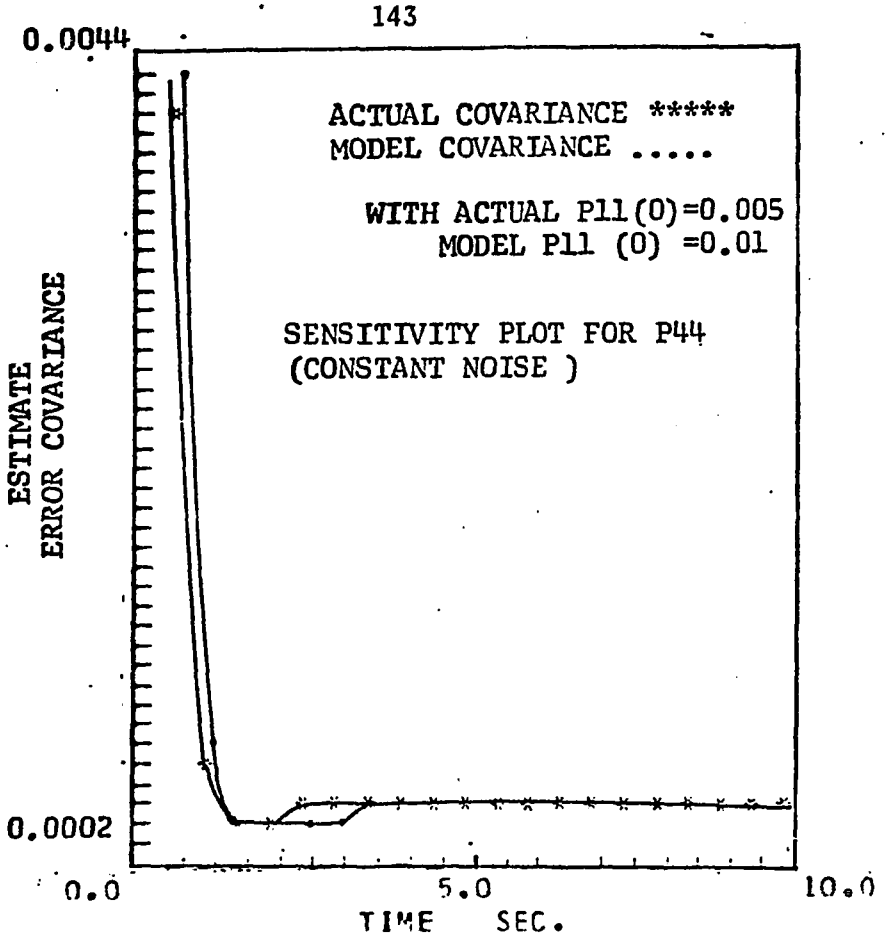


FIG. 7.2 ERROR SENSITIVITY DUE TO P(0)
NONLINEAR EXTENDED KALMAN FILTER
SENSITIVITY ANALYSIS DUE TO PRIOR STATISTICS

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0042	0.0044
1.0	0.0005	0.0006
1.5	0.0001	0.0001
2.0	0.0001	0.0001
2.5	0.0002	0.0001
3.0	0.0002	0.0002
3.5	0.0002	0.0002
4.0	0.0002	0.0002
4.5	0.0002	0.0002
5.0	0.0002	0.0002
5.5	0.0002	0.0002
6.0	0.0002	0.0002
6.5	0.0002	0.0002
7.0	0.0002	0.0002
7.5	0.0002	0.0002
8.0	0.0002	0.0002
8.5	0.0002	0.0002
9.0	0.0002	0.0002
9.5	0.0002	0.0002
10.0	0.0002	0.0002

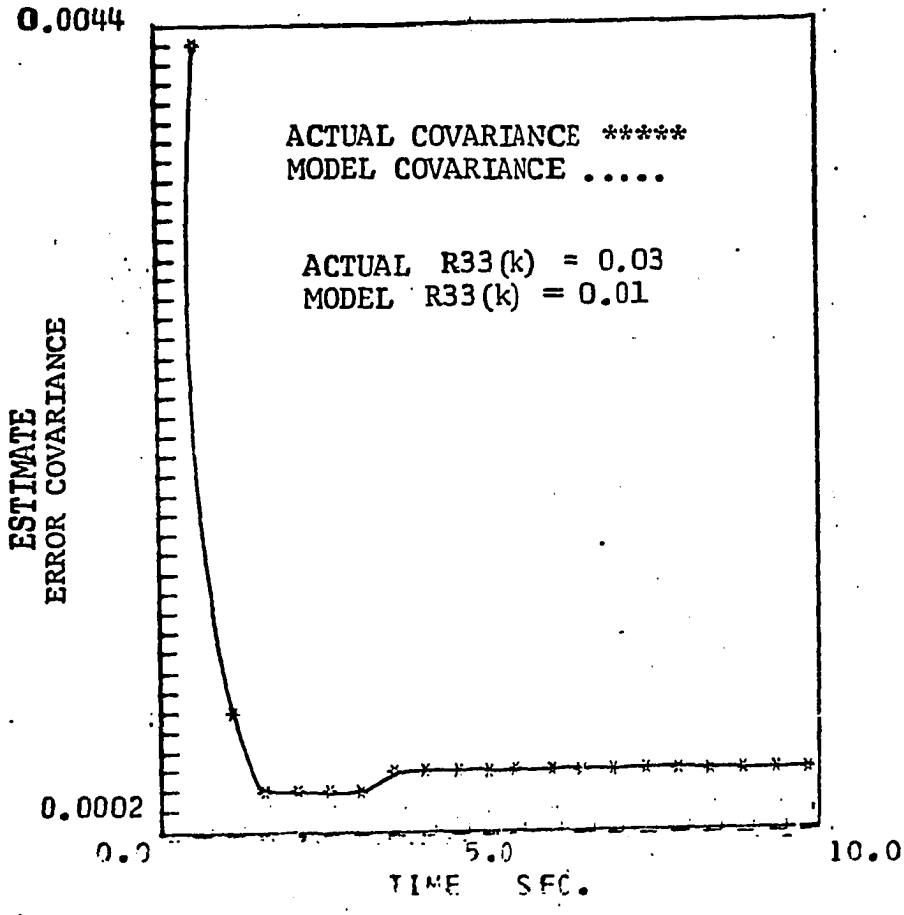


FIG. 7.3 SENSITIVITY DUE TO MEASUREMENT NOISE ERROR

NONLINEAR EXTENDED KALMAN FILTER
CONSTANT NOISE, SENSITIVITY PLOT FOR P44

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0044	0.0044
1.0	0.0006	0.0006
1.5	0.0001	0.0001
2.0	0.0001	0.0001
2.5	0.0001	0.0001
3.0	0.0002	0.0002
3.5	0.0002	0.0002
4.0	0.0002	0.0002
4.5	0.0002	0.0002
5.0	0.0002	0.0002
5.5	0.0002	0.0002
6.0	0.0002	0.0002
6.5	0.0002	0.0002
7.0	0.0002	0.0002
7.5	0.0002	0.0002
8.0	0.0002	0.0002
8.5	0.0002	0.0002
9.0	0.0002	0.0002
9.5	0.0002	0.0002
10.0	0.0002	0.0002

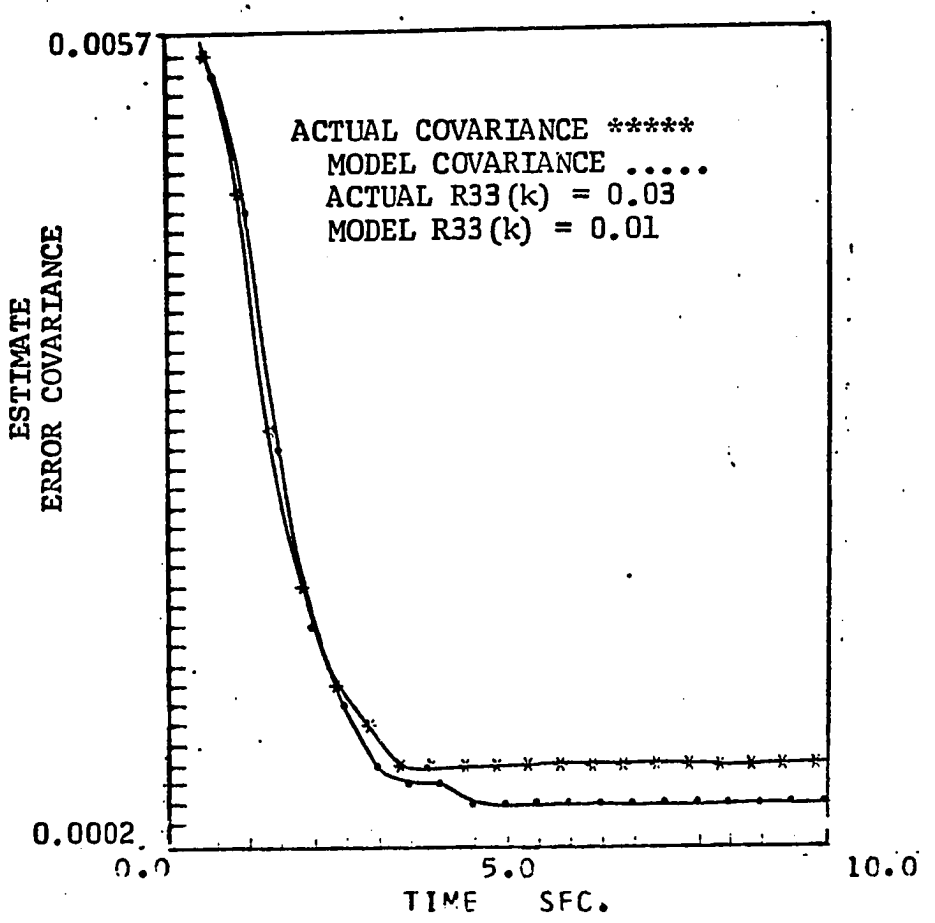


FIG. 7.4 SENSITIVITY DUE TO MEASUREMENT NOISE

NONLINEAR EXTENDED KALMAN FILTER
CONSTANT NOISE, SENSITIVITY PLOT FOR P33

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0057	0.0056
1.0	0.0046	0.0045
1.5	0.0029	0.0027
2.0	0.0017	0.0015
2.5	0.0010	0.0008
3.0	0.0007	0.0005
3.5	0.0005	0.0003
4.0	0.0004	0.0002
4.5	0.0004	0.0002
5.0	0.0004	0.0002
5.5	0.0004	0.0002
6.0	0.0004	0.0002
6.5	0.0004	0.0002
7.0	0.0004	0.0002
7.5	0.0004	0.0002
8.0	0.0004	0.0002
8.5	0.0004	0.0002
9.0	0.0004	0.0002
9.5	0.0004	0.0002
10.0	0.0004	0.0002

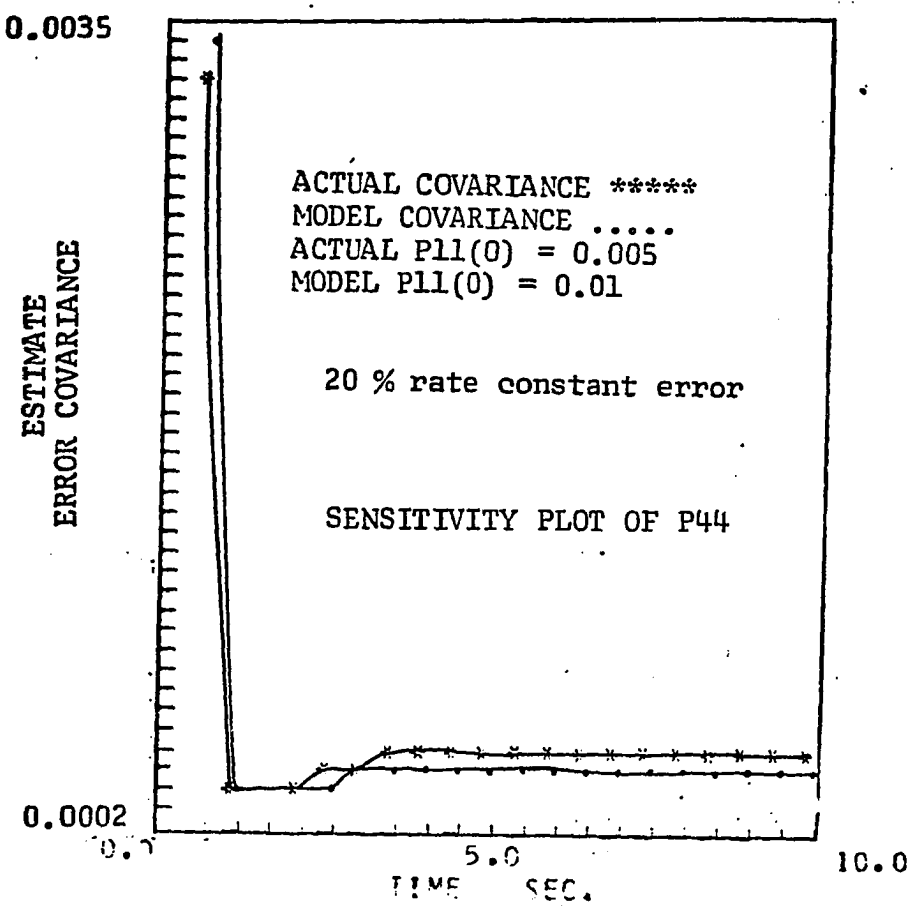


FIG. 7.5 SENSITIVITY DUE TO MODEL AND P(0) ERROR
NONLINEAR EXTENDED KALMAN FILTER
CONSTANT NOISE COMPENSATION

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0034	0.0035
1.0	0.0000	0.0001
1.5	0.0000	-0.0000
2.0	0.0001	0.0001
2.5	0.0002	0.0001
3.0	0.0002	0.0002
3.5	0.0002	0.0002
4.0	0.0002	0.0002
4.5	0.0002	0.0002
5.0	0.0002	0.0002
5.5	0.0002	0.0002
6.0	0.0003	0.0002
6.5	0.0003	0.0002
7.0	0.0003	0.0002
7.5	0.0003	0.0002
8.0	0.0003	0.0002
8.5	0.0003	0.0002
9.0	0.0003	0.0002
9.5	0.0003	0.0002
10.0	0.0003	0.0002

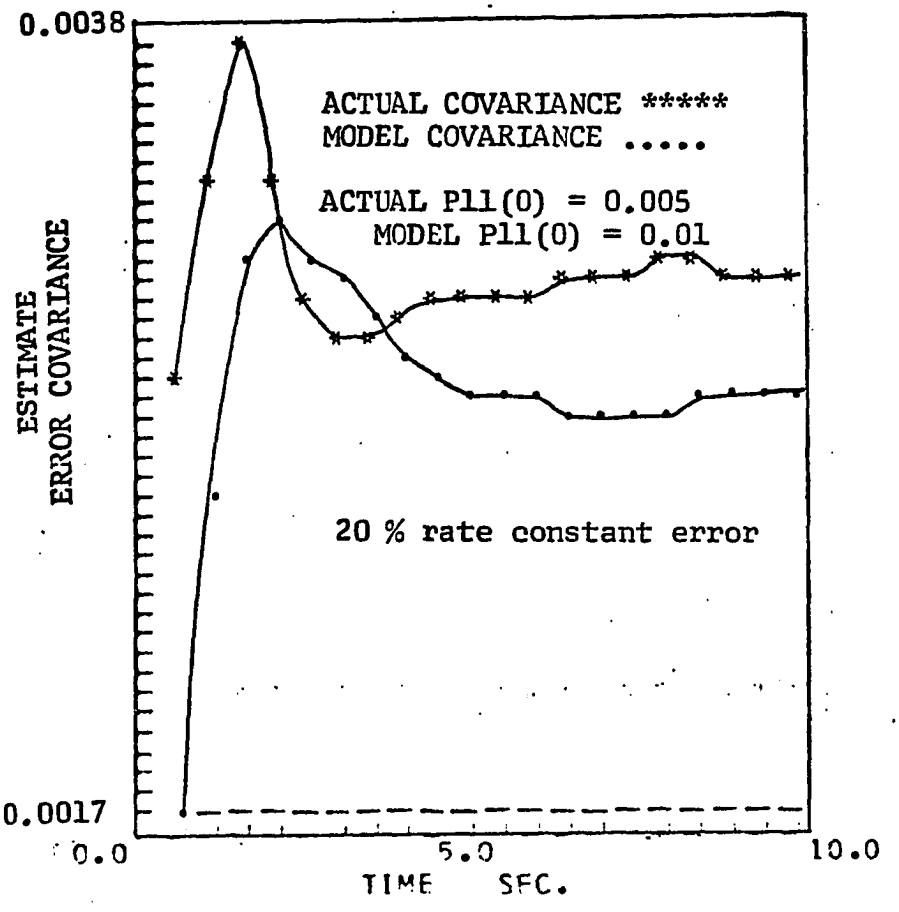


FIG. 7.6 SENSITIVITY DUE TO MODEL &PRIOR STATISTICS

NONLINEAR EXTENDED KALMAN FILTER
CONSTANT NOISE COMPENSATION
SENSITIVITY PLOT OF P11

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0018	-0.0007
1.0	0.0030	0.0011
1.5	0.0038	0.0025
2.0	0.0030	0.0027
2.5	0.0023	0.0026
3.0	0.0020	0.0024
3.5	0.0020	0.0022
4.0	0.0022	0.0020
4.5	0.0023	0.0019
5.0	0.0024	0.0018
5.5	0.0024	0.0017
6.0	0.0024	0.0017
6.5	0.0024	0.0017
7.0	0.0024	0.0017
7.5	0.0024	0.0017
8.0	0.0025	0.0017
8.5	0.0025	0.0017
9.0	0.0025	0.0017
9.5	0.0024	0.0017
10.0	0.0024	0.0017

A sample output for sensitivity analysis in this section can be found in Fig. 7.1 through Fig. 7.6.

VII.4 Sensitivity Analysis Due to Model Errors

Various model errors exist in the filtering process. Some are due to incomplete knowledge of the physical process leading to uncertainty in model parameters. A typical example in our reactor model is the uncertainty in the reactor rate constant or the reaction mechanism for complex reactions. In addition, some model errors are introduced deliberately for numerical simplification or when approximating a higher dimensional model by a lower order model. This saves both in computer storage and computation time. In addition, one can approximate a nonlinear model by a pseudo-linear model in order to obtain analytical results and utilize linear analysis.

In this section the sensitivity analysis of both the extended Kalman filter and the second order nonlinear filter is studied when the following effects are considered:

A. Uncertainty in the model parameters.

B. Model errors due to deliberate simplification in the model.

C. Model errors due to the combination of deliberate model simplification and error in the prior statistics.

1. Uncertainty in the model parameters:

An uncertainty in the model parameters will cause an undesirable model response and this may either degrade the filter performance or cause instability and divergence. In order to use our general

sensitivity analysis algorithm, we can treat the model uncertainty via $\Delta\Phi = \Phi_c - \Phi$. Φ is the transition matrix of the actual model when using accurate knowledge of the model parameter. Φ_c is the filter model transition matrix with an uncertain model parameter. We substituted this relationship into the algorithm and permanent offset and divergence between the actual and the computed error covariance was observed. This offset and divergence depended on the particular model response to the uncertainty in model parameters. This explains the reason for filter divergence due to uncertainty in model parameters.

The error sensitivity due to the uncertainty in model parameters is evaluated via $\Delta\Phi = \Phi_c - \Phi$ which can be achieved by substituting into the general algorithm:

$$P(k+1,k) = \Phi_c P_c(k,k)\Phi_c'(k+1,k) + \Gamma(k)Q_c(k+1)\Gamma'(k) + \Delta\Phi(k+1,k)X(k)\Delta\Phi'(k+1,k) + \Delta\Phi(k+1,k)C(k,k)\Phi_c'(k+1,k) + \Phi_c(k+1,k)C'(k,k)\Delta\Phi'(k+1,k) \qquad \text{VII-19}$$

$$P(k+1,k+1) = [I-K_c(k+1)H_c(k+1]P(k+1,k)[I-K_c(k+1)H_c(k+1)]' + K_c(k+1)R(k+1)K_c'(k+1) \quad . \qquad \text{VII-20}$$

Combining equations VII-19 and VII-20 we have on the right hand side, dropping the index for convenience:

$$P(k+1,k+1) = (I-K_cH)\Phi_c P\Phi_c(I-K_cH)' + F(k) \quad . \qquad \text{VII-21}$$

This is a linear inhomogeneous difference equation with the forcing term:

$$F(k) = (I-K_c H)(\Gamma Q_c \Gamma' + \Delta\Phi X \Delta\Phi' + \Delta\Phi C \Phi_c' + \Phi_c C \Delta\Phi)(I-K_c H) + K_c R K_c' \qquad \text{VII-22}$$

where X(k) and C(k,k) are as defined before. Note that if $\Delta\Phi = 0$, all terms with $\Delta\Phi$ in VII-22 drop out, and we have

$$F(k) = K_c R K_c' + (I-K_c H)(\Gamma Q_c \Gamma') \quad .$$

Substitute this into Equation VII-21,

$$P(k+1,k+1) = (I-K_c H)(\Phi_c P \Phi_c')(I-K_c H)' + K_c R K_c' + (I-K_c H)(\Gamma Q \Gamma')(I-K_c H)'$$

$$= (I-K_c H)(\Phi P \Phi' + \Gamma Q \Gamma')(I-K_c H)' + K_c R K_c' \quad .$$

This is exactly the same as the filter model error covariance. The sensitivity matrix as defined by

$$DP(k+1,k+1) = P(k+1,k+1) - P_c(k+1,k+1) \quad .$$

Therefore, if $\Delta\Phi = 0$, we will have $DP(k+1,k+1) = 0$ or $P(k+1,k+1) = P_c(k+1,k+1)$.

2. Model errors due to deliberate simplification:

As mentioned before, deliberate model simplification has been used very often in simulation and filtering realization and usually takes the form of replacing a higher order complex model by a lower dimensional simpler one. Here again, we can use the general computer program to make a sensitivity analysis via either $\Phi = \Phi_c - \Phi$ or $\Delta\Psi = \Psi_c - \Psi$ to take care of the errors either in reduction in dimension or non-linearity. In this work, we evaluate the sensitivity due to integration

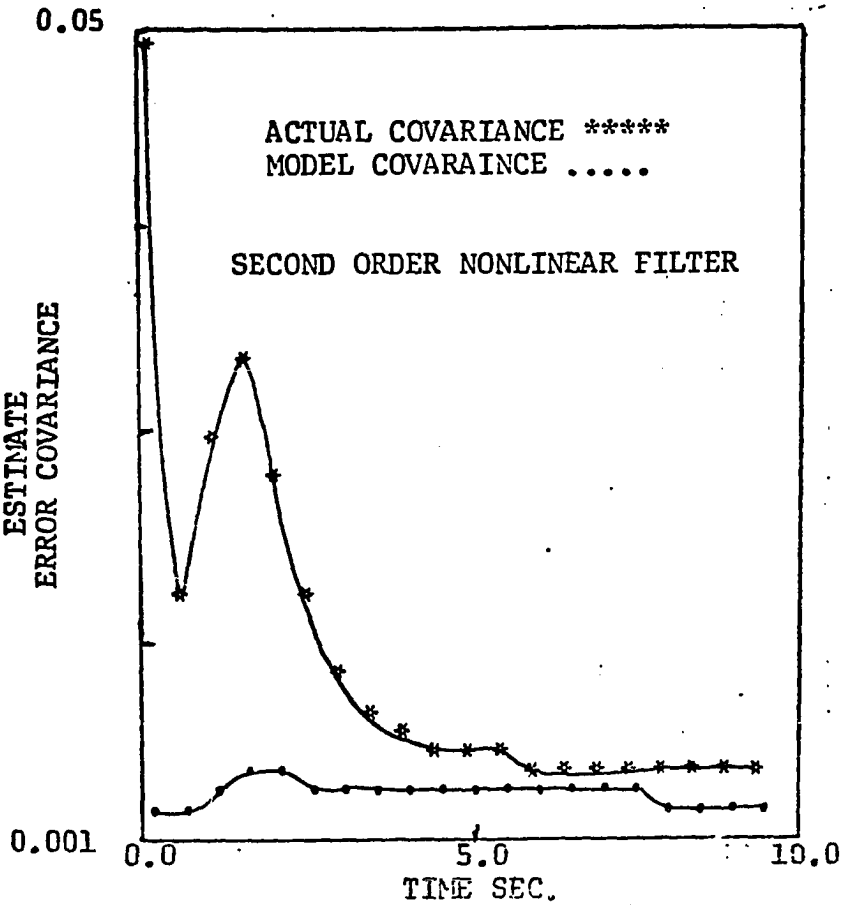


FIG. 7.7 SENSITIVITY ANALYSIS PLOT FOR P11, USING HAUNG'S ADAPTIVE I TO COMPENSATE FOR EULER'S INTEGRATION ERROR AGAINST RUNGE-KUTTA MERSON

TIME SEC.	ACTUAL P	MODEL P
0.50	0.0518	0.0010
1.00	0.0158	0.0008
1.50	0.0259	0.0027
2.00	0.0310	0.0034
2.50	0.0235	0.0029
3.00	0.0154	0.0027
3.50	0.0104	0.0025
4.00	0.0078	0.0024
4.50	0.0063	0.0022
5.00	0.0053	0.0020
5.50	0.0047	0.0019
6.00	0.0042	0.0018
6.50	0.0039	0.0017
7.00	0.0036	0.0017
7.50	0.0034	0.0016
8.00	0.0032	0.0015
8.50	0.0031	0.0015
9.00	0.0030	0.0014
9.50	0.0030	0.0014
10.00	0.0029	0.0013

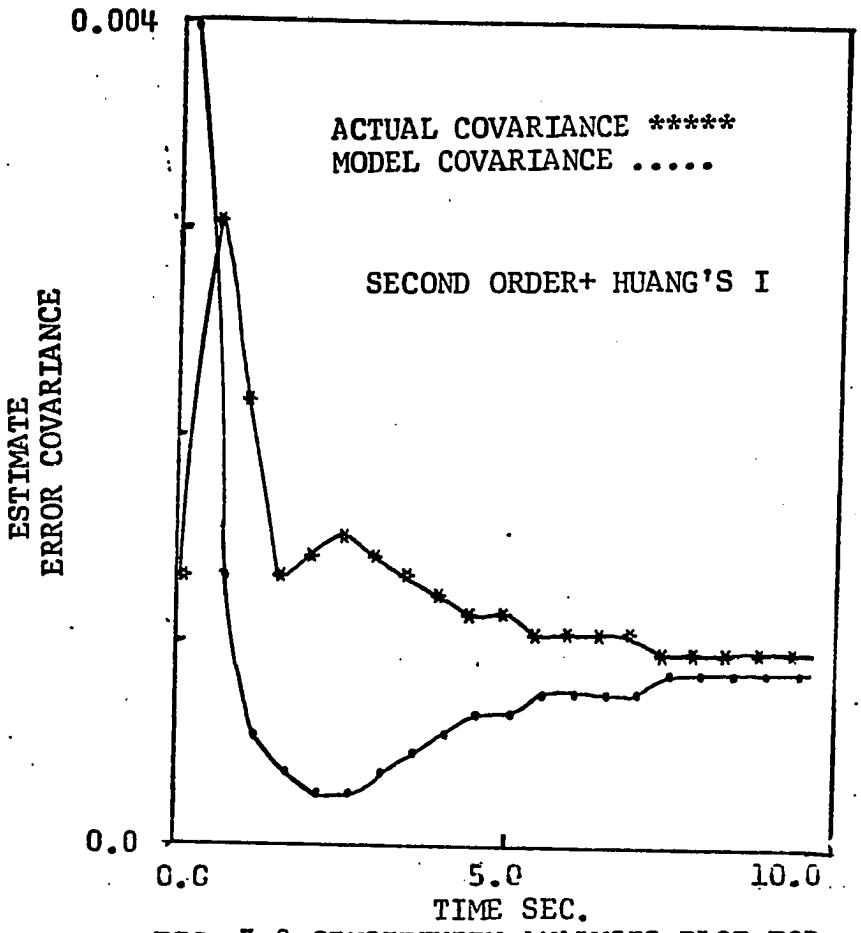


FIG. 7.8 SENSITIVITY ANALYSIS PLOT FOR P22,USING HUANG'S ADAPTIVE I TO COMPENSATE FOR EULER'S INTEGRATION ERROR AGAINST RUNGE-KUTTA MERSON

TIME SEC.	ACTUAL P	MODEL P
0.50	0.0012	0.0039
1.00	0.0030	0.0011
1.50	0.0020	0.0004
2.00	0.0011	0.0001
2.50	0.0012	0.0001
3.00	0.0013	0.0001
3.50	0.0013	0.0002
4.00	0.0012	0.0003
4.50	0.0011	0.0003
5.00	0.0010	0.0004
5.50	0.0009	0.0005
6.00	0.0009	0.0005
6.50	0.0009	0.0005
7.00	0.0008	0.0006
7.50	0.0008	0.0006
8.00	0.0008	0.0006
8.50	0.0008	0.0006
9.00	0.0008	0.0006
9.50	0.0008	0.0006
10.00	0.0007	0.0006

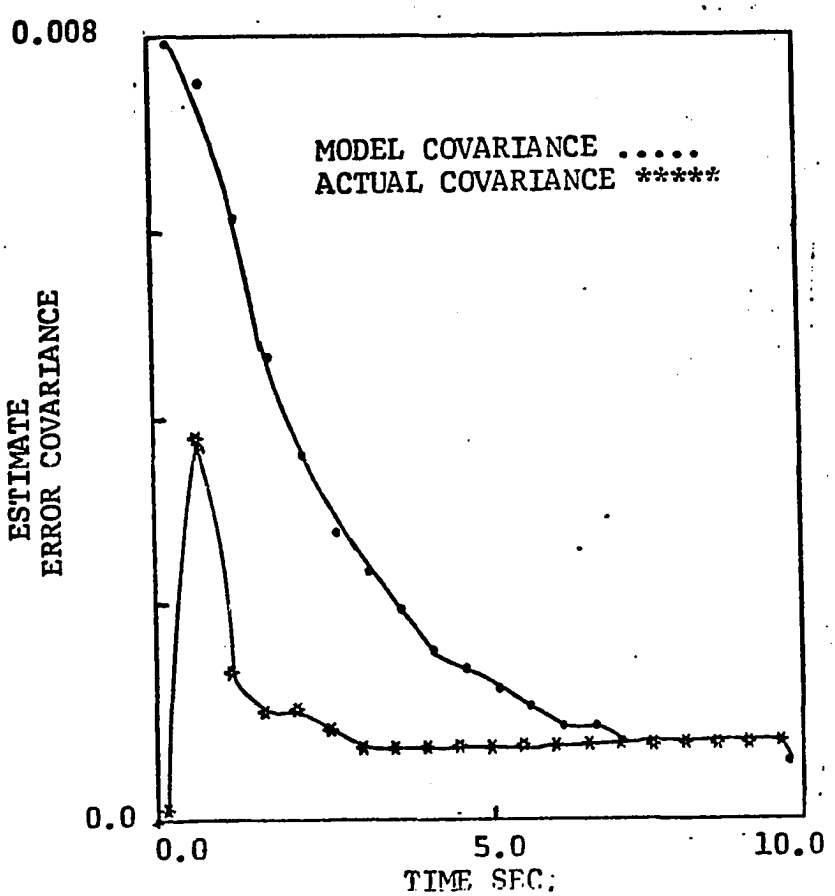


FIG. 7.9 SENSITIVITY ANALYSIS PLOT FOR P33, USING HUANG'S ADAPTIVE I TO COMPENSATE FOR EULER'S INTEGRATION ERROR AGAINST RUNGE-KUTTA MERSON NONLINEAR SECOND ORDER FILTER

TIME SEC.	ACTUAL P	MODEL P
0.50	-0.0001	0.0080
1.00	0.0038	0.0076
1.50	0.0014	0.0060
2.00	0.0009	0.0046
2.50	0.0009	0.0036
3.00	0.0007	0.0029
3.50	0.0006	0.0023
4.00	0.0005	0.0019
4.50	0.0005	0.0015
5.00	0.0005	0.0013
5.50	0.0005	0.0011
6.00	0.0005	0.0009
6.50	0.0005	0.0008
7.00	0.0005	0.0007
7.50	0.0005	0.0006
8.00	0.0005	0.0005
8.50	0.0005	0.0005
9.00	0.0005	0.0004
9.50	0.0005	0.0004
10.00	0.0005	0.0004

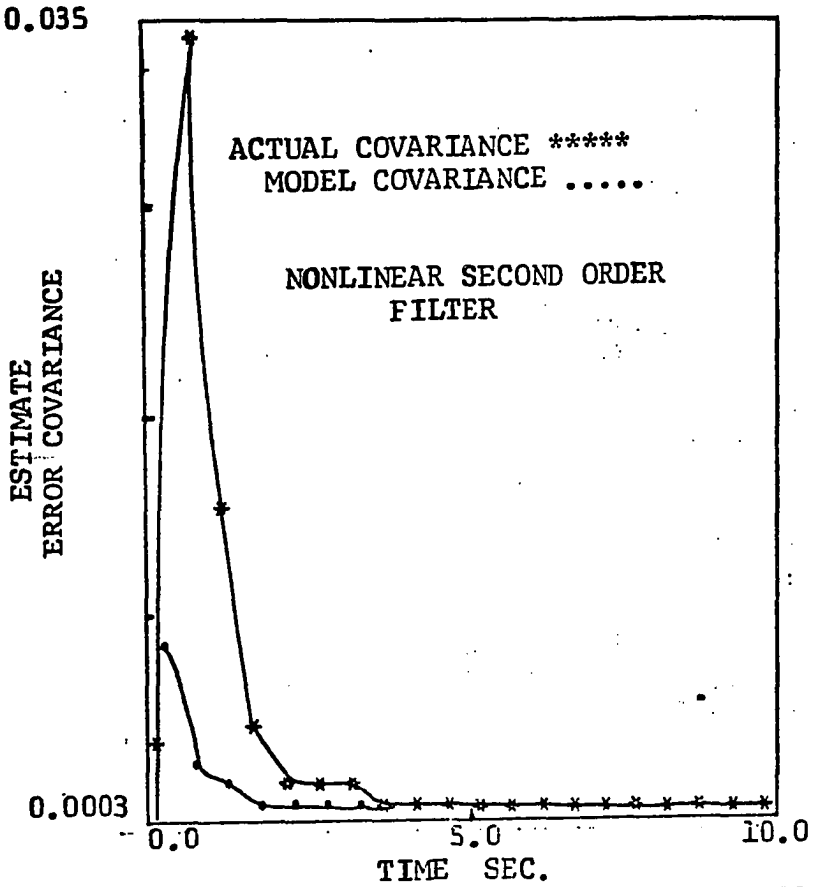


FIG. 7.10 SENSITIVITY ANALYSIS PLOT FOR P44, USING HUANG'S ADAPTIVE I TO COMPENSATE FOR EULER'S INTEGRATION ERROR AGAINST RUNGE-KUTTA MERSON

TIME SEC.	ACTUAL P	MODEL P
0.50	0.0025	0.0069
1.00	0.0343	0.0015
1.50	0.0132	0.0009
2.00	0.0031	0.0004
2.50	0.0010	0.0002
3.00	0.0006	0.0001
3.50	0.0005	0.0000
4.00	0.0004	-0.0000
4.50	0.0003	-0.0000
5.00	0.0003	-0.0000
5.50	0.0002	-0.0000
6.00	0.0002	-0.0000
6.50	0.0002	0.0000
7.00	0.0002	0.0000
7.50	0.0002	0.0000
8.00	0.0002	0.0000
8.50	0.0002	0.0000
9.00	0.0002	0.0000
9.50	0.0002	0.0000
10.00	0.0002	0.0000

error (using simple Euler's integration) via $\Delta\Phi$ as derived in A, while Fig. 7.7 to Fig. 7.10 shows the rapid disappearing of the difference by using adaptive techniques introduced by this work.

3. Errors due to the combination of prior statistics and model simplification.

A combination of errors in prior statistics and model simplification may also degrade the filter performance. The sensitivity due to these errors can also be studied by our general algorithm. Here, we treat again the model simplification via $\Delta\Psi = \Psi_c - \Psi$. In this case the recursive relation for the error covariance is:

$$P(k+1,k) = \Phi(k+1,k)P(k,k)\Phi'(k+1,k) + \Gamma(k)Q(k+1)\Gamma'(k)$$
$$+ \Delta\Psi(k)\Delta\Psi'(k) + \Phi(k+1,k)\Delta m(k)\Delta\Psi'(k) + \Delta\Psi(k)\Delta m'(k,k)\Phi'(k+1,k)$$

$$P(k+1,k+1) = [I-K_c(k+1)H(k+1)]P(k+1,k)[I-K_c(k+1)H(k+1)]'$$
$$+ K_c(k)R(k+1)K_c'(k+1) \cdot .$$

Combining these equations

$$P(k+1,k+1) = [I-K_c(k+1)H(k+1)]\Phi(k+1,k)P(k,k)\Phi'(k+1,k)$$
$$\cdot[I-K_c(k+1)H(k+1)]' + F(k) \quad .$$

Where F(k) is

$$F(k) = [I-K_c(k+1)H(k+1)][\Gamma(k)Q(k+1)\Gamma'(k)+\Delta\Psi(k)\Delta\Psi'(k)$$
$$+ \Phi(k+1,k)\Delta m(k,k)\Delta\Psi'+\Delta\Psi(k)\Delta m'(k,k)\Phi'(k+1,k)]$$
$$\cdot[I-K_c(k+1)H(k+1)]' + K_c(k+1)R(k+1)K_c'(k+1)$$

$$\Delta m(k+1,k+1) = [I-K_c(k+1)H(k+1)]\Phi(k+1,k)\Delta m(k,k) + f(k)$$

$$f(k) = [I-K_c(k+1)H(k+1)]\Delta\Psi(k) \ .$$

Let

$$\Psi_c(k+1,k) = [I-K_c(k+1)H(k+1)]\Phi(k+1,k) \ .$$

The following theorem is due to Price [P3]. If the dynamic system is uniformly completely observable and controlable and if F(k) is uniformly bounded and P(0) is bounded, then the error covariance matrix P(k,k) is uniformly bounded for all k. Since the solution for P(k,k) is:

$$P(k,k) = \Psi(k,0)P(0)\Psi_c'(k,0) + \sum_{i=0}^{k-1} \Psi_c(k,i+1)F(i)\Psi_c'(k,i+1) \ ,$$

the uniform boundedness of the error covariance matrix assures us that the filtering estimation error falls below the bound. Therefore we can set up the filter model such that its error covariance matrix stays within our acceptable region.

We can see that the boundness of the forcing term of the covariance equation requires the boundness of X(k), C(k,k) m(k) and m(k,k). These are all related to the uniformly asymptotic stability of the actual system dynamics. This is essentially assuming uniformly asymptotically stable system dynamics.

VII.5 Error Sensitivity Due to Measurement Error

We will discuss two types of incorrect measurement errors in this section using our general algorithm.

A. Incorrect measurement statistics.

B. Incorrect measurement matrix in the case of linear or incorrect measurement function in the nonlinear case.

1. Incorrect measurement statistics:

Here, we are concerned with the incorrect measurement noise covariance, R(k). Assuming everything of the filter model except the measurement noise is the same as the actual model, then substituting into our general algorithm:

$$P(k+1,k) = \Phi_c(k+1,k)P(k,k)\Phi_c'(k+1,k) + \Gamma(k)Q(k+1)\Gamma'(k)$$

$$P(k+1,k+1) = [I-K_c(k+1)H(k+1)]P(k+1,k)[I-K_c(k+1)H(k+1)]' + K_c(k+1)R(k+1)K_c'(k+1) \quad .$$

Therefore, similar results can be obtained as in the section of sensitivity due to prior statistics. The boundedness of the error covariance is also related to the asymptotic stability of the process transition matrix.

2. Incorrect measurement matrix:

An incorrect measurement matrix in the linear case or an incorrect measurement matrix in the nonlinear case can also degrade the filter performance. The sensitivity analysis study of these effects can also be achieved by our general algorithm using $\Delta H(k) = H(k)-H_c(k)$; here $H(k)$ is the actual measurement matrix while $H_c(k)$ is the filter model incorrect measurement matrix.

Substituting this relationship into the general algorithm we have:

$$P(k+1,k) = \Phi(k+1,k)P(k,k)\Phi_c'(k+1,k) + \Gamma(k)Q(k+1)\Gamma'(k)$$

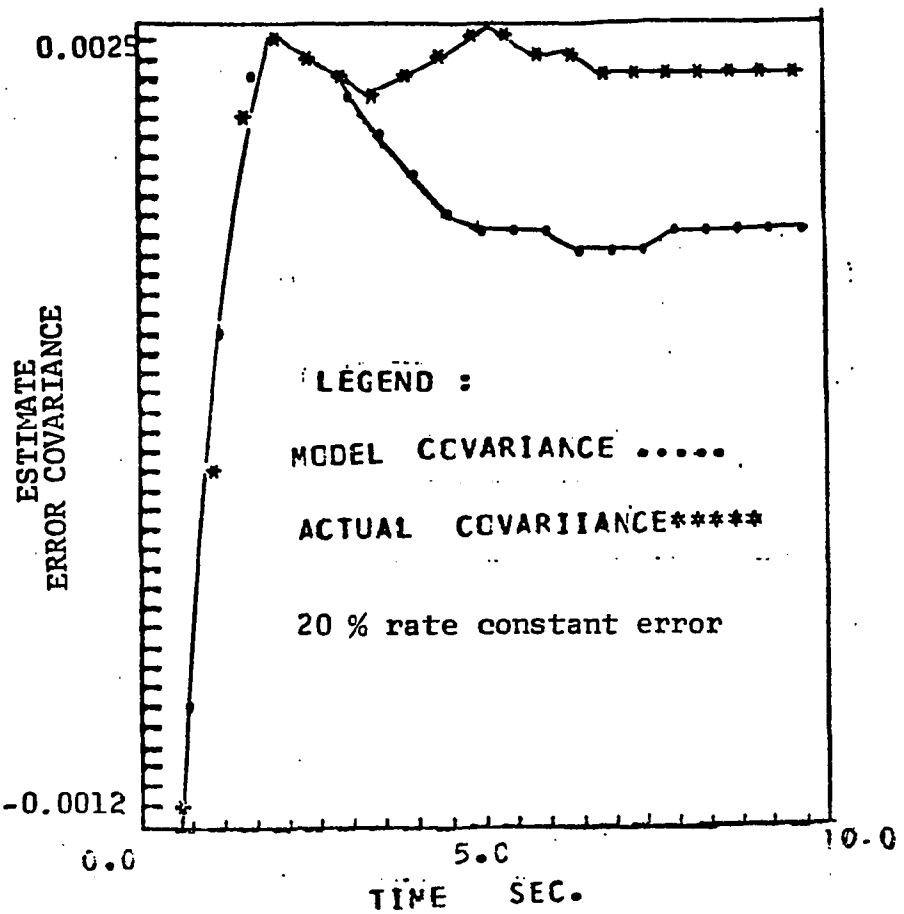


FIG. 7.11 FILTER DIVERGENCE
SECOND ORDER NONLINEAR FILTER
SENSITIVITY ANALYSIS DUE TO MODEL ERROR
CONSTANT NOISE PLOT FOR P11

TIME SEC.	ACTUAL P	MODEL P
0.5	-0.0012	-0.0007
1.0	0.0005	0.0012
1.5	0.0023	0.0025
2.0	0.0027	0.0027
2.5	0.0026	0.0026
3.0	0.0025	0.0024
3.5	0.0024	0.0022
4.0	0.0025	0.0020
4.5	0.0026	0.0018
5.0	0.0026	0.0017
5.5	0.0027	0.0017
6.0	0.0026	0.0017
6.5	0.0026	0.0016
7.0	0.0025	0.0016
7.5	0.0025	0.0016
8.0	0.0025	0.0017
8.5	0.0025	0.0017
9.0	0.0025	0.0017
9.5	0.0025	0.0017
10.0	0.0025	0.0017

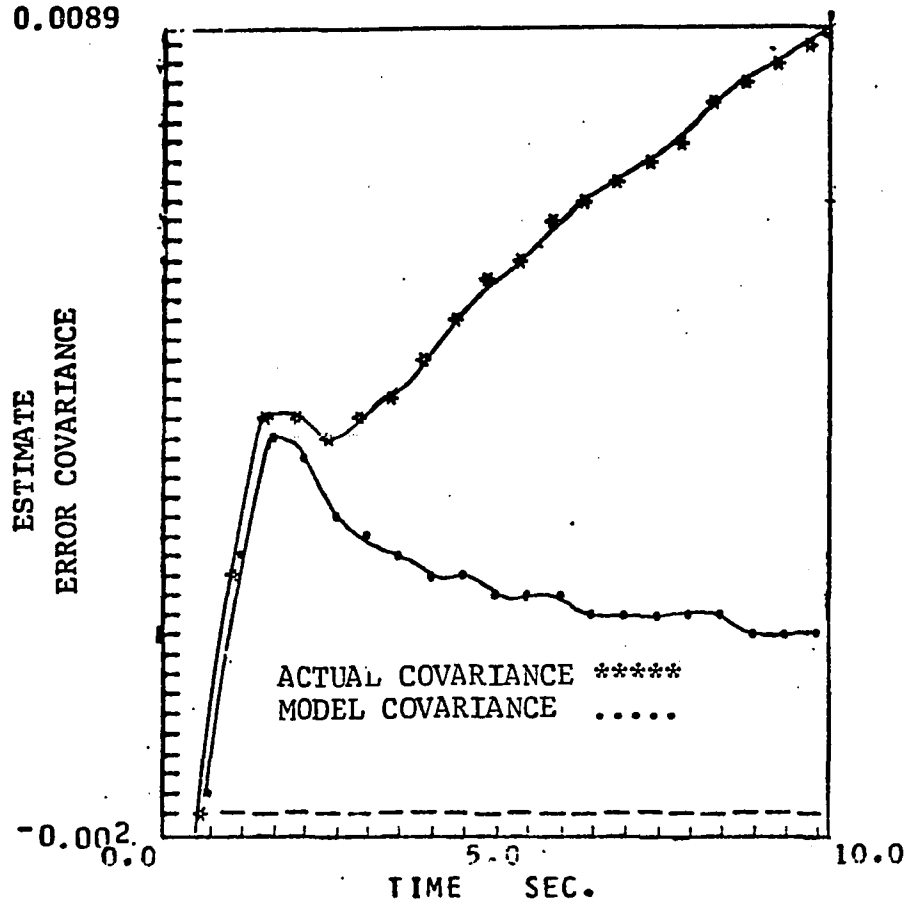


FIG. 7.12 FILTER DIVERGENCE
SECCND CRDER NONLINEAR FILTER
SENSITIVITY ANALYSIS DUE TC MODEL ERROR
NO NOISE CCMPENSATICN

TIME SEC.	ACTUAL P	MCDEL P
C.5	-C.0023	-0.CC20
1.0	0.0011	0.CC14
1.5	C.0032	0.0029
2.0	C.0031	0.CC25
2.5	0.0030	0.0020
3.0	C.0031	0.0015
3.5	C.CC35	0.CC12
4.0	0.0040	0.CC10
4.5	0.0046	0.0009
5.0	C.0051	0.CCC8
5.5	0.0055	0.0007
6.0	C.CC59	0.CCC6
6.5	C.0062	0.CCC5
7.0	0.0065	C.CCC5
7.5	C.CC69	0.CC04
8.0	0.0072	0.CCC4
8.5	0.0075	0.0003
9.0	C.CC78	0.CC03
9.5	C.CC82	0.CCC3
10.0	0.0085	C.CC02

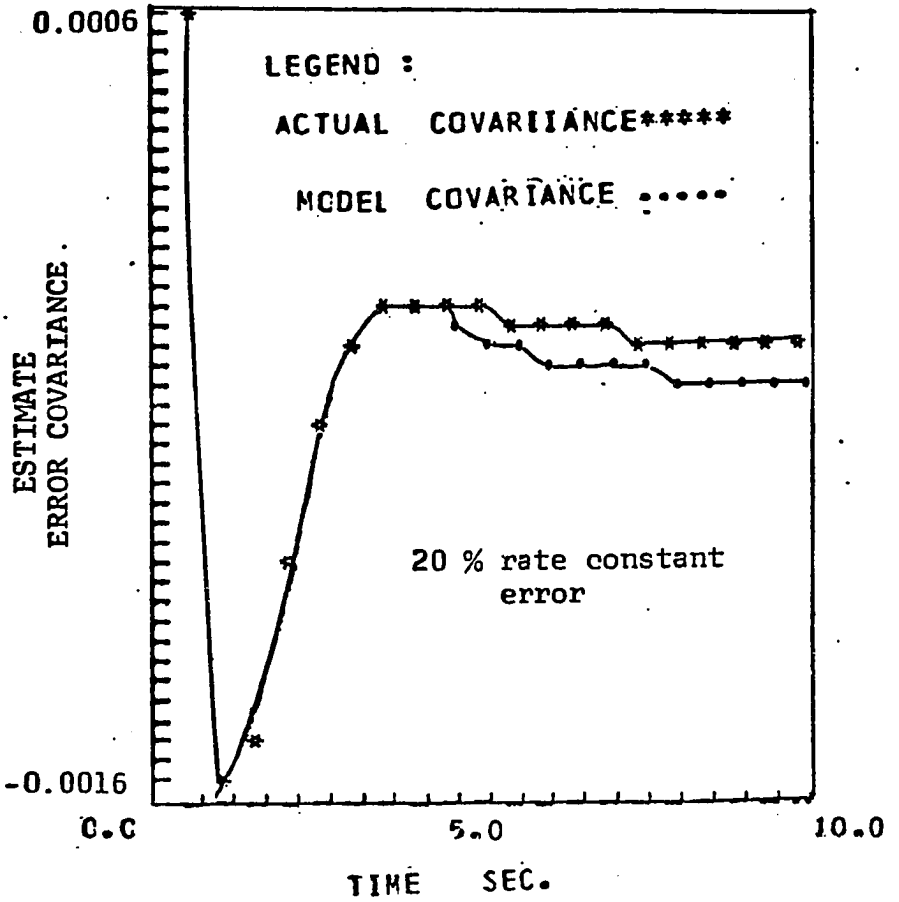


FIG. 7.13 FILTER DIVERGENCE
SECOND ORDER NONLINEAR FILTER
SENSITIVITY ANALYSIS DUE TO MODEL ERROR
CONSTANT NOISE PLOT FOR P22

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0016	0.0016
1.0	-0.0010	-0.0010
1.5	-0.0009	-0.0008
2.0	-0.0003	-0.0002
2.5	0.0002	0.0002
3.0	0.0005	0.0005
3.5	0.0006	0.0006
4.0	0.0006	0.0006
4.5	0.0006	0.0006
5.0	0.0006	0.0005
5.5	0.0006	0.0005
6.0	0.0006	0.0004
6.5	0.0006	0.0004
7.0	0.0005	0.0004
7.5	0.0005	0.0004
8.0	0.0005	0.0004
8.5	0.0005	0.0004
9.0	0.0005	0.0004
9.5	0.0005	0.0004
10.0	0.0005	0.0004

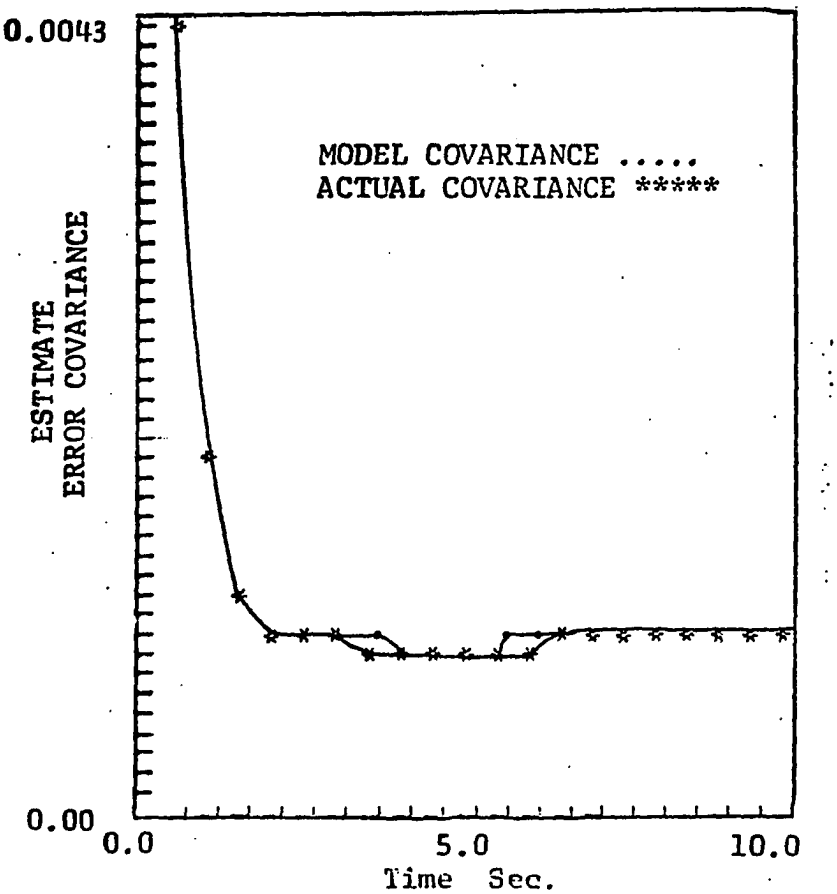


FIG. 7.14 APPROXIMATING A 4 thORDER MODEL
BY A SECOND ORDER MODEL
SENSITIVITY ANALYSIS DUE TO MODEL ERROR
PLOT FOR REACTANT CONCENTRATION
ERROR COVARIANCE MATRIX USING HUANGS III

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0043	0.0043
1.0	0.0019	0.0019
1.5	0.0011	0.0011
2.0	0.0009	0.0009
2.5	0.0009	0.0009
3.0	0.0008	0.0008
3.5	0.0008	0.0008
4.0	0.0008	0.0008
4.5	0.0008	0.0008
5.0	0.0008	0.0008
5.5	0.0008	0.0008
6.0	0.0008	0.0008
6.5	0.0008	0.0009
7.0	0.0008	0.0009
7.5	0.0008	0.0009
8.0	0.0009	0.0009
8.5	0.0009	0.0009
9.0	0.0009	0.0009
9.5	0.0009	0.0009
10.0	0.0009	0.0009

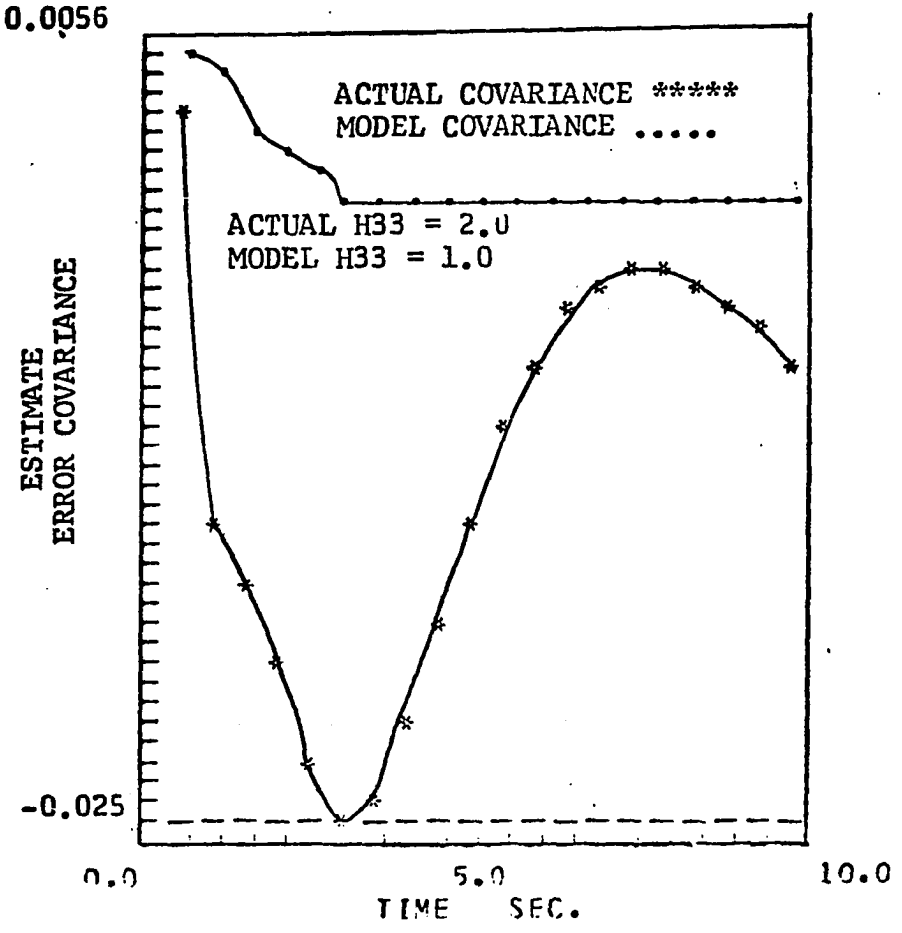


FIG. 7.15 SENSITIVITY DUE TO MEASUREMENT ERROR

NONLINEAR EXTENDED KALMAN FILTER
SENSITYVITY ANALYSIS DUE TO MEASUREMENT
MATRIX ERROR,SENSITIVITY PLOT OF P33

TIME SEC.	ACTUAL P	MODEL P
0.5	0.0036	0.0056
1.0	-0.0130	0.0045
1.5	-0.0154	0.0027
2.0	-0.0190	0.0015
2.5	-0.0230	0.0008
3.0	-0.0251	0.0005
3.5	-0.0245	0.0003
4.0	-0.0215	0.0002
4.5	-0.0174	0.0002
5.0	-0.0132	0.0002
5.5	-0.0095	0.0002
6.0	-0.0067	0.0002
6.5	-0.0047	0.0002
7.0	-0.0036	0.0002
7.5	-0.0032	0.0002
8.0	-0.0033	0.0002
8.5	-0.0039	0.0002
9.0	-0.0047	0.0002
9.5	-0.0056	0.0002
10.0	-0.0067	0.0002

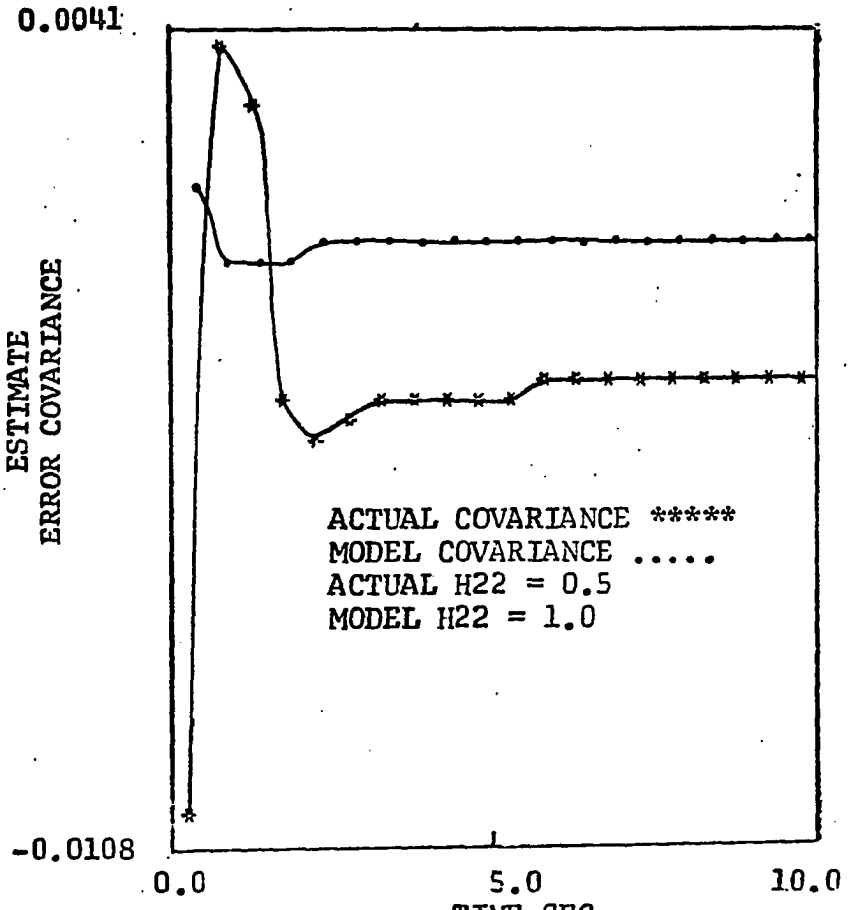


FIG. 7.16 SENSITIVITY ANALYSIS DUE TO ERROR IN MEASUREMENT MATRIX
USE NONLINEAR EXTENDED KALMAN FILTER
SENSITIVITY ANALYSIS PLOT OF P22

TIME SEC.	ACTUAL P	MODEL P
0.50	-0.0108	0.0016
1.00	0.0041	-0.0002
1.50	0.0030	-0.0003
2.00	-0.0027	-0.0001
2.50	-0.0036	0.0001
3.00	-0.0032	0.0003
3.50	-0.0030	0.0003
4.00	-0.0029	0.0003
4.50	-0.0028	0.0003
5.00	-0.0027	0.0003
5.50	-0.0026	0.0003
6.00	-0.0026	0.0003
6.50	-0.0025	0.0003
7.00	-0.0025	0.0004
7.50	-0.0025	0.0004
8.00	-0.0024	0.0004
8.50	-0.0024	0.0004
9.00	-0.0024	0.0004
9.50	-0.0023	0.0004
10.00	-0.0023	0.0004

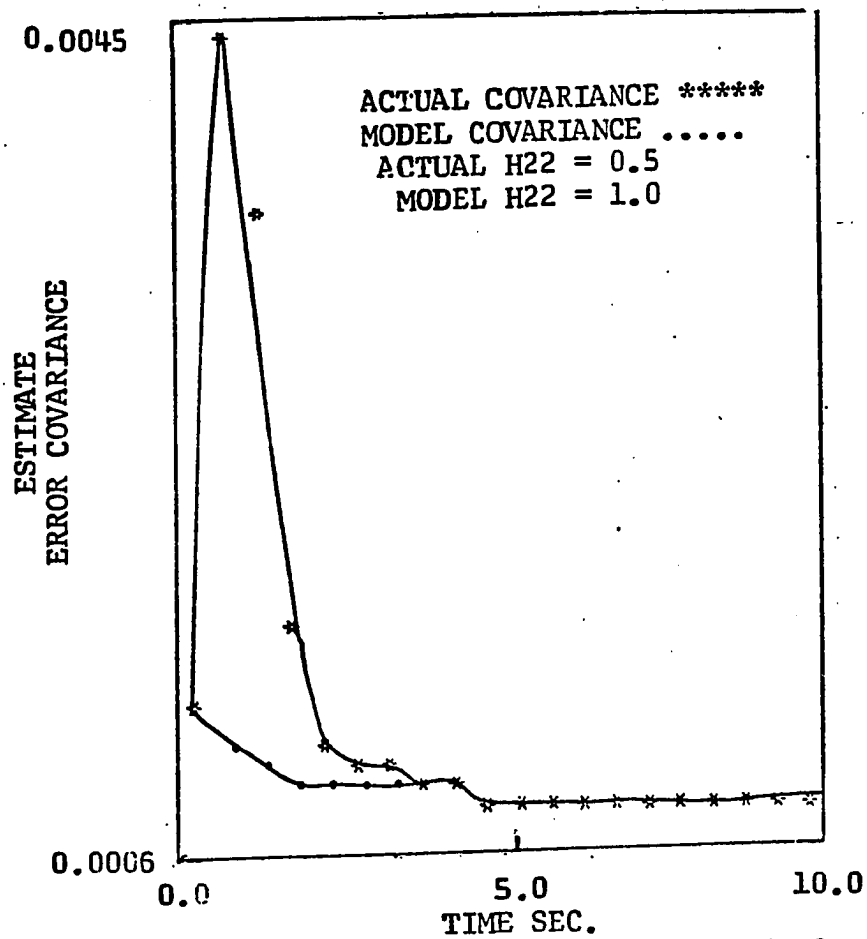


FIG. 7.17 SENSITIVITY ANALYSIS DUE TO MEASUREMENT MATRIX ERROR
NONLINEAR EXTENDED KALMAN FILTER
SENSITIVITY ANALYSIS PLOT FOR P44

TIME SEC.	ACTUAL P	MODEL P
0.50	0.0043	0.0044
1.00	0.0275	0.0025
1.50	0.0212	0.0019
2.00	0.0070	0.0016
2.50	0.0026	0.0014
3.00	0.0020	0.0012
3.50	0.0017	0.0011
4.00	0.0014	0.0010
4.50	0.0011	0.0010
5.00	0.0010	0.0009
5.50	0.0009	0.0009
6.00	0.0008	0.0008
6.50	0.0008	0.0008
7.00	0.0007	0.0008
7.50	0.0007	0.0008
8.00	0.0007	0.0007
8.50	0.0007	0.0007
9.00	0.0007	0.0007
9.50	0.0006	0.0007
10.00	0.0006	0.0007

$$P(k+1,k+1) = [I-K_c(k+1)H(k+1)P(k+1,k)[I-K_c(k+1)H(k+1)]'$$

$$+ K_c(k+1)R(k+1)K_c'(k+1) - \underline{K_c(k+1)\Delta H(k+1)C'(k+1,k)}$$

$$\underline{[I-K_c(k+1)H(k+1)]' - [I-K_c(k+1)H_c(k+1)]}$$

$$\underline{C'(k+1,k)\Delta H'(k+1)K_c'(k+1)}$$

$$+ \underline{K_c(k+1)\Delta H(k)X(k+1)\Delta H'(k+1)K_c'(k+1)} \ .$$

This is also a linear inhomogeneous difference equation with the underlined part as the forcing terms. Thus, if $\Delta H = 0$, the sensitivity matrix is equal to zero.

In our reactor model, an incorrect measurement element on the jacket temperature has no effect on the sensitivity of the reactant concentration. Even a constant offset on the jacket temperature error covariance matrix does not offset the concentration. This is due to the very light interaction between the jacket temperature and the reactant concentration. This is shown in Fig. 7.16 and Fig. 7.17.

VII.6 Relation to Adaptivity

The purpose of applying an adaptive or error compensation technique during the filtering process is to minimize any error due to imprecise knowledge of prior statistics, measurement errors or uncertain model parameters. It is intuitively clear that the adaptive technique will adjust the filter such that the error covariance of the filter is close to that of the actual model. This is also shown in Fig. 7.7 and Fig. 7.9 in applying to our reactor model.

Thus an adaptive filter can be used to minimize the sensitivity matrix or to desensitize the system to any error in the filter

process. Mathematically, this is also very clear as shown in the previous section. The errors in filtering always appear in the forcing term of the inhomogeneous linear difference equation in calculating the actual error covariance matrix. Thus, the adaptive filter is trying to minimize the errors in filtering or to minimize the forcing term, thus minimizes the sensitivity matrix. A perfect adaptive filter will lead to a zero sensitivity matrix.

VII.7 Numerical Examples and Discussions

Sensitivity analysis for different error categories were performed on our reactor model. Generally speaking, we can reach the following conclusions:

1. Errors due to different initial error covariance will disappear after the filter reaches the steady state.

2. Errors due to measurement noise on temperatures only effect the temperature, not the concentration (due to the damping of reactor wall and fluids).

3. Errors due to model errors or integration errors can be desensitized by a proper error compensation technique.

4. Errors due to the combination of model error and prior statistics is less serious than that of having model error alone, since the error in statistics may desensitize the system and compensate for part of the model error.

5. Small errors due to measurement matrix in temperature do not effect the concentration significantly. This is also related to conclusion 2. The result is shown in Fig. 7.16 and Fig. 7.17. However,

CHAPTER VIII

ERRORS IN FILTERING

VIII.1 Introduction.

In this chapter we discuss various types of round off errors in modeling and filtering. The round off errors in filtering and modeling can cause filtering instability and even divergence. The accumulation and propagation of errors during the filtering process sometimes is so serious it can quickly produce catastrophic errors and a runaway situation.

Round off errors in the filtering part are due mostly to the filter gain calculation which is in turn related to the error covariance matrix calculation and matrix inverse calculations. Ill-conditioning of the process model transition matrix and too short a computer word length can easily introduce round off errors in the matrix inverse and matrix operations carried out in the gain calculation.

Our reactor model is a typical case with very sensitive dynamics. Propagation of round off errors in the early filtering stages often causes runaway of the reactor dynamics. In order to reduce the possible round off errors, the following areas are important:

A. Use a better matrix inverse algorithm to treat the ill-conditioned matrix.

B. Normalize the process dynamics state variables and use proper scaling to reduce the sensitivity due to round off and therefore reduce the possible ill-conditioning of the process transition matrix. i.e. reduce the wide spread of the eigenvalues of the system and thus reduce the system stiffness.

C. Choose proper sensor measurement or take several sensor measurements to reduce the uncertainty before starting the filtering.

D. Choose a very small initial error covariance of the estimate, therefore decreasing the dynamics range or the ill-conditioning of the matrix. Otherwise we will violate the linearization assumption.

E. As a result of this study, it is much better to make an off-line calculation of the upper and lower bounds of the error covariance matrix of the estimate before the filtering is started. Then, we have an idea of the best we can hope for and the worst we can get in the future filtering. Hence model instability can be prevented and divergence avoided by tuning the various model parameters and prior statistics and increasing the confidence in on-line filtering and thereby reducing the on-line computational load.

VIII.2 Modeling Errors and Approximations

As stated in the previous chapters and the last section, the modeling error consists of uncertainty in modeling parameters (as rate constants in the reactor model) or round off in integration, stiffness or instability of the model dynamics, and some deliberate simplification of the model dynamics for practical implementation. All of the above errors can cause filter divergence and instability.

A. Uncertainty in model parameters.

For a less sensitive model the uncertainty in a parameter can degrade the filter performance and give filter offset in the estimate. Some models are very sensitive when operating in certain region while less sensitive in another region. The reactor model is a typical case with reduced rate constant RK = 1.0, the model is not sensitive enough to cause modeling instability while raising RK to 1.5 causes and exponential growth in the rate term and runaway in both the dynamics and the filtering estimate. It is better to operate a filter in a less sensitive parameter region to insure better filter performance, however, sometimes it can be unrealistic to do so due to physical and economical reasons.

Here, the model error compensation and proper control will play a very important role in confining the filtering in a stable operable region before a runaway will occur. By using bang-bang control the dynamics and filtering estimates are stable even with RK=6.0 as it is shown in Fig. 3.3, and Fig. 3.4.

B. Round off in integration.

Round off error and truncation error in the integration of the model dynamics often causes oscillation and overshooting in the dynamics. This behavior is unacceptable when the filter is operating on a sensitive model. An intensive study has been made in this work of the effects round off and truncation error on the filtering performance. An adaptive type model error compensation technique was developed to reduce the sensitivity and overshooting.

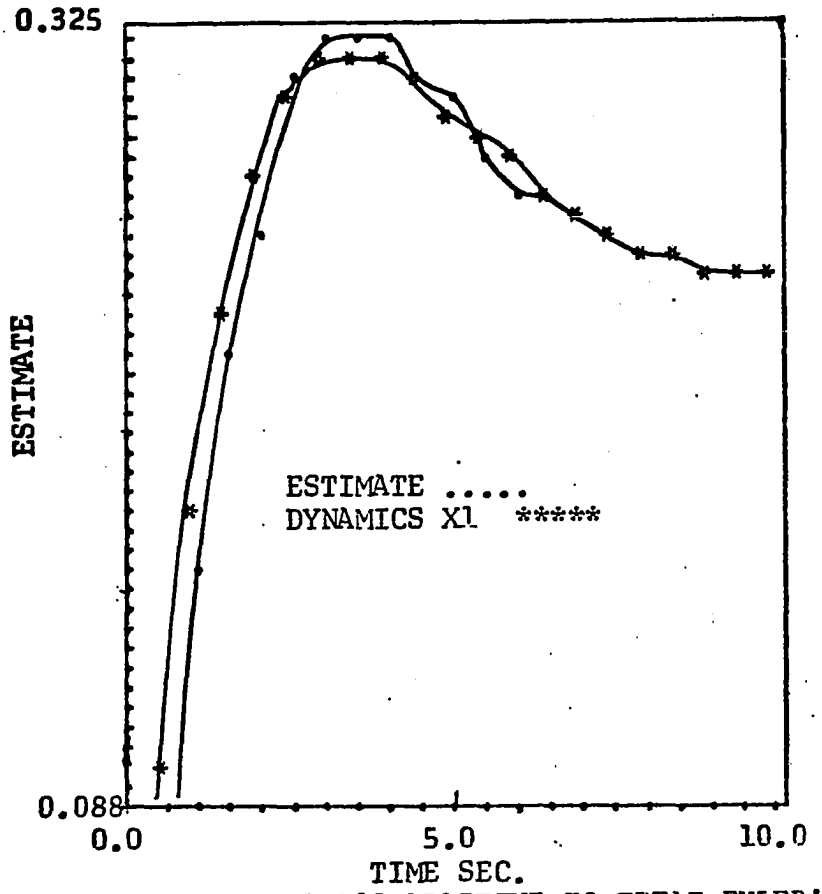


FIG.8.1 HUANG'S ADAPTIVE TO TREAT EULER'S INTEGRATION ERROR AND SIMULATE THE RESULT OF RUNGE-KUTTA MERSON SECOND ORDER NONLINEAR FILTER

TIME SEC.	X 1	ESTIMATE
0.50	0.0988	0.0843
1.00	0.1810	0.1617
1.50	0.2427	0.2292
2.00	0.2845	0.2685
2.50	0.3097	0.3205
3.00	0.3222	0.3315
3.50	0.3255	0.3298
4.00	0.3228	0.3288
4.50	0.3165	0.3205
5.00	0.3083	0.3100
5.50	0.2995	0.2920
6.00	0.2907	0.2829
6.50	0.2827	0.2796
7.00	0.2755	0.2732
7.50	0.2695	0.2671
8.00	0.2645	0.2646
8.50	0.2607	0.2626
9.00	0.2578	0.2585
9.50	0.2559	0.2533
10.00	0.2547	0.2585

Euler's integration method was used in generating the filtering estimate. Using the extended Kalman filter with Euler integration caused oscillation and overshooting (also shown Wells [W1]). We have been very successful in introducing a fictitious model noise proporational to the amount of the truncation error of the integration routine. Significant improvement has been shown in using an extended Kalman filter combined with model error compensation. Also, it has been found that a combination of the second order non-linear approximate filter with model error compensation shows even better performance.

The truncation error for the integration routine used is closely related to the integration interval h. For Euler integration, this error is proportional to h^2, while for the Runge-Kutta Merson it is h^5. The numerical stability of the integration is also related to the particular interval chosen for integration. In the reactor model, numerical instability is observed when h = 0.1 sec for the Euler integration. However, h = 0.5 for Runge-Kutta Merson routine led to instability. The smaller the step interval, a heavier computational load is necessary, while too large a step size not only causes serious round off and truncation errors but also causes ill-conditioning of the process transition matrix, introducing errors in the matrix inverse, and then instability in the process dynamics. Recent studies on the efficient integration routines for stiff differential equation system are reported by Lapidus and Willoughby [S10], [W5].

C. Model approximation.

In many filtering applications, one should be willing to trade off some filter accuracy by having model simplification. This can be

done by the following approaches:

1. Use a pseudo-linear analysis by dropping the non-linear term, or approximate the non-linear terms by linearization.

2. Reduction of the number of state variables of the system, thus approximating a higher complex system by a lower dimensional system.

3. Decompose the higher dimension system into several small subsystems, and then perform the filtering for each of the subsystem.

Bucy [B8] made several theoretical contributions to the above subjects. We apply these successfully to our reactor model.

VIII.3 Round Off Errors in Matrix Inversion

For the time varying non-linear filtering case, matrix inversion is required whenever the process transition matrix changes. Round off error can become very serious when the transition matrix is ill-conditioned. In the reactor model, for instance, the ill-conditioning of the transition matrix is worse than that of a Hilbert matrix, if we use h = 0.2 or higher for the Euler integration. In order to reduce possible round off in the matrix inverse, we choose a combination of the Shur relations and matrix interaction to carry out the matrix inversion in the filter gain calculation.

A. Matrix inversion by the Shur relation.

The Shur relation expands the determinant by determinant of subblocks (Gantmacher [G1]). Suppose we are interested in finding the inverse of matrix D, an NxN matrix. Let D_m be the mxm matrix formed by the first m rows and columns.

Let

$$D_m = \begin{pmatrix} A_m & b_m \\ c_m' & d_m \end{pmatrix} = \begin{pmatrix} D_{m-1} & b_m \\ c_m' & d_m \end{pmatrix} .$$

Then a recursive scheme is developed as:

$$D_m^{-1} = E_m = \begin{pmatrix} E_{m-1} + \alpha_m E_{m-1} b_m c_m' E_{m-1} & -\alpha_m E_{m-1} b_m \\ -\alpha_m c_m' E_{m-1} & \alpha_m \end{pmatrix} . \qquad \text{VIII-1}$$

Where

$$\alpha_m^{-1} = d_m - c_m' E_{m-1} b_m$$

$$E_1 = 1/D_{11} \; ; \quad d_2 = D_{22}; \quad b_2 = D_{12}; \quad c_2 = D_{21} .$$

B. Matrix iteration.

Suppose we already have the matrix inverse, B of matrix A, then an improved matrix inverse can be found by the following iteration scheme:

$$B_{i+1} = B_i(2I - AB_i) \quad . \qquad \text{VIII-2}$$

VIII.4 Numerical Examples and Discussions

Extensive evaluations have been made in this work on various errors in the filtering studies of our reactor model. Generally speaking, round off errors due to matrix gain calculation or integration truncation is not serious if we use filter sample period of 0.5 sec and integration interval of 0.05 sec. Round off errors of these types can easily be compensated by the adaptive techniques introduced in this work as shown in Fig. 8.1, and Fig. 8.2. Sometimes it is desirable to use various model simplifications as listed in the previous

section. The errors introduced due to model simplification often bring permanent offset, due to the nongaussian properties of the errors.

A. Pseudo linear analysis.

The error introduced due to pseudo-linear analysis has been treated successfully by Bucy [B8] in terms of the development of upper bounds for the error in process transition matrix:

The norm of a vector is defined as:

$$||x|| = (\sum_i x_i^2)^{\frac{1}{2}} \quad ,$$

while the norm of a matrix is

$$||G|| = (\sum_{ij}\sum G_{ij}^2)^{\frac{1}{2}} \quad .$$

Let

$$a = ||G^{-1}\delta G|| ; \qquad \alpha = ||\delta G G^{-1}|| .$$

Then, for $a \ll 1$ we have

$$\frac{|\delta\lambda_i|}{\lambda_i} \leq 2\alpha \quad , \qquad \text{VIII-3}$$

where

$$\lambda_i = \text{eigenvalue of } P_* \quad .$$

The error in the calculation of error covariance is related to that of the process transition matrix. Let

$$\Theta = (\frac{1}{a})\delta G \quad .$$

Then

$$\begin{aligned} P_* &= (G + a\Theta)\bar{P}(G + a\Theta)' \\ &= P + a[G\bar{P}\Theta' + \Theta\bar{P}G' + a\Theta\bar{P}\Theta'] \quad . \end{aligned} \qquad \text{VIII-4}$$

By applying Equation VIII-4 n times, we have

$$\frac{|\delta\lambda_i|}{\lambda_i} \leq 2n\alpha \quad .$$

This implies that the criterion of accuracy of the process transition matrix G should be

$$||\delta G G^{-1}|| << \frac{1}{2n}$$

where n is the number of computing intervals, $n\Delta t$ is equal to the length of time that the approximation is used. Thus, this presents an upper bound for the linearization error in the extended Kalman filter.

B. Reduction of number of dimension of the state variables.

As a numerical example, we approximate our 4th order reactor model by a second order model. Thus we use a second order model filter to estimate the 4th order reactor model. The actual model:

$$\dot{x}_1 = -(c_1+c_4)x_1+c_3(1+x_4)^2\exp\left(\frac{K_1x_1}{1+x_1}\right)+ C_4x_2-c_3 \qquad \text{VIII-5}$$

$$\dot{x}_2 = -(c_5+c_6)x_2+c_5x_1+c_6x_3 \qquad \text{VIII-6}$$

$$\dot{x}_3 = -(c_7+c_8)x_3+c_8x_2 \qquad \text{VIII-7}$$

$$\dot{x}_4 = - c_1x_4 - c_2(1+x_4)^2\exp\left(\frac{K_1x_1}{1+x_1}\right)+ c_2 \quad . \qquad \text{VIII-8}$$

The filter model: dropping x_2 and x_3 and Equations VIII-6 and VIII-7

$$\dot{x}_1 = -(c_1+c_4)x_1 + c_3(1+x_4)^2\exp\left(\frac{K_1x_1}{1+x_1}\right) - c_3 \qquad \text{VIII-8}$$

$$\dot{x}_4 = - c_1x_4 - c_2(1+x_4)^2\exp\left(\frac{K_1x_1}{1+x_1}\right)+ c_2 \qquad \text{VIII-9}$$

Satisfactory results are shown in Fig. 8.5 and Fig. 8.6 by using the smoothing and adaptive techniques of the work. Slight offset after 5 seconds is due to the steady increasing x_2 which we ignored in our simplified model (this violates the Gaussian assumption of the dynamic noise.

C. Decomposition of large system into subsystem.

Bucy [B8] presented the idea of partitioning a large complex system into several smaller systems and perform the filtering for each of the subsystems. This approach not only saves computation storage and reduces the computation load but also reduces the possible round off errors in the calculations of filter gain matrix. With the help of a decomposition algorithm, a very large unmanageable system can be solved by partitioning it into several easily manageable subsystems. This can be presented as follows:

Suppose we intend to decompose the system state vector $\underline{x}$ into m subsystems $\underline{\tilde{x}}_i$; then $\underline{\tilde{x}}_i = D_i\underline{x}$. In order to be able to recover the $\underline{x}$ vector from the $\underline{\tilde{x}}_i$, D_i have to satisfy

$$\sum_{i=1}^{k} D_i^{+}D = I \quad , \qquad \text{VIII-10}$$

where D_i is the pseudo inverse of D_i. Therefore,

$$\sum_{i=1}^{k} D_i^{+}\underline{\tilde{x}}_i = \underline{x} \quad . \qquad \text{VIII-11}$$

Then the system equations can be written as:

$$\underline{\tilde{x}}_1(k+1) = D_i \, \Phi \sum_{j=1}^{m} D_j^{+}\underline{\tilde{x}}_j(k) + D_i\underline{u} \quad . \qquad \text{VIII-12}$$

The measurement will be decomposed as:

$$\underline{z}_i = E_i\underline{z} = E_iH\underline{x} + E_i\underline{v} = E_iH\sum_{j=1}^{m} D_j^+\underline{\bar{x}}_j + E_i\underline{v} \quad . \qquad \text{VIII-13}$$

The suboptimal decomposed filter will be:

Prediction -

$$\underline{\hat{\bar{x}}}_i(k+1,k) = D_i\Phi(k)\sum_{j=1}^{m} D_j^+\underline{\bar{x}}_j(k,k) \qquad \text{VIII-14}$$

$$\underline{\hat{\bar{x}}}_i(k+1,k+1) = \underline{\hat{\bar{x}}}_i(k+1,k) + K_i[z_i(k+1)-E_iH\sum_{j=1}^{m} D_j^+\underline{\hat{\bar{x}}}_j(k+1,k)]. \qquad \text{VIII-15}$$

Let

$$N_i = E_iHD_i^+ \; ; \qquad \theta_i(k) = D_i\Phi(k)D_i^+ \quad .$$

Then

$$K_i(k) = P_i(k)N_i^![N_iP_i(k)N_i^!+R_i]^{-1} \qquad \text{VIII-16}$$

$$P_i(k+1) = \theta_i(k)\bar{P}_i(k)\theta_i^!(k) + Q_i \qquad \text{VIII-17}$$

$$C_i(k+1) = I-K_i(k+1)N_i(k+1) \qquad \text{VIII-18}$$

$$P_i(k+1) = C_i(k+1)P_i(k+1) \qquad \text{VIII-19}$$

where $P_i(0)$, Q_i and R_i are the initial error covariance, process noise covariance and measurement noise covariance matrix for the subsystems. The state variables which produce small effects can be separated into $\bar{x}_2$.

In applying this idea to our reactor model, we separate our four-state variablesproblem into two two-variable subsystems, i.e. x_1 and x_4. This is more sensitive and is separated from x_2 and x_3. Therefore we have

$$\underline{\bar{x}}_1 = \begin{pmatrix} x_1 \\ x_4 \end{pmatrix} \qquad \underline{\bar{x}}_2 = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} \quad .$$

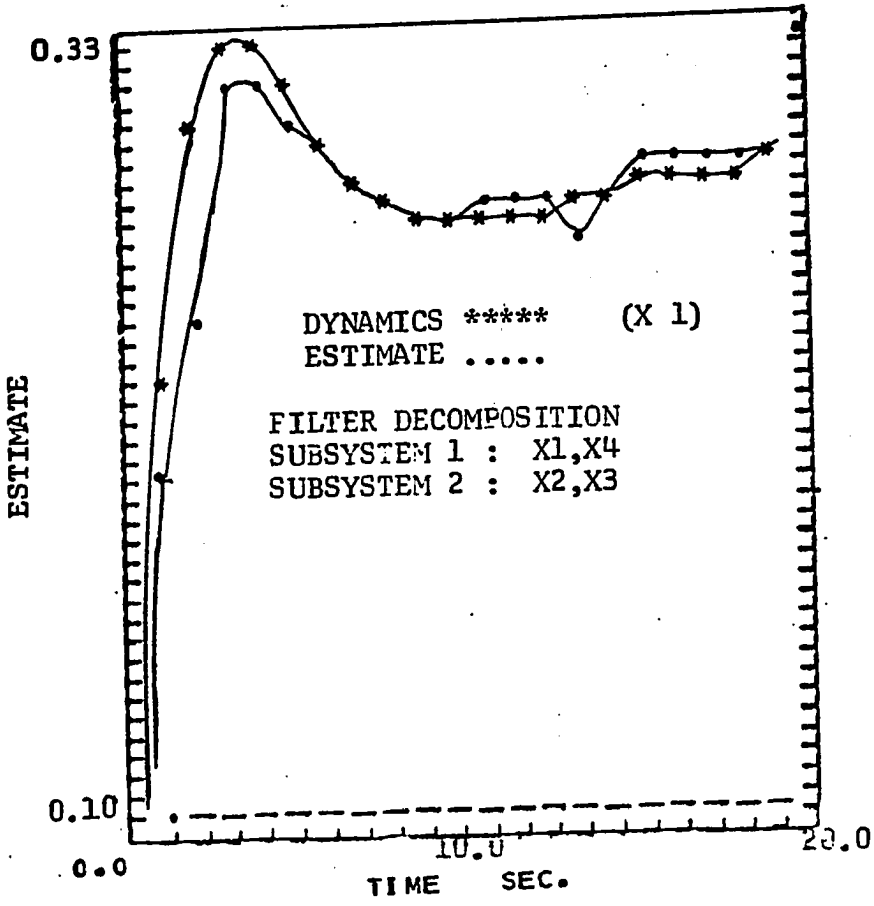


FIG. 8.2 DOUBLE THE SAMPLING PERIOD

FILTER DECOMPOSITION ,SUBSYSTEM 1

TIME SEC.	X 1	ESTIMATE
1.0	0.1856	0.1509
2.0	0.2927	0.2828
3.0	0.3292	0.3148
4.0	0.3265	0.3126
5.0	0.3087	0.2936
6.0	0.2888	0.2877
7.0	0.2723	0.2742
8.0	0.2610	0.2584
9.0	0.2545	0.2493
10.0	0.2520	0.2543
11.0	0.2521	0.2597
12.0	0.2540	0.2584
13.0	0.2568	0.2614
14.0	0.2599	0.2475
15.0	0.2631	0.2624
16.0	0.2661	0.2744
17.0	0.2690	0.2783
18.0	0.2716	0.2785
19.0	0.2741	0.2767
20.0	0.2765	0.2807

ACCUMULATED MEAN SQUARE ERROR = 0.2126E-03

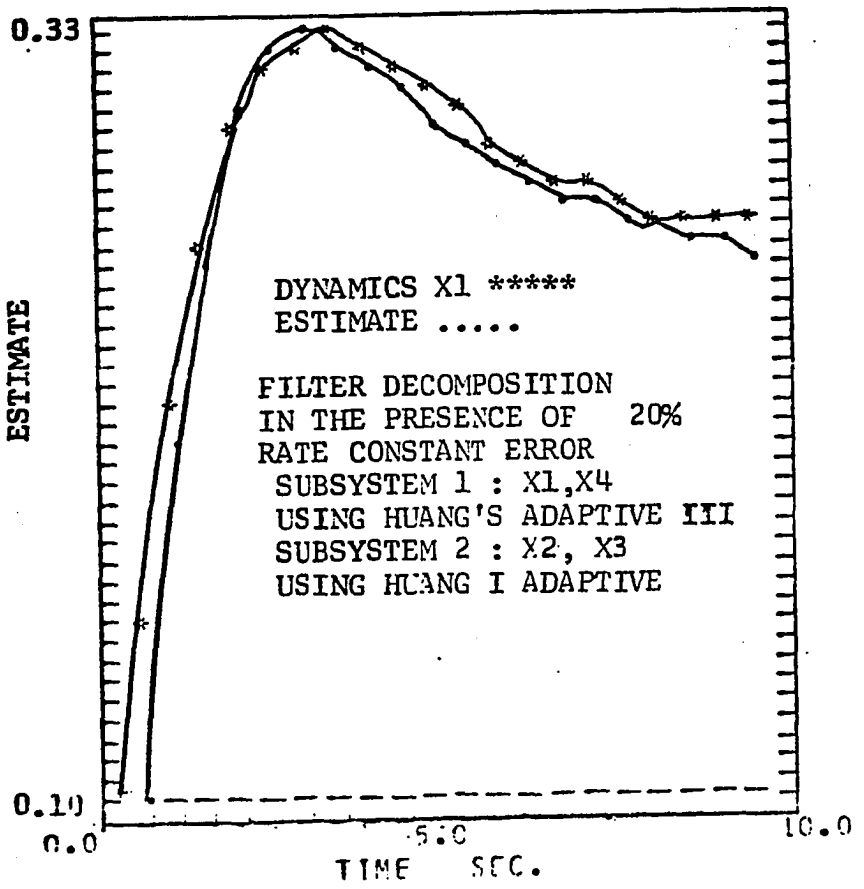


FIG. 8. 3. FILTER DECOMPOSITION IN THE PRESENCE OF 20 % RATE CONSTANT ERROR

TIME SEC.	X 1	ESTIMATE
0.5	0.1002	0.0331
1.0	0.1842	0.1700
1.5	0.2475	0.2399
2.0	0.2902	0.2980
2.5	0.3153	0.3269
3.0	0.3270	0.3319
3.5	0.3293	0.3273
4.0	0.3254	0.3174
4.5	0.3179	0.3079
5.0	0.3087	0.2951
5.5	0.2989	0.2846
6.0	0.2895	0.2748
6.5	0.2809	0.2709
7.0	0.2734	0.2648
7.5	0.2671	0.2598
8.0	0.2621	0.2543
8.5	0.2583	0.2537
9.0	0.2556	0.2490
9.5	0.2538	0.2454
10.0	0.2523	0.2425

ACCUMULATED MEAN SQUARE ERROR = 0.3139E-03

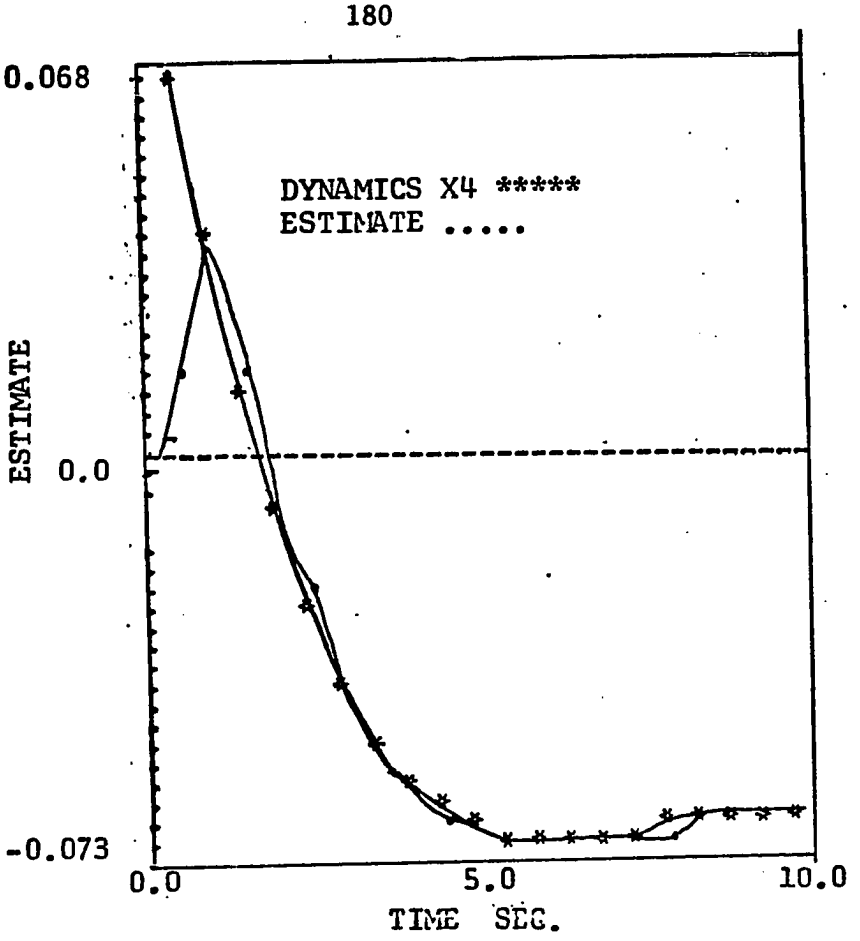


FIG. 8.4 USING HUANG'S ADAPTIVE I TO COMPENSATE FOR EULER"S INTEGRATION ERROR TO SIMULATE THE RESULT OF USING RUNGE-KUTTA MERSON INTEGRATION

TIME SEC.	X 4	ESTIMATE
0.50	0.0683	0.0153
1.00	0.0377	0.0379
1.50	0.0107	0.0124
2.00	-0.0120	-0.0105
2.50	-0.0301	-0.0257
3.00	-0.0440	-0.0425
3.50	-0.0543	-0.0546
4.00	-0.0616	-0.0626
4.50	-0.0666	-0.0677
5.00	-0.0698	-0.0708
5.50	-0.0716	-0.0726
6.00	-0.0725	-0.0727
6.50	-0.0727	-0.0729
7.00	-0.0724	-0.0726
7.50	-0.0718	-0.0720
8.00	-0.0711	-0.0712
8.50	-0.0703	-0.0705
9.00	-0.0696	-0.0696
9.50	-0.0689	-0.0686
10.00	-0.0684	-0.0688

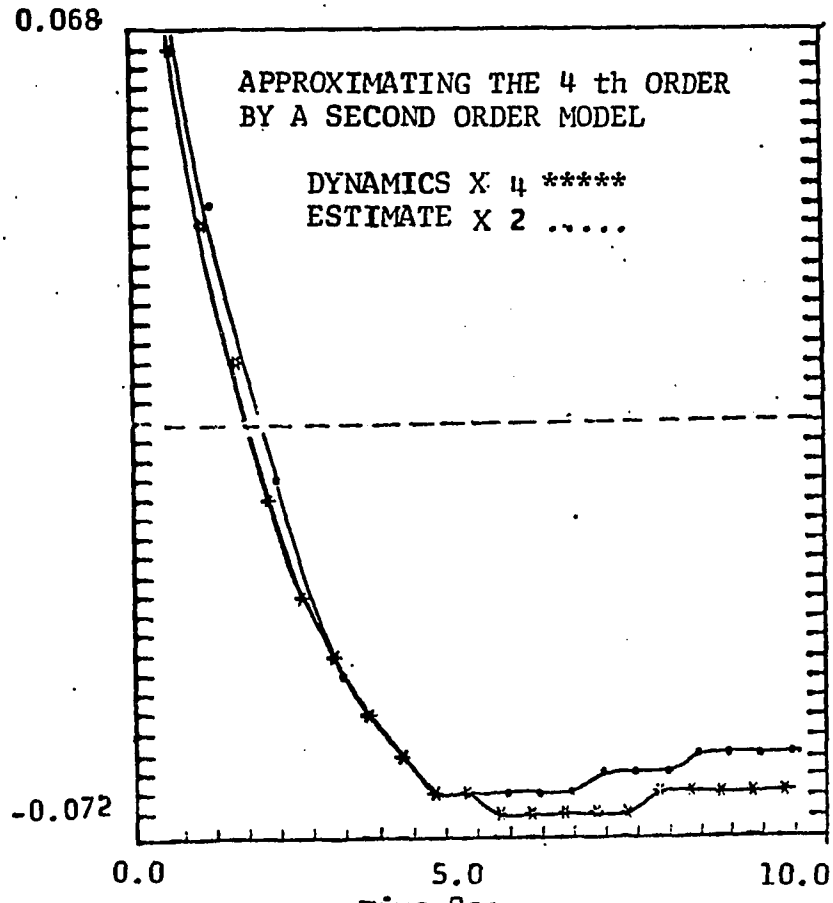


FIG.8.5 MODEL APPROXIMATION
SECOND ORDER NONLINEAR FILTER
HUANG'S III

TIME SEC.	X 2	ESTIMATE
0.5	0.0683	0.0668
1.0	0.0373	0.0388
1.5	0.0097	0.0114
2.0	-0.0135	-0.0125
2.5	-0.0319	-0.0317
3.0	-0.0459	-0.0461
3.5	-0.0562	-0.0564
4.0	-0.0634	-0.0633
4.5	-0.0682	-0.0682
5.0	-0.0712	-0.0696
5.5	-0.0727	-0.0704
6.0	-0.0733	-0.0700
6.5	-0.0732	-0.0693
7.0	-0.0727	-0.0678
7.5	-0.0719	-0.0667
8.0	-0.0710	-0.0656
8.5	-0.0702	-0.0640
9.0	-0.0693	-0.0626
9.5	-0.0686	-0.0616
10.0	-0.0680	-0.0612

ACCUMULATED MEAN SQUARE ERROR = 0.1519E-04

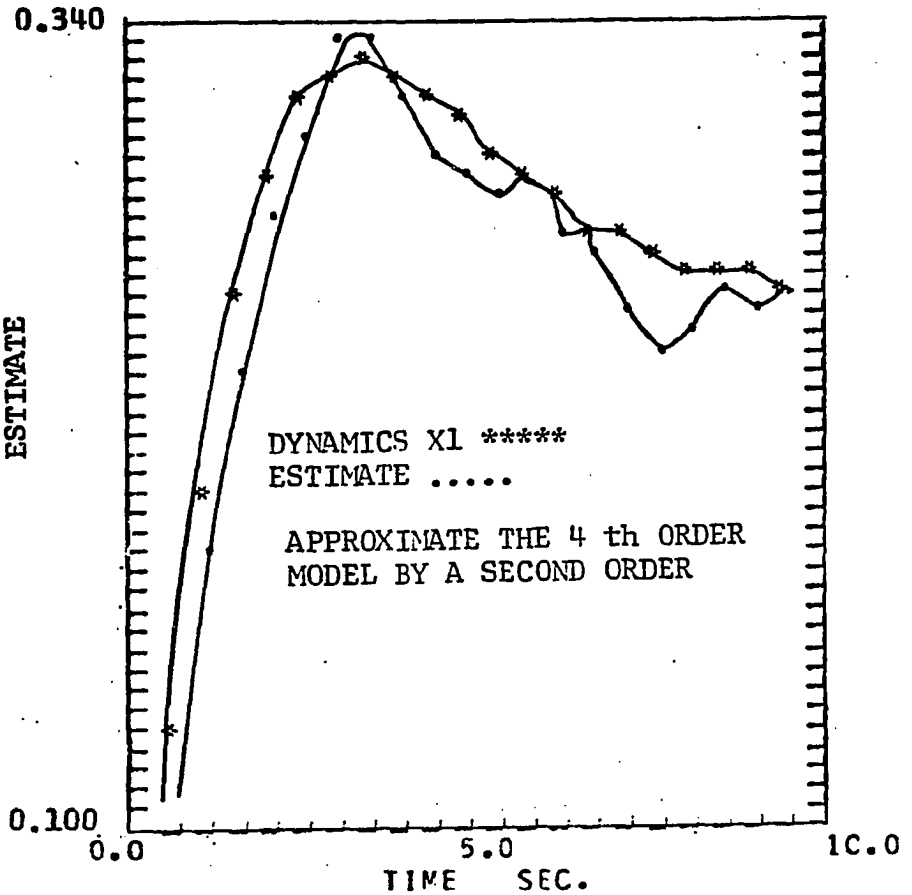


FIG. 8.6 MODEL DIMENSION REDUCTION
SECCND CRDER NCALINEAR FILTER
HUANGS ADAPTIVE III

TIME SEC.	X 1	ESTIMATE
0.5	C.1002	0.C764
1.0	0.1842	0.1656
1.5	0.2475	0.2238
2.0	C.2902	0.2773
2.5	0.3153	C.3C54
3.0	0.3270	0.3370
3.5	C.3293	0.3372
4.0	0.3254	0.3180
4.5	0.3179	0.2988
5.0	C.3087	C.2879
5.5	0.2989	0.2851
6.0	0.2895	0.2895
6.5	C.2809	0.2686
7.0	0.2734	C.2604
7.5	0.2671	0.2435
8.0	0.2621	C.2334
8.5	0.2583	0.2342
9.0	0.2556	0.2524
9.5	C.2538	0.2425
10.0	0.2528	0.2489

ACCUMULATED MEAN SQUARE ERRCR = 0.2685E-03

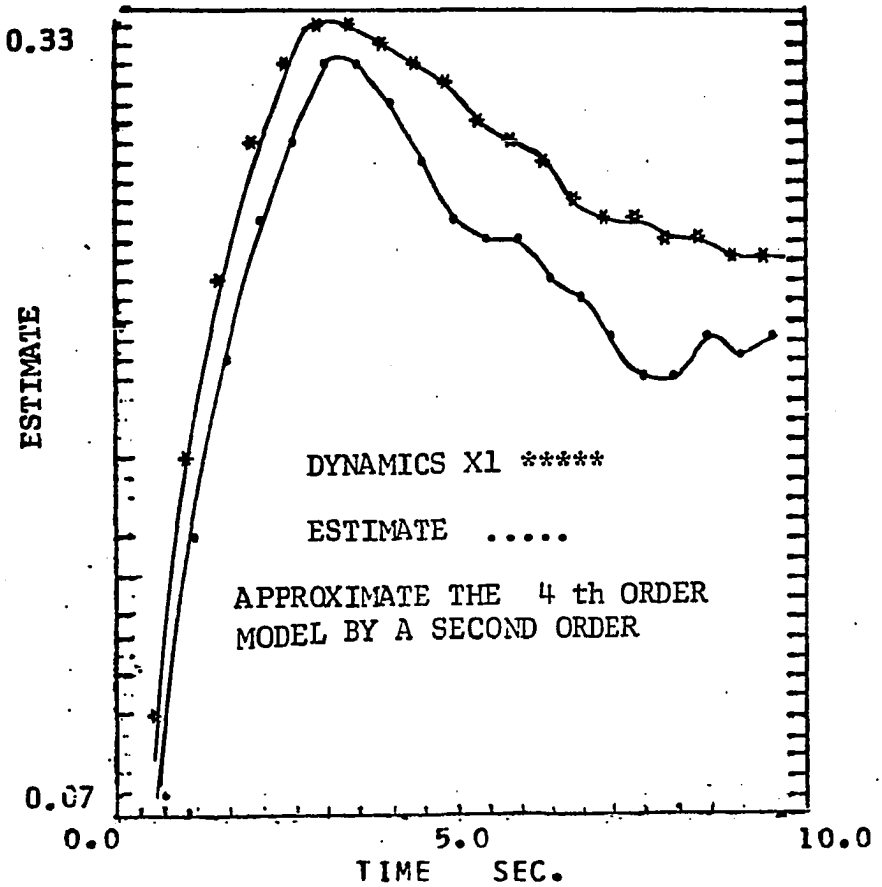


FIG. 8.7 CONSTANT NOISE
SECOND ORDER NONLINEAR FILTER
CONSTANT NOISE

TIME SEC.	X 1	ESTIMATE
0.5	0.1002	0.0764
1.0	0.1842	0.1620
1.5	0.2475	0.2178
2.0	0.2902	0.2659
2.5	0.3153	0.2911
3.0	0.3270	0.3149
3.5	0.3293	0.3158
4.0	0.3254	0.3003
4.5	0.3179	0.2813
5.0	0.3087	0.2670
5.5	0.2989	0.2600
6.0	0.2895	0.2605
6.5	0.2809	0.2456
7.0	0.2734	0.2379
7.5	0.2671	0.2240
8.0	0.2621	0.2140
8.5	0.2583	0.2121
9.0	0.2556	0.2247
9.5	0.2538	0.2201
10.0	0.2528	0.2257

ACCUMULATED MEAN SQUARE ERROR = 0.1057E-02

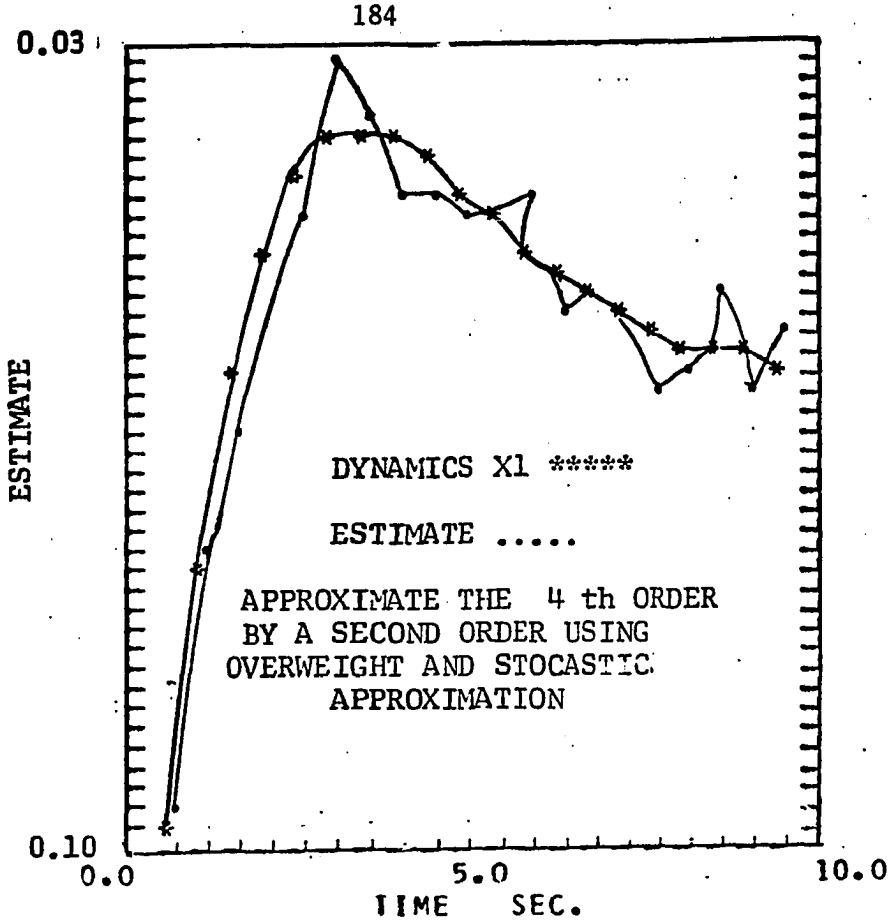


FIG. 8.8 MODEL APPROXIMATION BY HUANGS IV SECOND ORDER NONLINEAR FILTER

TIME SEC.	X 1	ESTIMATE
0.5	0.1002	0.1069
1.0	0.1842	0.1887
1.5	0.2475	0.2269
2.0	0.2902	0.2881
2.5	0.3153	0.3031
3.0	0.3270	0.3536
3.5	0.3293	0.3359
4.0	0.3254	0.3088
4.5	0.3179	0.3088
5.0	0.3087	0.3028
5.5	0.2989	0.3021
6.0	0.2895	0.3079
6.5	0.2809	0.2666
7.0	0.2734	0.2746
7.5	0.2671	0.2422
8.0	0.2621	0.2459
8.5	0.2583	0.2507
9.0	0.2556	0.2745
9.5	0.2538	0.2418
10.0	0.2528	0.2653

ACCUMULATED MEAN SQUARE ERROR = 0.1971E-03

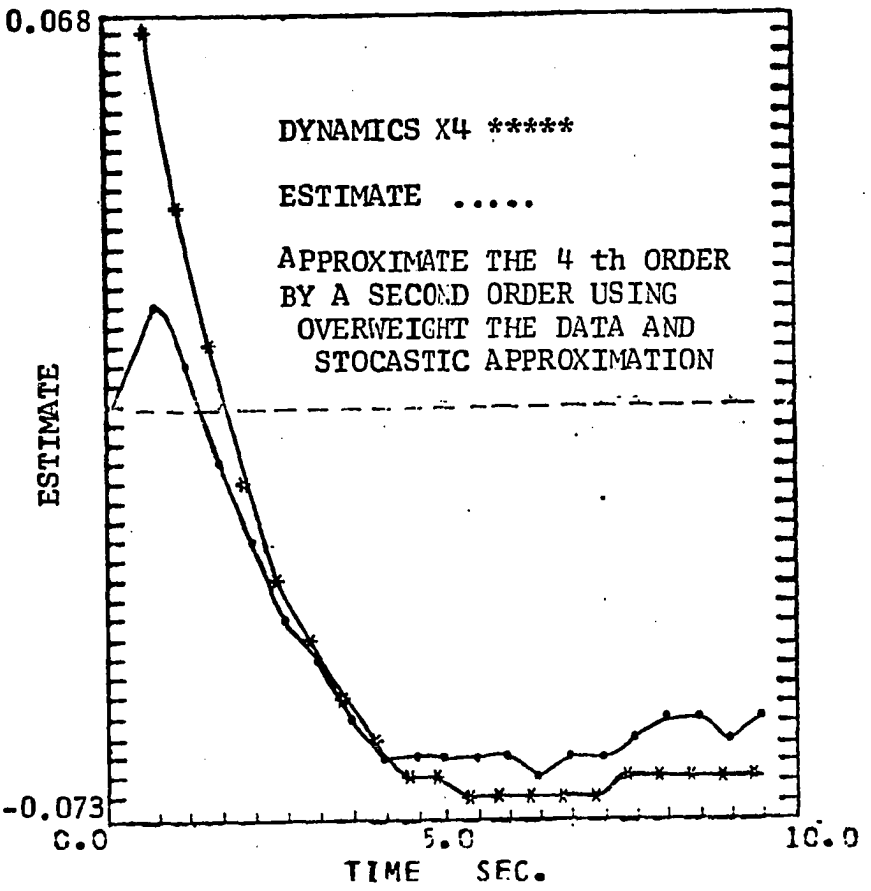


FIG. 8.9 MODEL APPROXIMATION BY HUANG'S IV SECOND ORDER NONLINEAR FILTER

TIME SEC.	X 2	ESTIMATE
0.5	0.0683	0.0173
1.0	0.0373	0.0074
1.5	0.0097	-0.0115
2.0	-0.0135	-0.0247
2.5	-0.0319	-0.0400
3.0	-0.0459	-0.0479
3.5	-0.0562	-0.0593
4.0	-0.0634	-0.0652
4.5	-0.0682	-0.0657
5.0	-0.0712	-0.0666
5.5	-0.0727	-0.0667
6.0	-0.0733	-0.0662
6.5	-0.0732	-0.0690
7.0	-0.0727	-0.0651
7.5	-0.0719	-0.0652
8.0	-0.0710	-0.0616
8.5	-0.0702	-0.0594
9.0	-0.0693	-0.0578
9.5	-0.0686	-0.0612
10.0	-0.0680	-0.0581

ACCUMULATED MEAN SQUARE ERROR = 0.2441E-03

The results by decomposing the reactor system into two subsystems show better performance than those of reduction of dimension, i.e. ignore the presence of x_2 and x_3 at all as this is intuitively clear.

Filter decomposition of large complex system not only saves computation time and storage but also offers better filtering result than those of filtering the original system. This is true for the large stiff (or ill-condition) system, since the ill-conditioning is reduced by partitioning the original system into smaller subsystem thus a better filter gain is calculated for the subsystem as this is shown in Fig. 8.2 where the integration interval is doubled the result of decompose into two subsystem is better than those of filtering the original system as shown in Fig. 2.14.

The result shown on Fig. 8.3 where the decomposed filter also offers better alternatives than the original system in the presence of 20 per cent rate constant error. Applying different adaptive techniques to different subsystem based on its structure and sensitivity nature seems to be a better approach to the filtering in the presence of large model error. As this is shown in Fig. 8.3, we grouped x1, x4 into subsystem 1 as it is more sensitive to model errors and apply HUANG'S III adaptive to this subsystem and group x2 and x3 into subsystem 2, which is less sensitive; therefore we apply Huang's I adaptive.

CHAPTER IX

REAL TIME IMPLEMENTATION OF FILTERING

The growing importance of applying direct digital control (DDC) techniques to an industrial process is due to its great success in aerospace applications, As a result of the expanding aerospace industry in the last ten years, improvement in both the computer hardware and software coupled with advanced numerical mathematical techniques, modern control theory reflects the ability to use digital computer to solve complex problems which were considered impossible a few years ago. Recent advancement in the field of stochastic processes leads as a step closer to the real world.

The advantage of the Kalman filter over conventional regression and analysis and curve filtering are:

A. The Kalman filter uses a sequential recursive algorithm, which requires much less computer storage and computation time.

The following example given by Bucy [B8] is very helpful. Considering just the matrix operation and inversion, curve fitting requires $kn_x n_y(1+n_x) + n_x^2$ multiplication plus an $n_x x n_x$ inversion in processing one data point while the Kalman filter only requires $n_x n_y(3n_x+2n_y+2) + n_x^2 + 4n_x^3$ multiplications plus an $n_y x n_y$ inversion.

Where

k = sample points

n_x = number of state variables

n_y = number of measurement variables .

As an example, consider:

$k = 100$

$n_x = 10$

$n_y = 3$.

Then, for curve filtering, 33,100 multiplications and a 10x10 inversion are needed while only 5240 multiplications and a 3x3 inversion are required for the Kalman filter. Curve fitting requires the storage of all the observation data while the sequential kalman filter can process observations online, one at a time.

B. Kalman filtering can handle the estimation of dynamics under random forcing while curve fitting cannot. With the help of on-line digital control combined with a filter algorithm, DDC becomes possible. Real time filtering and the computer control system will detect any outside disturbance, then adjust itself to the new optimum state. Even the uncertainty in physical parameters in the design state will not effect the real plant performance when used with the help of an on-line error compensation technique.

We are interested in developing a reliable non-linear filter to fit the real time DDC environment and yet powerful enough to deal with most of the process disturbance and parameters uncertainties.

In the computational implementation of real filtering, we are faced with the following problems:

1. Limited memory.
2. imited speed.
3. Limited word length.

1. Limited memory: This is the physical limitation of the existing computer. A typical 6-state variable filter as is used in our reactor example requires about 600 words on the IBM-360. Generally speaking, the memory requirement varies roughly as the square of the number of state variable. Therefore, it is desirable to keep the state variables of the system as low as possible. An alternative to reducing the number of state variables is to partition (or decompose) the state variables x into two or more parts and construct separate filters for each part. This approach is called a 'suboptimal filter' by Bucy [B8]. He derived the suboptimal filter algorithm and applied it successfully to the Ranger IV, where it was possible to reduce a nine state variable problem to three separate three state variable problems. It takes less computing capacity to manipulate three 3x3 matrices than one 9x9 matrix. This also drastically reduces the computer storage since the storage requirement is proportional to the square of the number of state variables.

2. Limited Speed: The improvement of computer hardware and software make computational speed no longer a problem in on-line computation of a reasonable complex model. This is also true in real time filtering. Computing time varies roughly as the cube of the number

of state variables of the system. Our 4-variable reactor system takes about 1.0 second for processing one sample period.

3. Limited word length: The limited word length of the digital computer introduces round off errors into the solution of filter equations. It is also desirable to keep the number of state variables as low as possible to avoid the excessive accumulation of round off errors. However, selection of proper state variables and scaling (or normalization) can easily avoid the trouble due to the limited word length. This is especially true in our reactor model, we normalized the temperatures (usually vary from 100 to 1000) and the concentration (vary from 0.01 to 0.09). Without the normalization, this technique suffers from round off errors and leads to bad results. The introduction of the Shur relation and matrix iteration for the matrix inverse together with the Pine's machine error compensation also help reduce the round off errors to an acceptable degree.

We conclude this chapter by a summary of results (Table VIII-1) of various filters developed in this work and its feasibility as to the future real time implementation.

Table VIII-1. Summary of adaptive filters.

Filters	Memory Requirement (in standard word)	Word Length	Speed for 1 Period (in sec.)
Extended Kalman filter	600	no problem	1.0
Non-linear second order	620	no problem	1.2
Schmidt	620	no problem	1.3
Wolf's I	600	no problem	1.0
Wolf's II	600	no problem	1.0
Overweight the recent data	600	no problem	1.0
Pine's	600	no problem	1.0
This work I	600	no problem	1.0
This work II	620	no problem	1.2
This work III	600	no problem	1.0
This work IV	600	no problem	1.1
Limited memory	650	no problem	1.3
Dimension reduction*	500	no problem	0.7
Filter decomposition*	550	no problem	0.8

*combined with the adaptive filters introduced by this work.

(Approximately 0.3 second is used for generating the simulated observation.) It is obvious that the dimension reduction and filter decomposition are the most attractive filters in real time applications.

CHAPTER X

SUMMARY AND CONCLUSIONS

After this extensive study of filtering theory and applications, we feel more confident in the implementation of a filter analysis either on-line or off-line. It was shown that the sensitive nature of the non-linear filtering problem is due to a much smaller region for filter stability (observability and controlability) than that in the linear filtering. With the aide of error sensitivity analysis and error bound calculation, and the error compensation technique developed in this work, one is able to do the prefiltering study by using a simple lower dimensional model to approximate the complex model, design a better measurement sensor system through the sensitivity analysis and study the interaction among the variables. Thus, a reliable off-line filter model can be implemented for real time application where limited computer memory and speed are of great concern.

The introduction and analysis of a second order approximate filter by this work also provides an alternative for non-linear real time filtering where sampling is expensive, computer speed is limited and the model is highly nonlinear.

With the guide and formulation in Chapter III, possible future implementation of combined filtering and control with more sophisticated controller i.e. combined feed forwarded and feedback control can be done by dynamic programming as reported by West and McGuire [W3], Clifton and McGuire [C1], Jacobson [J1] or by applying conjugate gradient techniques and invariant embedding to solve the two point split boundary value problem via iteration in the control policy function space.

The introduction of several adaptive filtering and model error compensation techniques enables us to extend the Kalman filter to various non-linear cases with severe model error and non-Gaussian model and provides satisfactory filtering result.

The sophisticated filter decomposition algorithm developed by this work not only offer a successful approach to the filtering of large complex system with significantly saving in computation time and computer storage but also provides better filtering result for non-Gaussian non-linear systems.

BIBLIOGRAPHY

A1. Aoki, M. Optimization of Stochastic System, Academic Press, N. Y. 1967.

A2. Athans, M. A., Wishner, R. P., Bertolini, A. IEEE Trans. Automatic Control Vol. AC-13, Oct. 1968.

B1. Balakrishnan, A. V. Control Theory and the Calculus of Variations, Academic Press, 1969.

B2. Bass, R. W., Norum, V. D., Schwartz, L. J. Math. Anal. Appl., Vol. 16, 1966.

B3. Battin, R. H. Astronautical Guidance, McGraw-Hill, N. Y., 1964.

B4. Bellman, R. E., Kagiwada, H. H., Kalaba, R. E. J. Astronautical Science, 1966.

B5. Box, G., Jenkins, G. Time Series Analysis Forecasting and Control, Holden Day, San Francisco, 1970.

B6. Bryson, A. E. Jr., Henrikson, L. J. Harvard University, Div. Eng. Appl. Phys., Technical Report, 533, 1967.

B7. Bryson, A. E. Jr., Ho, Y. C. Applied Optimal Control, Blaisdall Publishing Co., Waltham, Mass., 1969.

B8. Bucy, R. A., Joseph, P. D. Filtering for Stochastic Process with Applications to Guidance, John Wiley & Sons, N. Y. 1968.

C1. Clifton, D. M. Ph.D. Dissertation, University of Oklahoma, 1970.

C2. Cramer, H. Mathematical Method of Statistics, Princeton Univ. Press, 1946.

D1. Denham, W. F., Pines, A. AIAA, J. Vol. 4, No. 6, 1966.

D2. Deyst, J. J., Price, C. F. IEEE Trans. Automatic Control, Jan. 1970.

F1. Fagin, S. L. JACC, 1968, University of Michigan, Ann Arbor, Michigan 1968.

F2. Fu, K. A. Pattern Recognition and Learning, Academic Press, N. Y. 1969.

G1. Gantmacher, F. R. The Theory of Matrices Vol. 1 and 11, Chelsea Publishing Co., N. Y. 1959.

G2. Gardner, L. A., Albert, F. A. Stochastic Approximation and Non-linear Regression, MIT Press, 1967.

G3. Gavalas, G. R., Seinfeld, J. H. Chemical Engr. Science, Jan. 1969.

G4. Griffin, R. E., Sage, A. P. IEEE Trans. Automatic Control, Vol. AC-13, No. 4, Aug. 1968 a, b.

H1. Haynes, G. W., Hermes, H. SIAM Journal on Control, Nov. 1970.

H2. Ho, Y. C. Rand Corp. Report, RM 3329-PR, 1963.

J1. Jazwinski, A. H. Stochastic Process and Filtering Theory, Academic Press, N. Y. 1970.

K1. Kalman, R. E. Trans. ASME J. Basic Engr. Vol. 82D, March, 1960.

K2. Kalman, R. E., Falb, P. L., Arbib, M. A. Topics in Mathematical System Theory, McGraw-Hill Co., N. Y., 1969.

K3. Kalman, R. E., Ho, Y. C., Narenda, K. S. Contributions to Differential Equation, Vol. 1, No. 2, 1962.

K4. Kushner, H. J. Stochastic Stability and Control, Academic Press, N. Y., 1967.

L1. Lapidus, L., Berger, A. AIChE Journal, March, 1969.

L2. Larson, R. E. Report on Estimation, Contract 14-03-79378, Bonneville Power Administration, Portland, Oregon, Nov. 1968.

L3. Lee, R.C.K. Optimal Estimation, Identification and Control, MIT Press, Mass. 1964.

L4. Lo, J. T. IEEE Symposium on Adaptive Process, Decision and Control, University of Texas, Austin, Dec. 1970.

L5. Lobry, C. SIAM Journal on Control, Nov. 1970.

M1. McGuire, M. L., Leucke, R. I&EC Fundamental 6, 1967.

M2. Meditch, J. S. Stochastic Optimal Linear Estimation and Control, McGraw-Hill Co., N. Y. 1969.

M3. Mehra, R. K. IEEE Symposium on Adaptive Process, Decision and Control, University of Texas, Austin, December 1970.

O1. Ortega, J. N., Rheinboldt, W. C. Iterative Solution of Non-linear Equations in Several Variables, Academic Press, N. Y., 1970.

P1. Payne, A. Advances in Control System, Academic Press, N. Y., 1969.

P2. Pine, S., Wolf, H., Bailie, A., Mohan, J. NASA Report, Contract NAS 5-2535, Analytical Mechanics Associates, Westbury, N. Y., 1964.

P3. Price, C. F., Deyst, J. J. IEEE Trans. Automatic Control, Jan. 1969.

S1. Sage, A. P. IEEE Symposium on Adaptive Process, Decision and Control, University of Texas, Austin, December 1970.

S2. Sage, A. P. Optimum System Control, Prentice Hall, Englewood, Cliff, N. Y., 1968.

S3. Sage, A. P., Melsa, J. L. Estimation Theory with Application to Communication and Control, McGraw-Hill, N. Y., 1971.

S4. Schmidt, S. F. Advances in Control System, Academic Press, N. Y., 1966.

S5. Schmidt, S. F. System Control Workshop, University of Michigan, 1968.

S6. Schwartz, L. Advances in Control System, Academic Press, 1969.

S7. Seinfeld, J. H., Gavalas, G. Chemical Engr. Science, Jan. 1969.

S8. Seinfeld, J. H. AIChE Journal, Nov. 1970.

S9. Seinfeld, J. H. Internal Journal on Control, Jan. 1969.

S10. Seinfeld, J. H., Lapidus, L., Huang, M. I&EC Fundamental, May, 1969.

S11. Sorenson, H. W. IEEE Symposium on Adaptive Process, Decision and Control, University of Texas, Austin, Dec. 1970.

W1. Wells, C. JACC-1969, Colorado (Joint Automatic Control Conference).

W2. Wells, C., Larson, R. E. JACC-1970, Georgia, 1970.

W3. West, H. H., McGuire, M. L. I&EC Fundamental, May, 1969.

W4. Wilde, D. Optimum Seeking Method, Prentice Hall, 1964.

W5. Willougby, R. A., Lininger, W. SIAM Journal on Numerical Analysis, March, 1970.

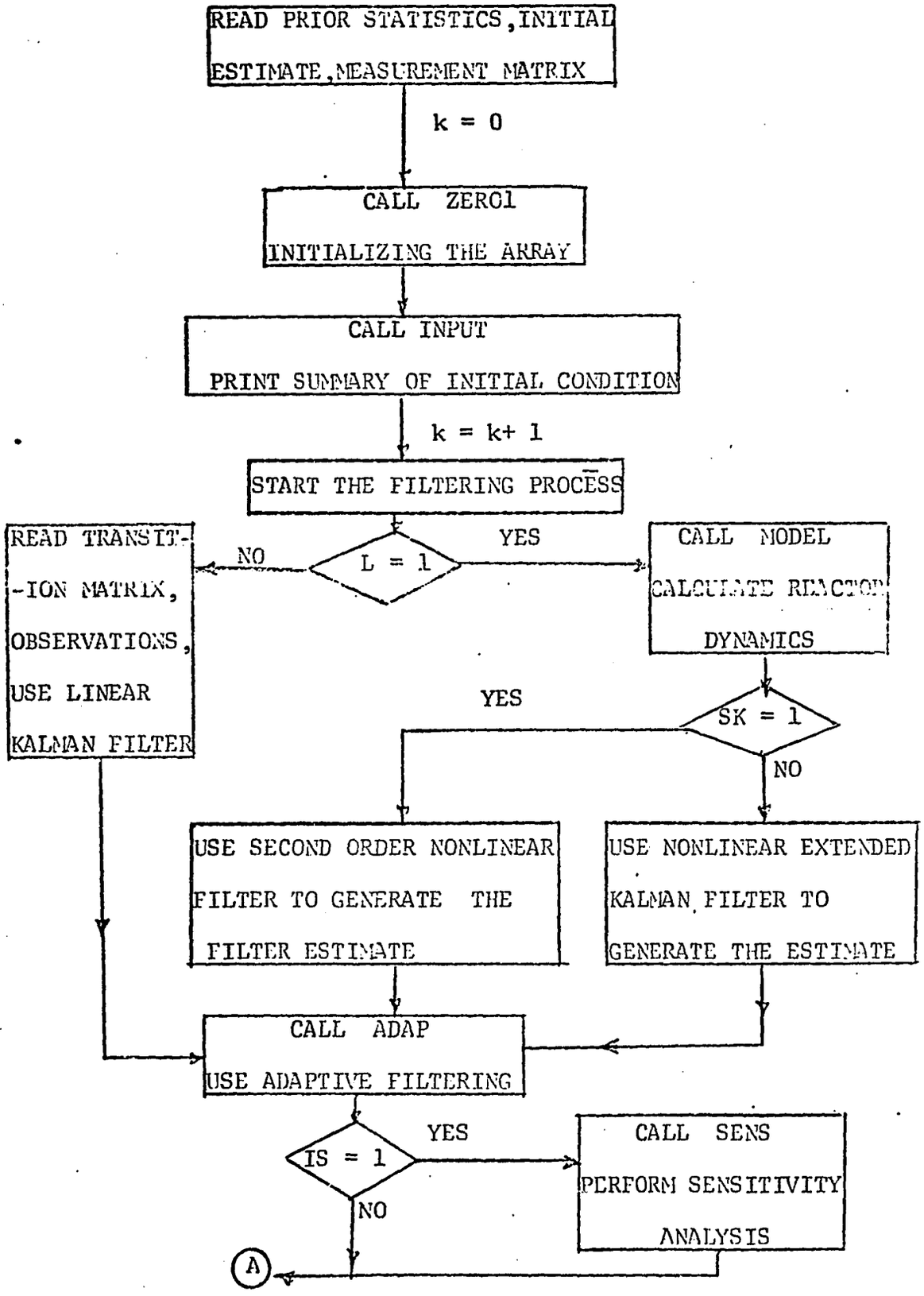
W6. Wolf, H. NASA Report, Analytical Mechanics Associates, Westbury, N. Y., 1968.

APPENDIX A : PROGRAM FLOW CHART

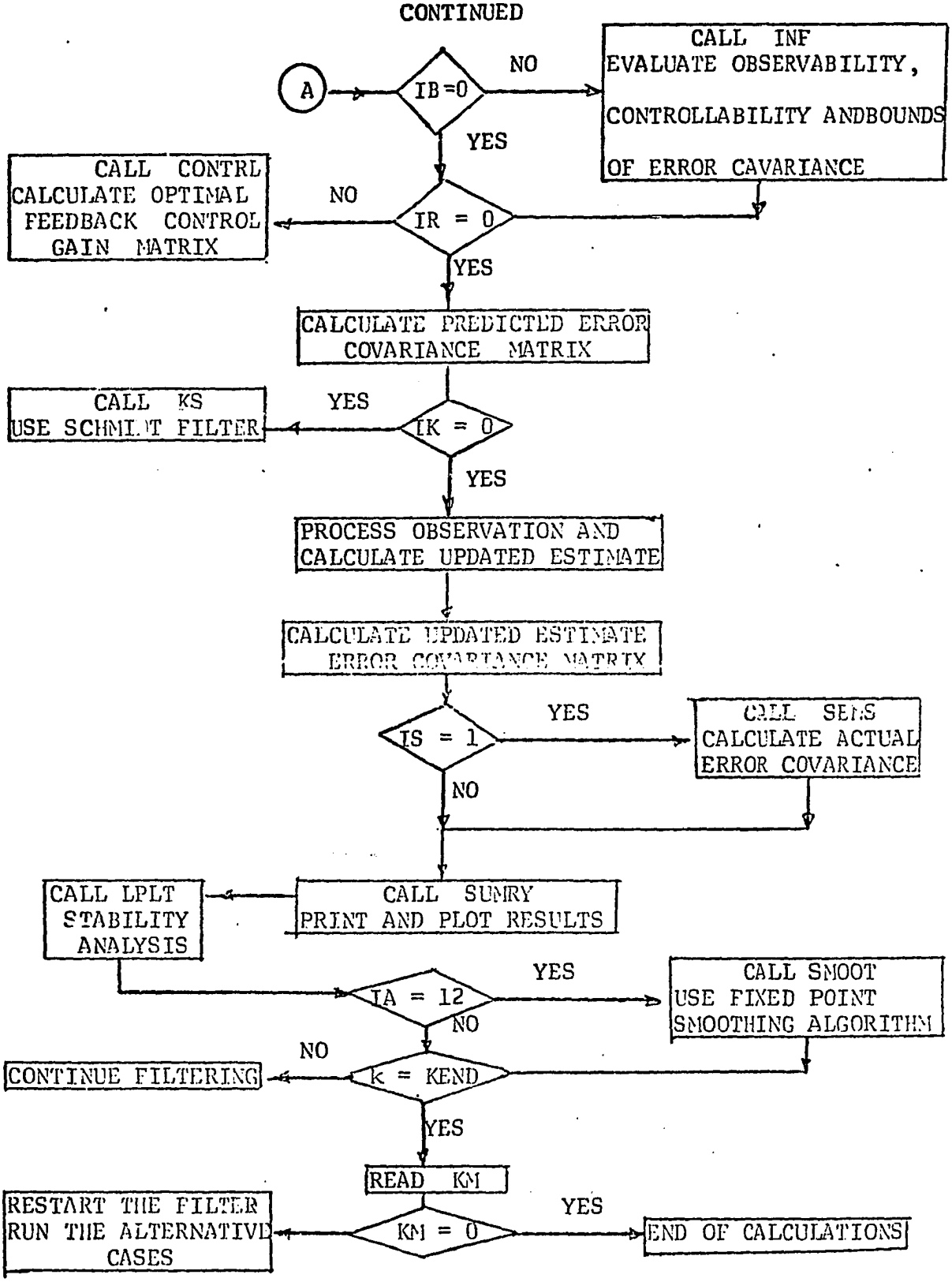
SUBROUTINE MAP FOR THE
COMPUTER PROGRAM

SUBROUTINE NAME	ORIGIN	STORAGE LENGTH	PURPOSE
MAIN	00	3842	CONTROL OF OVERAL CALCULATION LOGIC
SMOOT	3F4F	10AC	FIXED POINT SMOOTHING CALCULATION
SUMRY	4DD8	27A4	PRINT AND PLOT FILTERING RESULTS
MODEL	7350	10F8	GENERATING DYNAMICS AND OBSERVATION
FUN	8478	3AE	GENERATING REACTOR DYNAMICS
TRAN	8E74E	1C0	PROCESS TRANSITION MATRIX
INF	8E0E	4D3	UPPER AND LOWER BOUND CALCULATIONS
RKK	9A70	81D	EULER AND RUNGE-KUTTA INTEGRATION
DPLT	4EEC	27E	PLOT ROUTINE I
RANK	A8D0	5A0	RANK EVALUATION OF MATRIX
MATX	B2D0	F04	MATRIX INVERSION BY SHUR RELATION
ITX	C2C8	520	MATRIX INVERSION BY ITERATION
FUNA	C7E8	2F0	REACTOR DYNAMICS OF THE ACTUAL MODEL
FUNC	C9A0	2F8	REACTOR DYNAMICS OF FILTER MODEL
KS	CC98	E81	KALMAN -SCHMIDT FILTER
GAUS	D7A0	1F3	GENERATING GAUSSIAN NOISE
RANDU	D9F7	13C	RANDOM NUMBER GENERATOR
LOC	DB30	2A1	LOCATE MATRIX ELEMENTS
DPLOT	DDC8	1DD	PLOT ROUTINE II
SPLOT	DFAE	20C	PLOT ROUTINE III
SP	E1B7	7B4	SYMMETRIZING THE ERROR COVARIANCE
SENS	E93D	240C	SENSITIVITY ANALYSIS CALCULATION
ADAP	ECC4C	53E	ADAPTIVE FILTERING TECHNIQUE
ZERO	F19CE	13A	INITIALIZING THE ARRAY
RK4	F1708	8F1	RUNGE-KUTTA 4 th ORDER INTEGRATION
INPUT	F1DF8	61D	SUMMARY OF INPUT CONDITIONS
ZERO1	F25C8	13E	INITIALING THE ARRAY
CONTRL	F2708	D49	OPTINAL CONTROL GAIN CALCULATION
MULT	F3551	4D4	MATRIX PRODUCT EVALUATIONS
PRT	F3993	2BA	COPY PART OF THE MATRIX
OUTPT	F3C78	78C	PLOTING OUTPUT GENERATOR
STAB			STABILITY ANALYSIS
OVWT		6CE	OVERWEIGHT &STOCASTIC APPROX.

FLOW CHART : MAIN PROGRAM

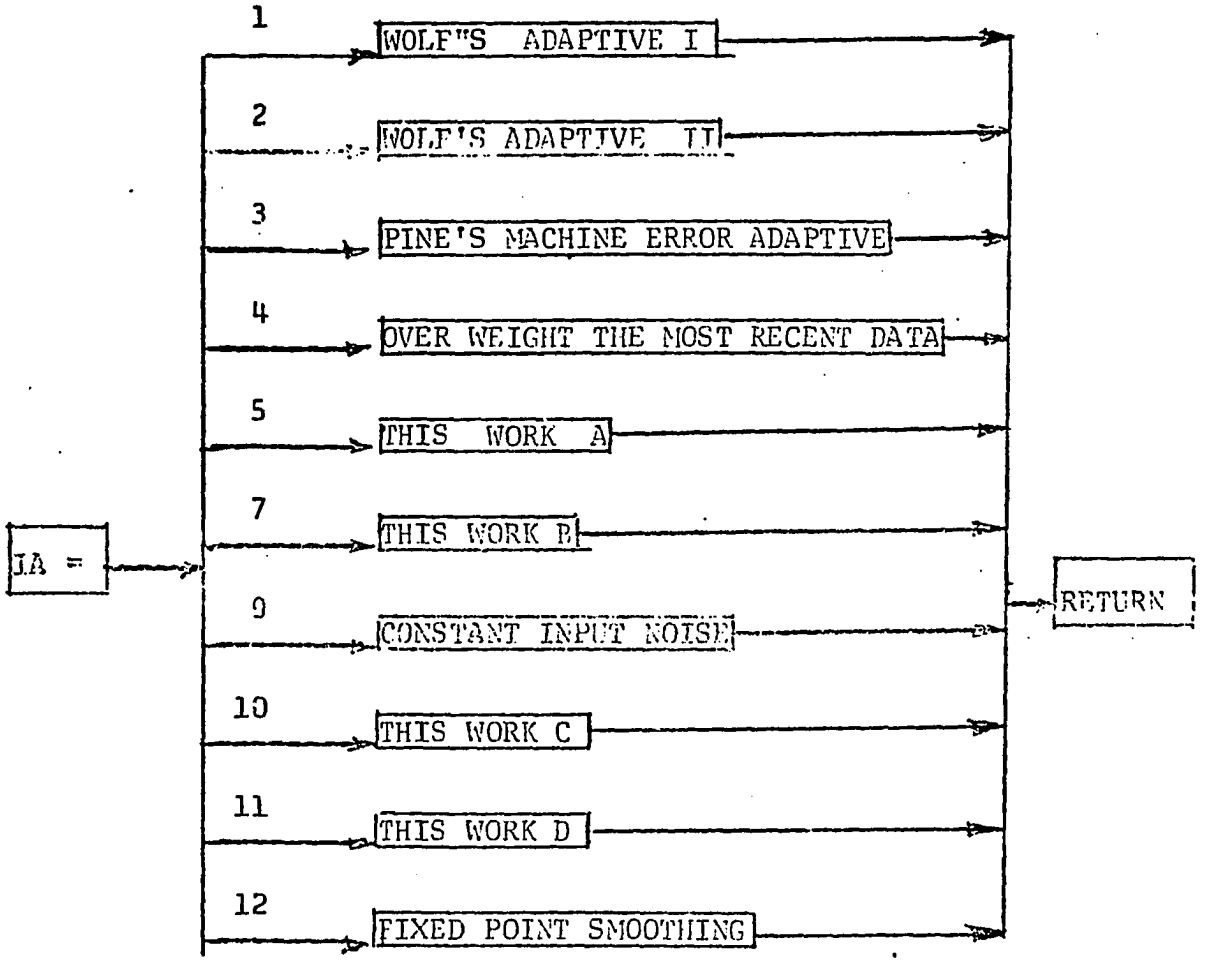


CONTINUED


A
IB=0
NO
CALL INF
EVALUATE OBSERVABILITY,
CONTROLLABILITY ANDBOUNDS
OF ERROR CAVARIANCE
YES
CALL CONTRL
CALCULATE OPTIMAL
FEEDBACK CONTROL
GAIN MATRIX
NO
IR = 0
YES
CALCULATE PREDICTED ERROR
COVARIANCE MATRIX
CALL KS
USE SCHMIDT FILTER
YES
IK = 0
YES
PROCESS OBSERVATION AND
CALCULATE UPDATED ESTIMATE
CALCULATE UPDATED ESTIMATE
ERROR COVARIANCE MATRIX
IS = 1
YES
CALL SENS
CALCULATE ACTUAL
ERROR COVARIANCE
NO
CALL LPLT
STABILITY
ANALYSIS
CALL SUMRY
PRINT AND PLOT RESULTS
IA = 12
YES
CALL SMOOT
USE FIXED POINT
SMOOTHING ALGORITHM
NO
NO
CONTINUE FILTERING
k = KEND
YES
READ KM
RESTART THE FILTER
RUN THE ALTERNATIVE
CASES
KM = 0
YES
END OF CALCULATIONS

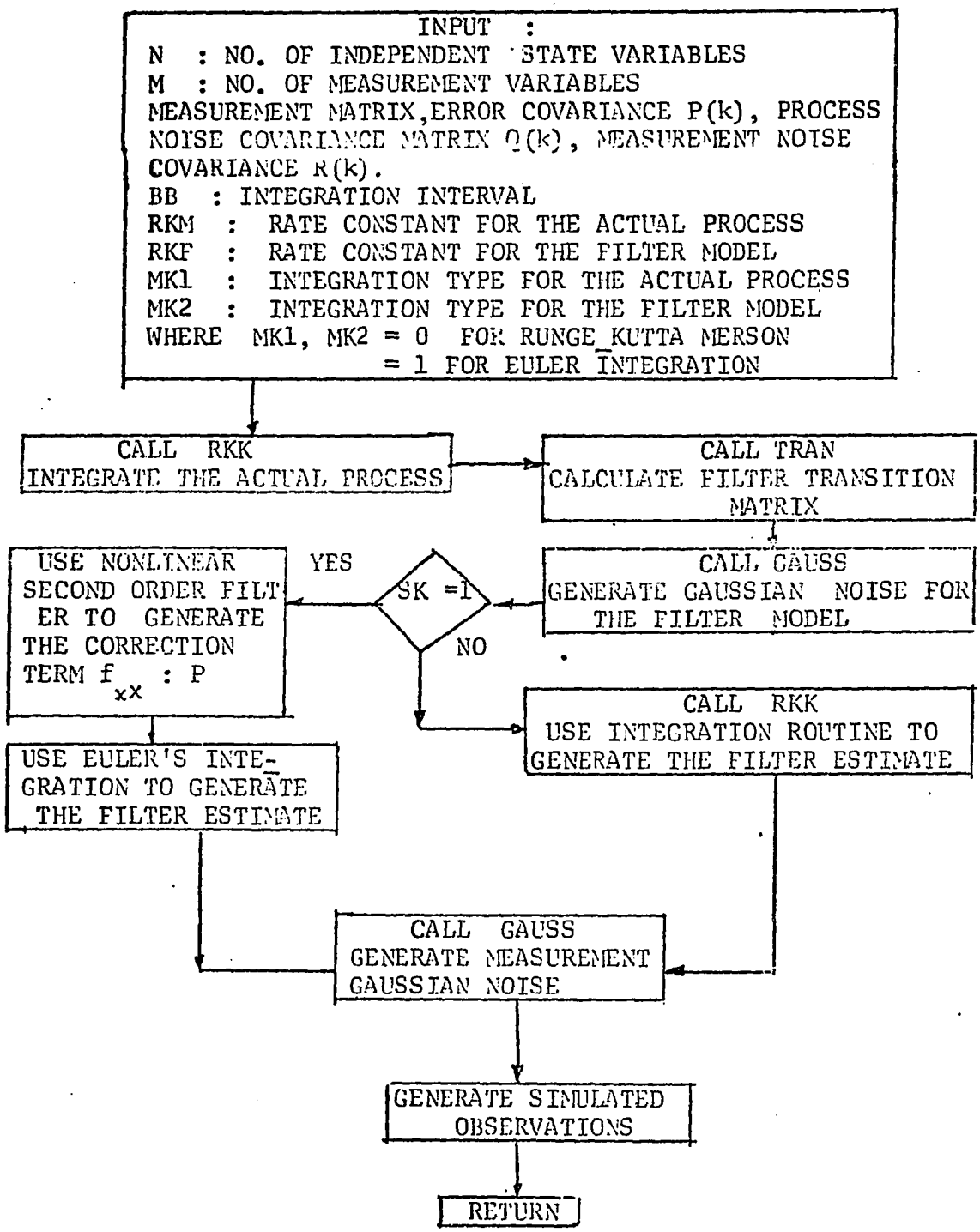
SUBROUTINE ADAP

ADAPTIVE FILTERS (10)



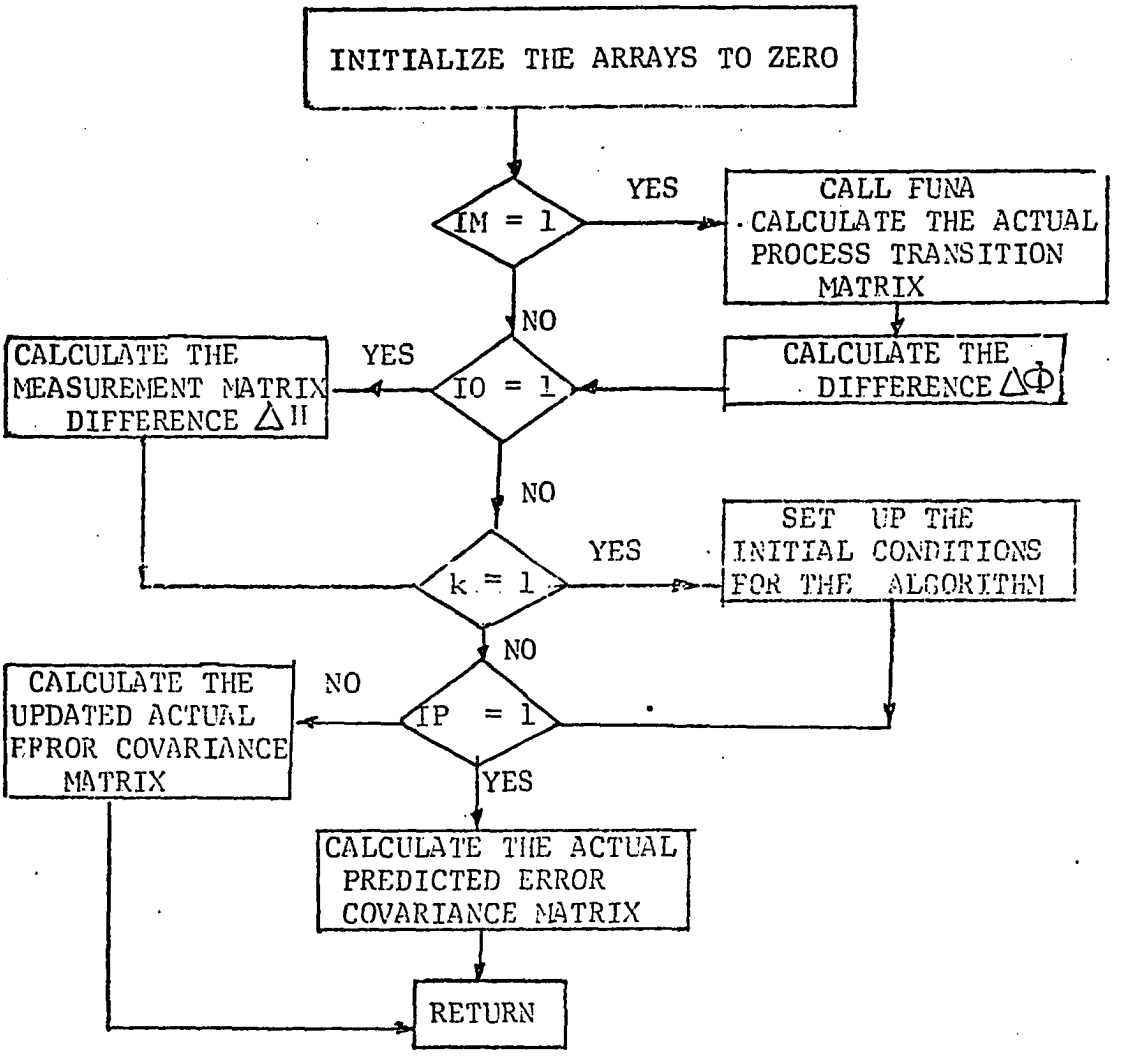
SUBROUTINE MODEL

CALCULATE REACTOR DYNAMICS AND TO
GENERATE SIMULATED OBSERVATION



SUBROUTINE SENS

GENERAL SENSITIVITY ANALYSIS



APPENDIX B. SAMPLE COMPUTER OUTPUT

APPROX IMATE THE 4ht ORDER MODEL BY A SECOND ORDER USING HUANG'S II
(ITERATED FIXED POINT SMOOTHING)

```
SAMPLE PERICD     1


  EULER INTEGRATION
 Y =      0.10016   DERY =      0.18821
 Y =      C.C683C   DERY =     -C.C6384
 Y =      0.01012   DERY =      0.03736
 Y =      0.CCC68   DERY =      0.C0386
 T=       0.5000   INTERVAL =      C.C5CCC
XE1=    0.105619   C.048251
SIMULATED CBSERVATICN  0.1C6950
FILTER ESTIMATE     0.1056     C.C482
KALMAN FILTER GAIN =  0.563966  0.C40883
NEW FILTER DATA  C.106370  C.C48306
NEW ERROR COVARIANCE MATRIX    0.00564C  0.C00409   C.000409   0.001704
SAMPLE PERICD     2


  EULER INTEGRATION
 Y =      0.18422   DERY =      C.15C1C
 Y =      0.03733   DERY =     -0.C5959
 Y =      0.03528   DERY =      0.05824
 Y =      0.CC52C   DERY =      0.C137C
 T=       1.0000   INTERVAL =      0.05CCC
XE1=    0.174C83   C.C23C80
SIMULATED CBSERVATION  0.188746
FILTER ESTIMATE     0.1741     C.0231
KALMAN FILTER GAIN =  C.539857  C.C43457
NEW FILTER DATA  0.181999  0.C23718
NEW ERROR COVARIANCE MATRIX    0.CC5399  C.000435   0.000435   0.001723
```

APPROXIMATE THE 4 th ORDER MODEL BY A SECOND ORDER USING HUANG'S II
(ITERATED FIXED POINT SMOOTHING)

```
SAMPLE PERIOD     3


  EULER INTEGRATION
 Y =      0.24750   DERY =      0.10736
 Y =      0.00967   DERY =     -0.05148
 Y =      0.06662   DERY =      0.06456
 Y =      0.01497   DERY =      0.02412
 T=       1.5000   INTERVAL =      0.05000
XE1=    0.229857 -0.000384
SIMULATED OBSERVATION  0.229742
FILTER ESTIMATE     0.2299    -0.0004
KALMAN FILTER GAIN =  0.523127  0.053108
NEW FILTER DATA  0.229797 -0.000390
NEW ERROR COVARIANCE MATRIX    0.005231  0.000531  0.000531  0.001685
SAMPLE PERIOD     4


  EULER INTEGRATION
 Y =      0.29017   DERY =      0.06830
 Y =     -0.01348   DERY =     -0.04109
 Y =      0.09844   DERY =      0.06193
 Y =      0.02942   DERY =      0.03238
 T=       2.0000   INTERVAL =      0.05000
XE1=    0.252480 -0.020471
SIMULATED OBSERVATION  0.288108
FILTER ESTIMATE     0.2525    -0.0205
KALMAN FILTER GAIN =  0.513054  0.059275
NEW FILTER DATA  0.270759 -0.010359
   SMOOTHED ESTIMATE PERIOD    0
    SMOOTHED ESTIMATE    -0.0017     0.1041     0.0000     0.0000
 SMOOTHED P =    0.00265   -0.00053   -0.00058   0.00092
NEW ERROR COVARIANCE MATRIX    0.005131  0.000593  0.000593  0.001637
```

APPROXIMATE THE 4 th ORDER MODEL BY A SECOND ORDER USING HUANG'S II
(ITERATED FIXED POINT SMOOTHING)

SAMPLE PERICD 1

EULER INTEGRATICN
Y = 0.1C016 DERY = 0.18831
Y = 0.C6830 DERY = -C.C6384
Y = 0.01012 DERY = 0.03736
Y = 0.CCC68 DERY = 0.C0386
T= 0.5000 INTERVAL = 0.05000
XF1= 0.107284 C.072271
SIMULATED CBSERVATICN 0.1C6950
FILTER ESTIMATE 0.1073 C.0723
KALMAN FILTER GAIN = 0.561190 0.037932
NEW FILTER DATA C.107097 0.072258
NEW ERRCR COVARIANCE MATRIX 0.006612 0.000379 C.0CC379 C.C017C1
SAMPLE PERIOD 2

EULER INTEGRATICN
Y = 0.18422 DERY = 0.15010
Y = 0.03733 DERY = -0.C595[illegible]
Y = 0.03528 DERY = 0.05824
Y = 0.0052C DERY = 0.C137C
T= 1.0000 INTERVAL = 0.05000
XF1= 0.2C2814 C.038985
SIMULATED CBSERVATICN 0.188746
FILTER ESTIMATE 0.2029 0.C3[illegible]0
KALMAN FILTER GAIN = C.546638 0.041C19
NEW FILTER DATA 0.195124 0.C38408
NEW ERRCR COVARIANCE MATRIX 0.005466 0.000410 C.00041C C.CC172C

APPROXIMATE THE 4 the ORDER MODEL BY A SECOND ORDER USING HUANG'S II
(ITERATED FIXED POINT SMOOTHING)

```
SAMPLE PERIOD     3


  EULER INTEGRATION
 Y =      0.24756   DERY =      0.10736
 Y =      0.00967   DERY =     -0.05348
 Y =      0.06662   DERY =      0.06456
 Y =      0.01497   DERY =      0.02412
 T=      1.5000   INTERVAL =      0.05000
XE1=    0.202204 -0.003017
SIMULATED OBSERVATION  0.229742
FILTER ESTIMATE      0.2022    -0.0030
KALMAN FILTER GAIN =  0.615066  0.085908
NEW FILTER DATA  0.219141 -0.000651
NEW ERROR COVARIANCE MATRIX    0.006151   0.000859   0.000859   0.001858
SAMPLE PERIOD     4


  EULER INTEGRATION
 Y =      0.29017   DERY =      0.06630
 Y =     -0.01348   DERY =     -0.04195
 Y =      0.02944   DERY =      0.06193
 Y =      0.02942   DERY =      0.03239
 T=      2.0000   INTERVAL =      0.05000
XE1=    0.241149 -0.019787
SIMULATED OBSERVATION  0.288108
FILTER ESTIMATE      0.2411    -0.0198
KALMAN FILTER GAIN =  0.558103  0.064650
NEW FILTER DATA  0.267357 -0.016751
   SMOOTHED ESTIMATE PERIOD    0
    SMOOTHED ESTIMATE     0.0244     0.0733     0.0000     0.0000
 SMOOTHED P =    0.00553   -0.00257   -0.00257    0.00267
NEW ERROR COVARIANCE MATRIX    0.005581   0.000646   0.000646   0.001665
```

APPROXIMATE THE 4 th ORDER MODEL BY A SECOND ORDER USING HUANG'S II
(ITERATED FIXED POINT SMOOTHING)

```
TOTAL NO OF PERIODS   20SAMPLE PERIOD IN SEC.      0.50
SAMPLE PERIOD    1


  EULER INTEGRATION
 Y =     0.10016   DERY =     0.18831
 Y =     0.06830   DERY =    -0.06384
 Y =     0.01012   DERY =     0.03736
 Y =     0.00068   DERY =     0.00386
 T=      0.5000   INTERVAL =     0.05000
XE1=    0.002468   0.001415
SIMULATED OBSERVATION  0.106950
FILTER ESTIMATE     0.0025     0.0014
KALMAN FILTER GAIN =  0.707816  0.170928
NEW FILTER DATA  0.076422  0.019274
NEW ERROR COVARIANCE MATRIX    0.007078  0.001709  0.001709  0.005000
SAMPLE PERIOD    2


  EULER INTEGRATION
 Y =     0.18422   DERY =     0.15010
 Y =     0.03733   DERY =    -0.05959
 Y =     0.03528   DERY =     0.05824
 Y =     0.00520   DERY =     0.01370
 T=      1.0000   INTERVAL =     0.05000
XE1=    0.110035   0.006567
SIMULATED OBSERVATION  0.188746
FILTER ESTIMATE     0.1100     0.0066
KALMAN FILTER GAIN =  0.669729  0.135183
NEW FILTER DATA  0.162750  0.017207
NEW ERROR COVARIANCE MATRIX    0.004697  0.001352  0.001352  0.002608
```

APPROXIMATE THE 4ht ORDER MODEL BY A SECOND ORDER USING HUANG'S II
(ITERATED FIXED POINT SMOOTHING)

```
SAMPLE PERIOD     ?


  EULER INTEGRATION
 Y =      0.24750   DERY =      0.10736
 Y =      0.00967   DERY =     -0.05148
 Y =      0.06662   DERY =      0.06156
 Y =      0.01497   DERY =      0.02412
 T=       1.5000   INTERVAL =      0.05000
XE1=    0.261896   0.008141
SIMULATED OBSERVATION  0.229742
FILTER ESTIMATE     0.2619     0.0081
KALMAN FILTER GAIN =  0.528665   0.051687
NEW FILTER DATA  0.244892   0.003480
NEW ERROR COVARIANCE MATRIX    0.005287   0.000517   0.000517   0.001678
SAMPLE PERIOD     4


  EULER INTEGRATION
 Y =      0.29017   DERY =      0.06030
 Y =     -0.01348   DERY =     -0.04199
 Y =      0.00844   DERY =      0.06193
 Y =      0.02942   DERY =      0.03238
 T=       2.0000   INTERVAL =      0.05000
XE1=    0.276633 -0.017154
SIMULATED OBSERVATION  0.288103
FILTER ESTIMATE     0.2766    -0.0172
KALMAN FILTER GAIN =  0.515769   0.058591
NEW FILTER DATA  0.282552 -0.010482
  SMOOTHED ESTIMATE PERIOD    0
   SMOOTHED ESTIMATE    -0.0067     0.0966     0.0000     0.0000
 SMOOTHED P =    0.00245   -0.00056   -0.00056    0.00091
NEW ERROR COVARIANCE MATRIX    0.005158   0.000586   0.000586   0.001628
```

APPROXIMATE THE 4 th ORDER MODEL BY A SECOND ORDER
USING HUANG'S IV (OVERWEIGHT THE DATA BY STOCASTIC APPROX.)

```
TOTAL NO OF PERIODS   20SAMPLE PERIOD IN SEC.      0.50
SAMPLE PERIOD    1


  EULER INTEGRATION
 Y =     0.10016  DERY =     0.18831
 Y =     0.06830  DERY =    -0.06394
 Y =     0.01012  DERY =     0.03736
 Y =     0.00069  DERY =     0.00386
 T=      0.5000  INTERVAL =     0.05000
XE1=    0.002468 -0.000585
SIMULATED OBSERVATION  0.106950
FILTER ESTIMATE     0.0025    -0.0006
KALMAN FILTER GAIN =  0.707816  0.170928
NEW FILTER DATA  0.106950  0.017274
NEW ERROR COVARIANCE MATRIX    0.007809  0.001709  0.001709
SAMPLE PERIOD    2                0.005000


  EULER INTEGRATION
 Y =     0.18422  DERY =     0.15010
 Y =     0.03733  DERY =    -0.05959
 Y =     0.03528  DERY =     0.05824
 Y =     0.00520  DERY =     0.01370
 T=      1.0000  INTERVAL =     0.05000
XE1=    0.142021  0.001264
SIMULATED OBSERVATION  0.188746
FILTER ESTIMATE     0.1420     0.0013
KALMAN FILTER GAIN =  0.680083  0.130745
NEW FILTER DATA  0.188746  0.007374
NEW ERROR COVARIANCE MATRIX    0.007601  0.001307  0.001307
SAMPLE PERIOD    3              0.002603


  EULER INTEGRATION
 Y =     0.24750  DERY =     0.10736
 Y =     0.00967  DERY =    -0.05148
 Y =     0.06662  DERY =     0.06456
 Y =     0.01497  DERY =     0.02412
 T=      1.5000  INTERVAL =     0.05000
XE1=    0.218565 -0.012360
SIMULATED OBSERVATION  0.229742
FILTER ESTIMATE     0.2186    -0.0124
KALMAN FILTER GAIN =  0.624769  0.080840
NEW FILTER DATA  0.226946 -0.011456
NEW ERROR COVARIANCE MATRIX    0.006482  0.000808  0.000808
                                0.001881
```

FILTER DECOMPOSITION SAMPLE COMPUTER OUTPUT

20 % rate constant error

SUBSYSTEM 1 : X1 ,X4 ; SUBSYSTEM 2 : X2, X3

```
SAMPLE PERIOD     2


SIMULATED OBSERVATION  0.198051  0.038281  0.037162
FILTER ESTIMATE     0.1199     0.0176
   SIMULATED MEASUREMENT FOR SUBSYSTEM     1 IS
        0.19805    0.03716
   FILTER GAIN FOR SUBSYSTEM     1
  0.64880    0.16530    0.16530    0.08025
  UPDATED ESTIMATE FOR THE SUBSYSTEM      1  IS
  0.17002    0.05328
NEW ERROR COVARIANCE MATRIX    0.006488  0.001653  0.001653  0.000802
  TIME :      0  MINUTE  0.7 SECONDS
SAMPLE PERIOD     2


FILTER ESTIMATE     0.1199     0.0176
   SIMULATED MEASUREMENT FOR SUBSYSTEM     2 IS
        0.03828
   FILTER GAIN FOR SUBSYSTEM     2
  0.17935    0.13614
  UPDATED ESTIMATE FOR THE SUBSYSTEM      2  IS
  0.02130    0.00650
NEW ERROR COVARIANCE MATRIX    0.001793  0.001361  0.001361  0.001301
  TIME :      0  MINUTE  0.0 SECONDS
```